

# Pharma SUG SDE Japan 2023

**Innovative Approaches for  
Clinical Trial Data Analysis and RWD Analysis :**

**SAS® Health Clinical Acceleration**

**SAS® Health Cohort Builder**

SAS Institute Japan

Toshiaki Habu • William Kuan



# **SAS® Life Science Analytics Framework (LSAF) SAS® Health Clinical Acceleration (HCA)**

# History of SDD/LSAF/HCA

## Previous Solution

SAS ®  
Drug Development  
(2005~2014)

SAS execution environment for regulatory requirements

- SAS9 based
- Java-based
- On-premise provision (Cloud provision in the latter half)
- Programming
- Comply with regulatory requirements (Part 11, etc.)
- Workspace functionality



## Current Solution

SAS ® Life Science Analytics Framework (2015~)

### Managed Cloud Service

- SAS9 based
- HTML5
- Provided in Cloud
- Workflow Function
- Enhanced Macro API

SAS® Health Clinical Acceleration (2023~)

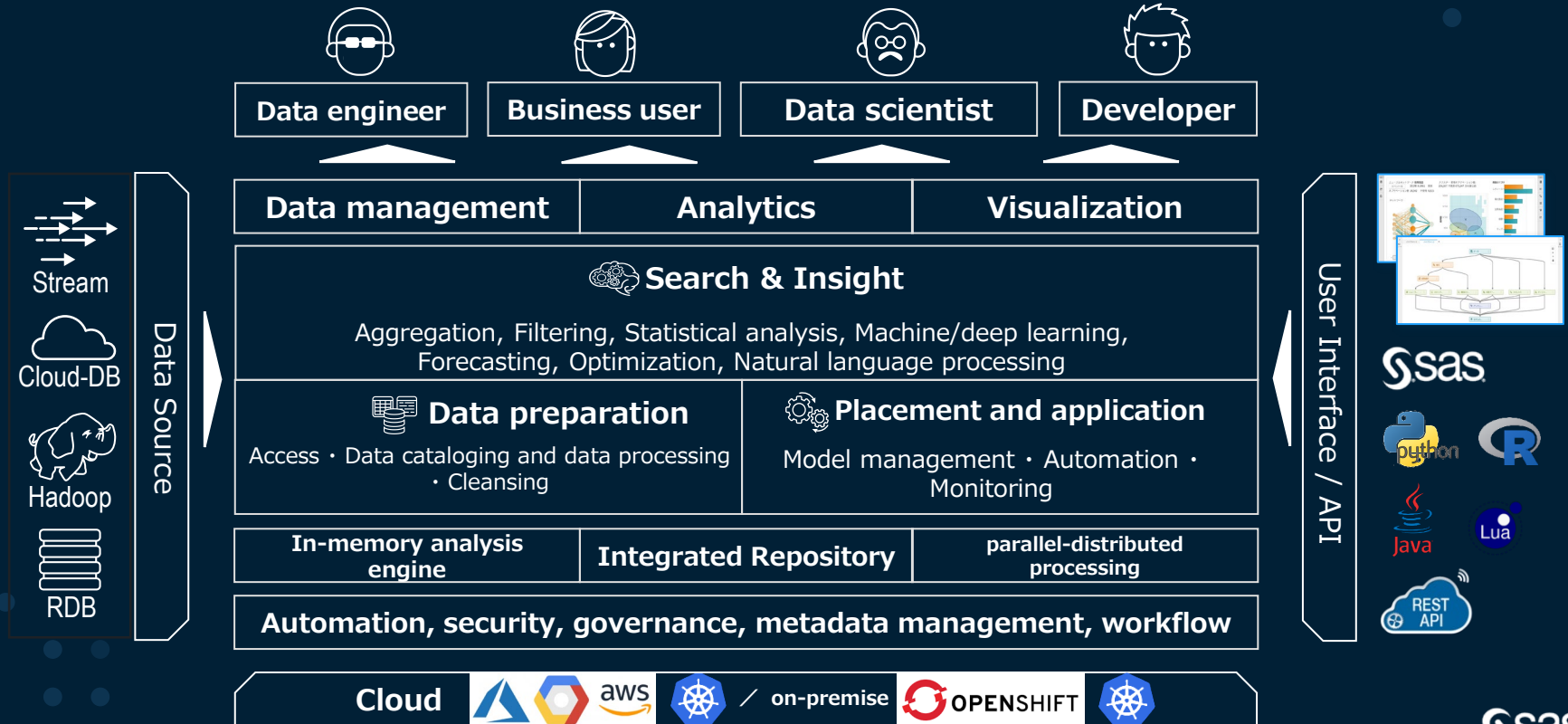
### Move to Viya-based from SAS9-based

- Engine modernization (based on Viya)
- Modularization of functions: further expansion planned

For about 18 years since SDD launch, it has been used by many companies, including global pharmaceutical companies, as a repository and SAS execution environment that complies with FDA CFR 21 Part 11 and other regulatory requirements.

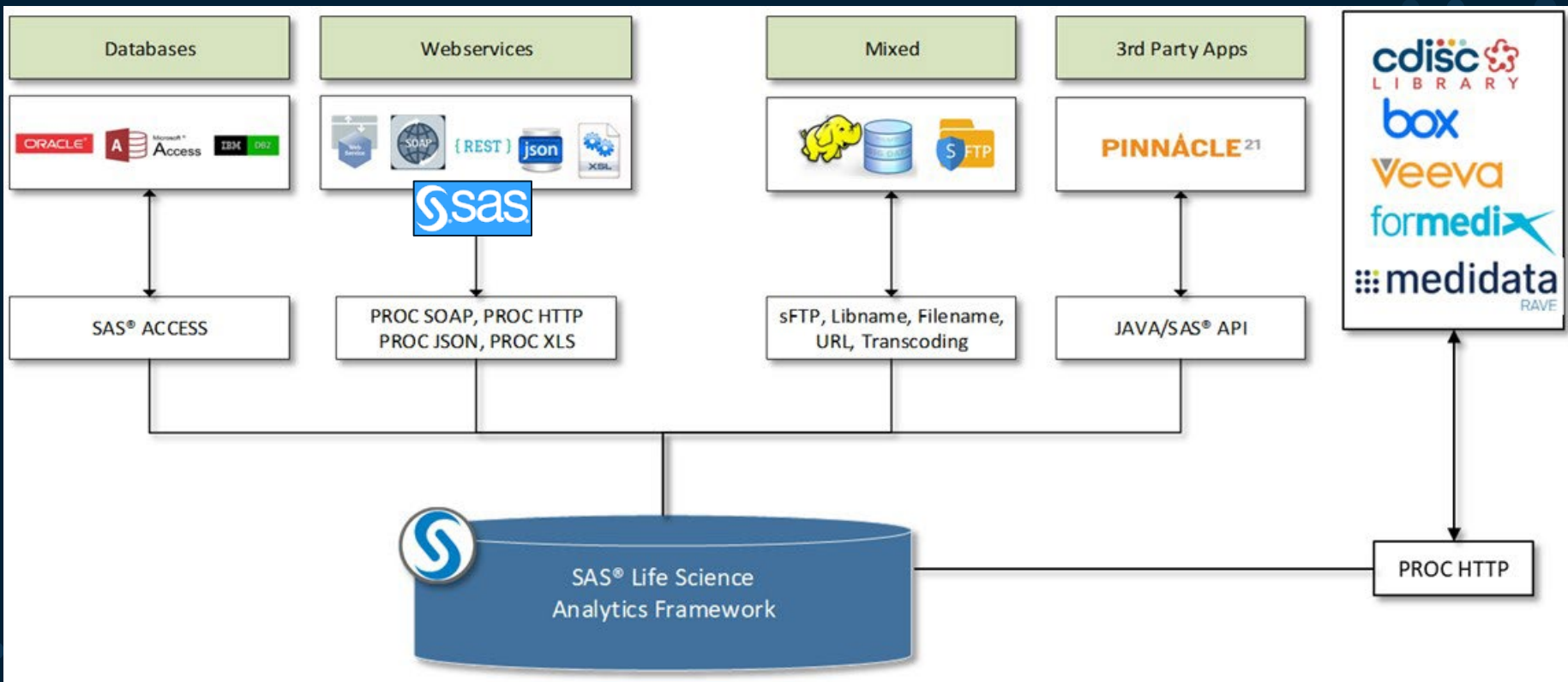
# SAS Viya

Provides the entire suite of functions and software necessary to promote data utilization on a single platform



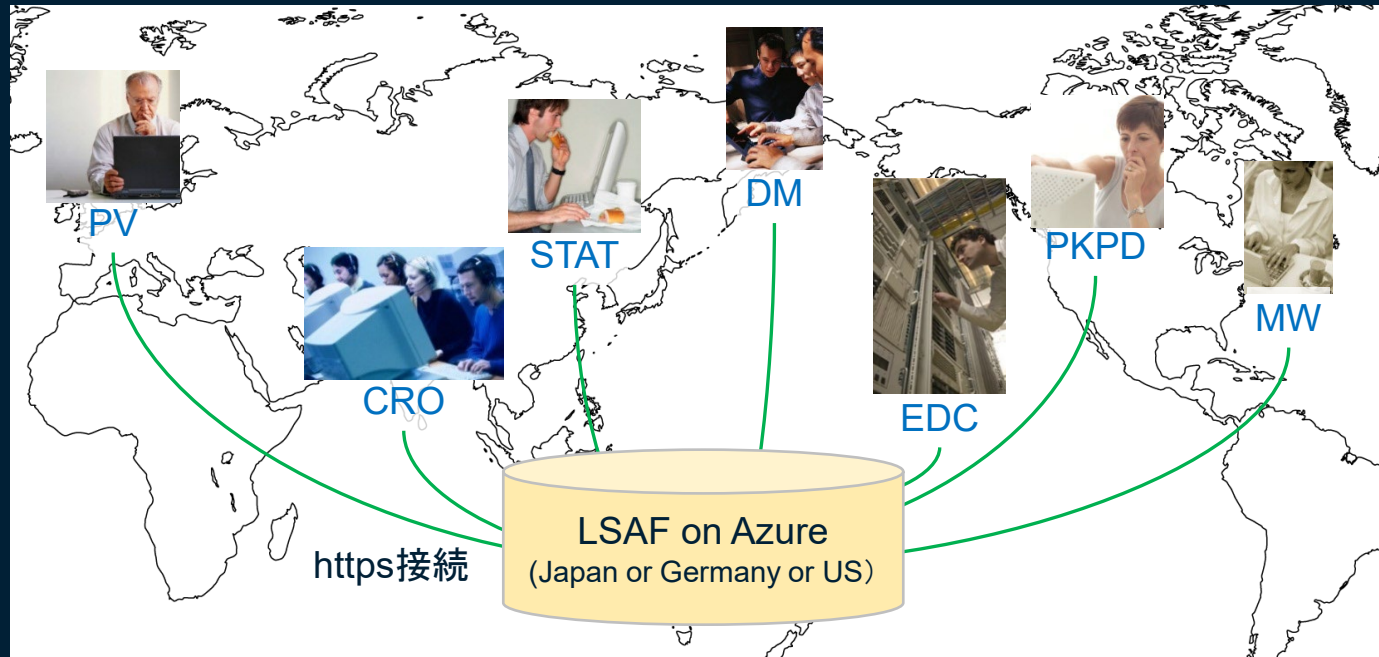
# SAS<sup>®</sup> Life Science Analytics Framework (LSAF) Overview

# Flexible Integration Capabilities



# Support collaborations within the organization and with external partners

Each user can connect to the LSAF server managed by SAS via https from a browser to develop and run programs and exchange SAS data sets and other files with other personnel, regions, and departments.



# Overview of LSAF (1/2)

LSAF is a solution that provides repository functions for clinical data and data standards, as well as program development and execution functions such as statistical analysis, in a secure cloud format for life science companies.

- **Server maintained and managed by SAS. Available via web browser (<https>)**

- Provided as a cloud service, reducing companies' burden on system construction, maintenance, and validation.
- Web-based, so it can be used without concern for the PC environment of each user.

- **Provides a clinical data repository (CDR) and statistical computing environment.**

- No restrictions on supported file formats, including SAS data sets, programs, and other SAS-related files, as well as CSV, Word, Excel, PDF, and other formats.
- Groups and roles (combinations of permission and privileges required for each role, such as programmer, medical writer, etc.) can be freely defined, and security can be efficiently configured in terms of access and privileges.
- Provides a personal area (development area) and a shared area (production area) that can be used for SAS program development and production execution.

Key Point!!

- **Provides regulatory compliance capabilities**

- Audit trails, implementation of version control, and traceability functions enable efficient electronic recording/electronic signature and GxP compliance.
- FDA 21 CFR Part 11 compliant



# Overview of LSAF (2/2)

- **Easy integration with other applications.**

- Extensive experience with EDC integration such as Medidata Rave and InForm, and Pinnacle 21 Enterprise/Community for the purpose of improving operational efficiency and security through automation.
- Easy deployment of machine learning, RWD, etc. by leveraging SAS Viya, Python and R, etc.

- **Workflow Functions**

- Workflow functionality improves operational efficiency and visualizes progress.
- Comply with industry standards (Business Process Model and Notation 2.0) for workflow design enables user-centered operations.

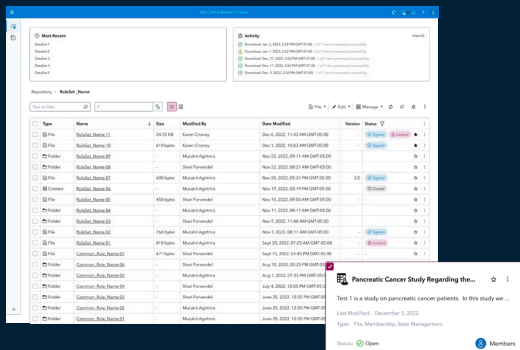
# SAS® Health Clinical Acceleration

# SAS® Health Clinical Acceleration at a Glance

## Overview

SAS® Health Clinical Acceleration (HCA) provides a modular, open, cloud-native content repository for managing, analyzing, reporting, and reviewing clinical research information. It is the successor to LSAF on Viya.

SAS HCA improves the way people work in a qualified environment and enables life science organizations to maintain data integrity, make informed business decisions, effectively assess the safety and efficacy of research compounds, collaborate across trials, phases, therapeutic areas, and ultimately streamline the drug development process and accelerate their submissions to the FDA.



**Platform:** Viya 4

## Value Propositions:

- Faster Clinical Analysis and Submission – Improved UX
- Regulated Clinical Trial Operations – Single Source of Truth
- Viya Enabled Programming, Analytics & Automation
- Open Integration – Bring your own tool

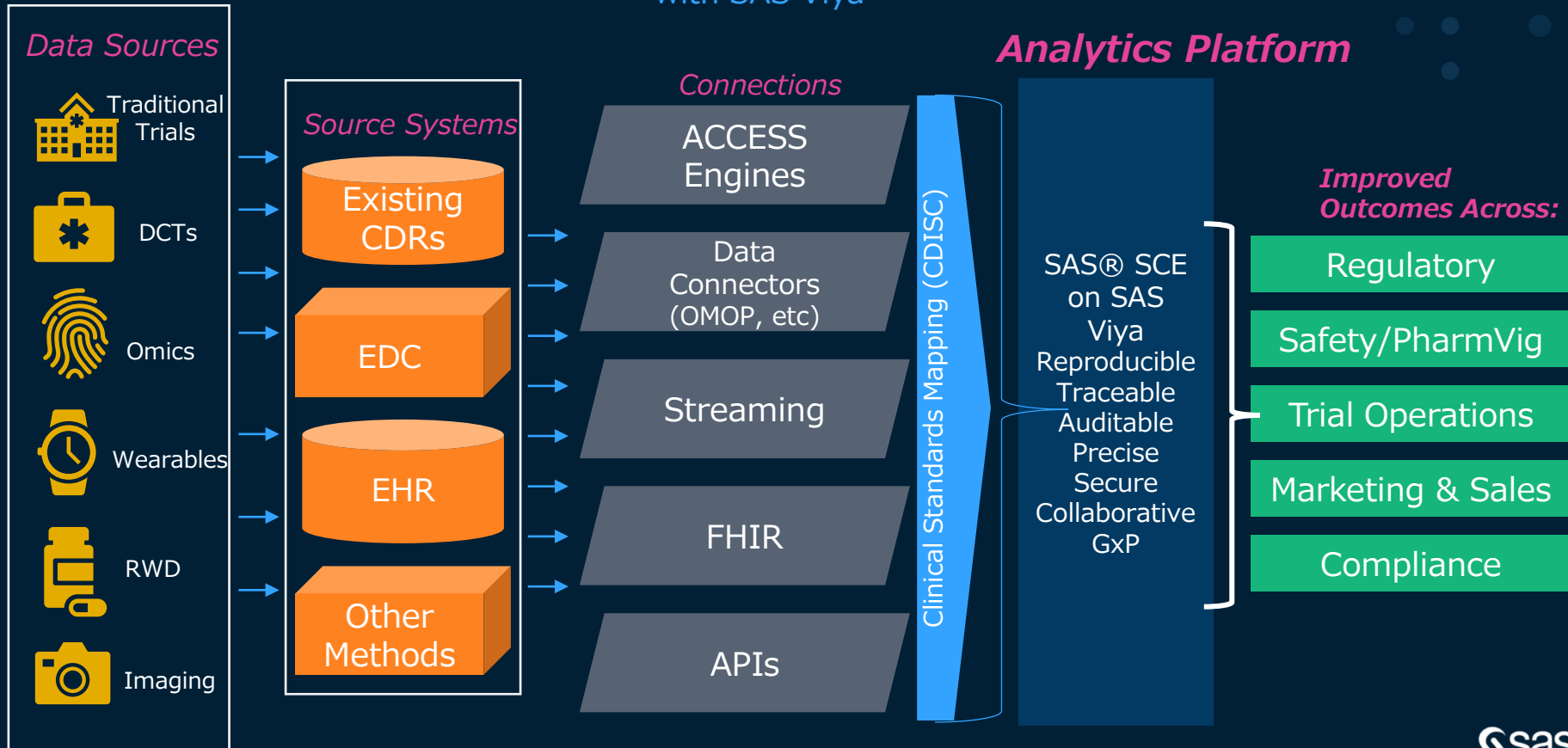
**Personas:** Head of IT, Head of Biostats, Study Lead, Biostatistician, Administrator, Programmer, Content Manager, Auditor/Regulator, IT Administrator

## Upcoming Features in MVP:

- Repository
  - Groups & Permissions
  - Roles & Privileges
  - Version Control
  - Electronic Signatures
  - Recycle Bin
  - Search
- Ownership Administration
- Audit History
- Action Status

# The Future of Efficient, Data-Driven Clinical Trials

with SAS Viya



# SAS<sup>®</sup> Health Cohort Builder

# Challenges in dealing with Real World Data

The amount of data to be processed is so large that SAS frozen (or appears to be frozen). In addition, the data wrangling and data cleansing steps took up my whole day.

There are concerns about the ability of our analytics platform on handling yearly increasing data volumes and data sources.



Data extraction is not feasible without skilled specialists.

# Challenges in dealing with Real World Data

SAS Viya technology accelerates data processing

Highly scalable analytics platform and easy to add additional computing nodes (processing servers), etc.

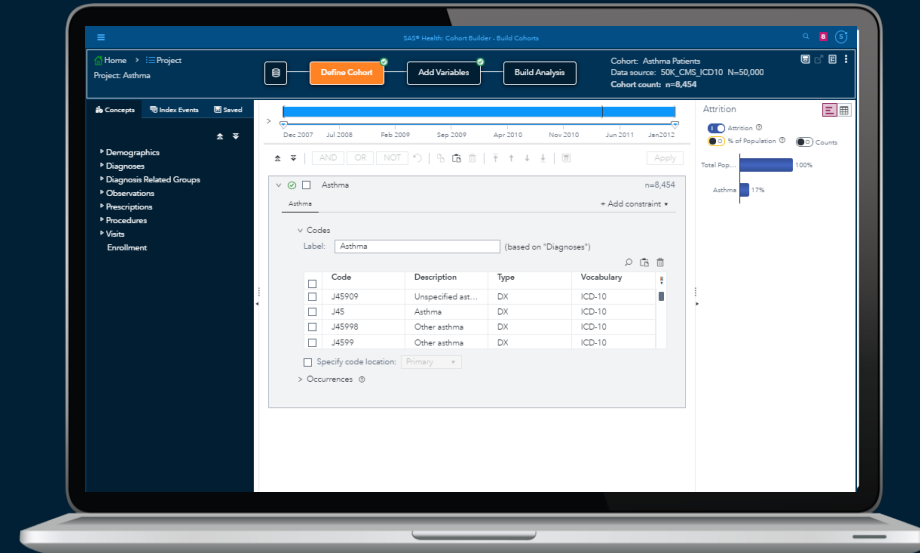


This system has a common data model and can handle data from any data source via GUI



## Easily and confidently build patient subsets within a unified analytics platform

- Interactive, drag-and-drop capability to support complex cohort queries with temporal relationships - no coding required

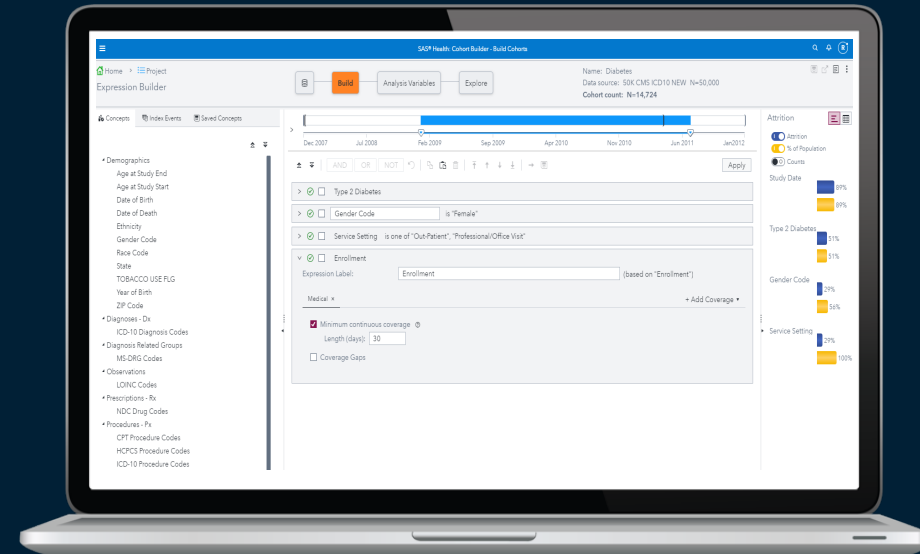






## Easily and confidently build patient subsets within a unified analytics platform

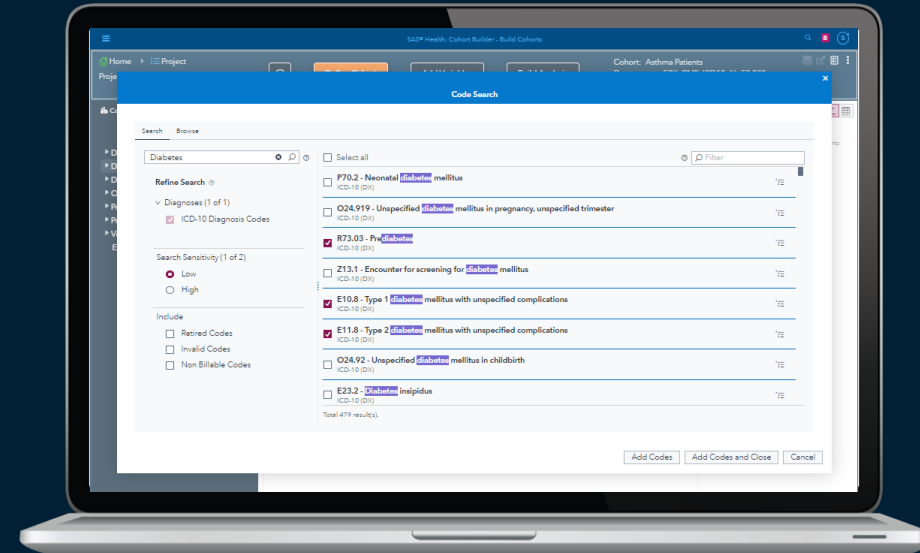
- Interactive, drag-and-drop capability to support complex cohort queries with temporal relationships - no coding required
- Explore cohort characteristics and the effect of inclusion / exclusion criteria on patient populations





## Easily and confidently build patient subsets within a unified analytics platform

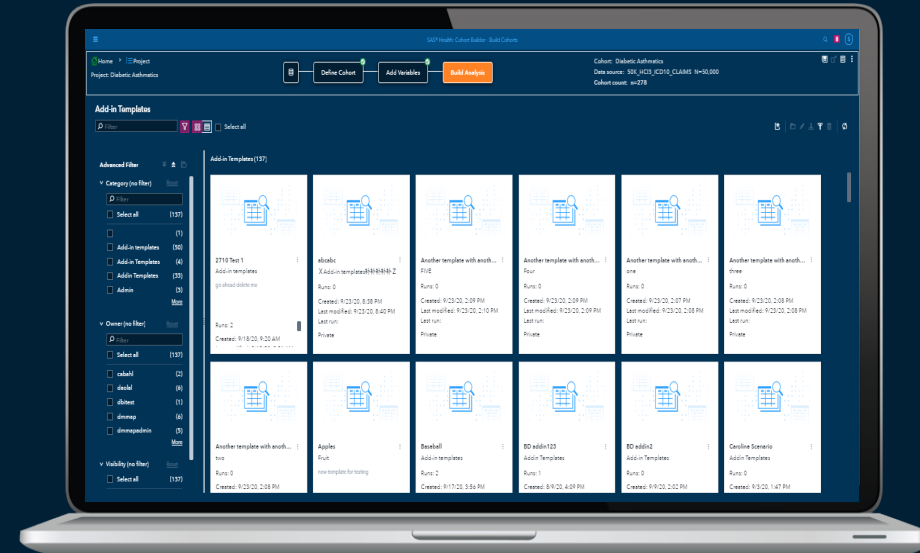
- Interactive, drag-and-drop capability to support complex cohort queries with temporal relationships - no coding required
- Explore cohort characteristics and the effect of inclusion / exclusion criteria on patient populations
- Utilize health care ontologies for selecting code sets to include within cohort criteria





## Easily and confidently build patient subsets within a unified analytics platform

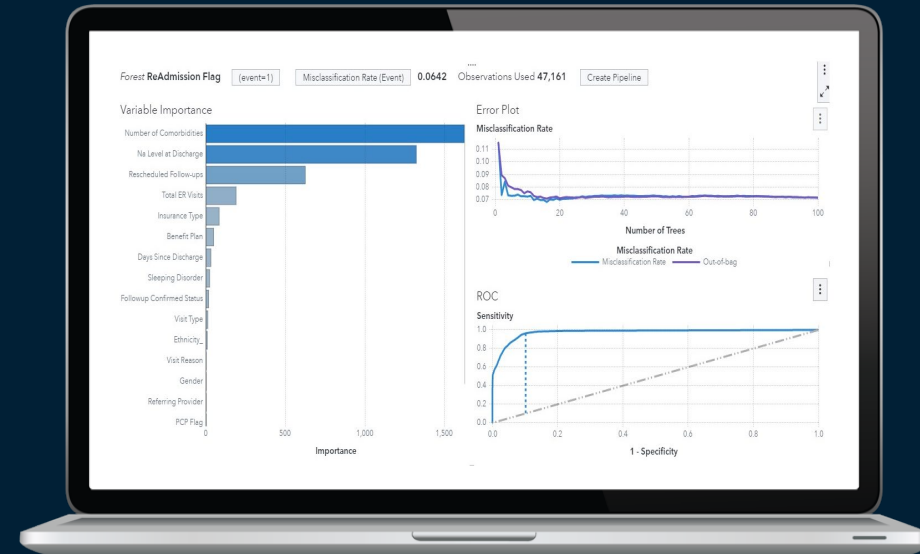
- Interactive, drag-and-drop capability to support complex cohort queries with temporal relationships - no coding required
- Explore cohort characteristics and the effect of inclusion / exclusion criteria on patient populations
- Utilize health care ontologies for selecting code sets to include within cohort criteria
- Leverage and extend your own analytic assets



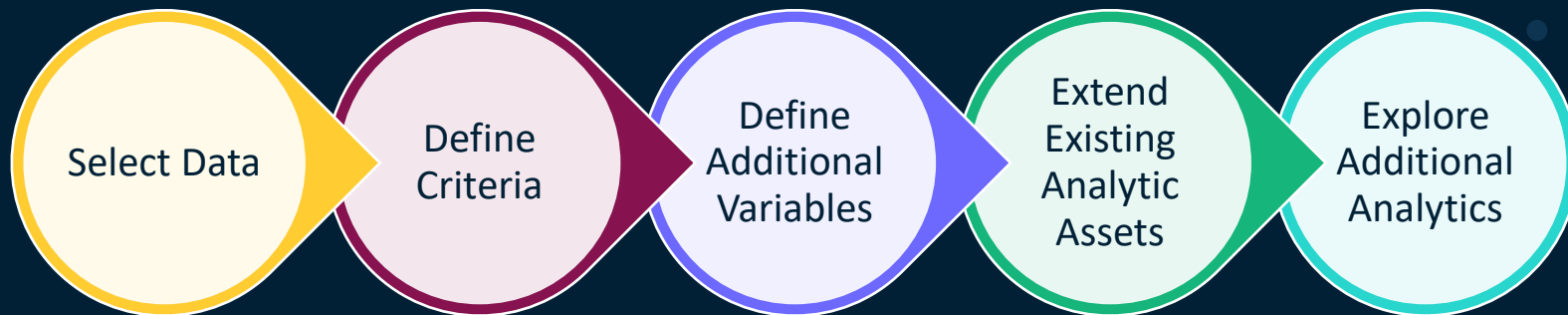


## Easily and confidently build patient subsets within a unified analytics platform

- Interactive, drag-and-drop capability to support complex cohort queries with temporal relationships - no coding required
- Explore cohort characteristics and the effect of inclusion / exclusion criteria on patient populations
- Utilize health care ontologies for selecting code sets to include within cohort criteria
- Leverage and extend your own analytic assets
- Gain fast, easy access to advanced health analytics and visualizations



# Overview of Process



- Map form
- Data valid
- No coding re
- Complex crit
- Share/reuse
- Predefined
- scores, etc
- Define and
- No co
- Trans
- SAS p
- Visual tools for exploration, analytics and AI/ML
- Coding interfaces for SAS and open source

# Project Management

The screenshot displays the SAS Health Cohort Builder interface. The top navigation bar includes a menu icon, the title "SAS® Health: Cohort Builder - Build Cohorts", a search icon, and a notification icon with the number "2". Below the navigation bar, there are tabs for "Projects" and "Data Sources". A search filter is present with the text "Filter".

The main area contains a grid of cohort cards. Two cards are highlighted with detailed information:

- Diabetes in our Population**  
Type: Cohort  
Owner: sheidz  
Shared: No
- Stroke in Afib Patie...**  
Type: Cohort  
Owner: sheidz  
Shared: Yes

The right-hand sidebar contains the following sections:

- Select one item to see its information.
- Summary Details Framework Menu
- [Add-in Manager](#)
- Cohorts
  - [Expression Manager](#)
- SAS Resources
  - [SAS Health](#)

# Data Source Selection

The screenshot displays the SAS Health Cohort Builder interface. The main area shows three data source cards, each with a bar chart and a 'Registered' status. The cards are:

- 100K\_3rdParty**: Date modified: Jan 30, 2021. Registered.
- 800K\_ICD10**: Date modified: Dec 19, 2020. Registered.
- Synthea\_100K**: Date modified: Dec 19, 2020. Registered.

The right sidebar contains navigation options: Summary, Profile, Framework Menu, Add-in Manager, Cohorts, Expression Manager, SAS Resources, and SAS Health.

**100K\_3rdParty**  
Date modified: Jan 30, 2021  
Registered

**800K\_ICD10**  
Date modified: Dec 19, 2020  
Registered

**Synthea\_100K**  
Date modified: Dec 19, 2020  
Registered

Select one item to see its information.  
Summary Profile Framework Menu  
Add-in Manager  
Cohorts  
Expression Manager  
SAS Resources  
SAS Health

- Map source data to standard format once
- Data load performs validation and profiling

# Define Inclusion and Exclusion Criteria

The screenshot displays the SAS Health: Cohort Builder - Build Cohorts interface. The top navigation bar includes 'Home', 'Project', and 'Project: Stroke in Afib Patients'. The main workflow consists of 'Define Cohort' (highlighted in orange), 'Add Variables', and 'Build Analysis'. The cohort name is 'Stroke in Afib Patients', the data source is '800K\_ICD10', and the cohort count is 'n=135,619'. The 'Define Cohort' step is active, showing a timeline from May 2007 to Jul 2013. Below the timeline, the 'Index Event' is 'Afib Dx'. The criteria section shows two inclusion criteria: 'Age at Index Date' is greater than or equal to 40, and 'Gender Code' is 'Male'. On the right, the 'Attrition' section shows a bar chart of cohort characteristics: Total Population (100%), Study Period (90%), Afib Dx (46%), Age ge 40 (45%), and Males (19%).

Selection of time period in the upper row. Filtering from data in the lower row

- No coding required
- Complex criteria supported
- Share/reuse criteria



# Define Inclusion and Exclusion Criteria

SAS Health: Cohort Builder - Build Cohorts

Project: Stroke in Afib Patients

Cohort: [redacted]  
Data source: 800K\_ICD10 N=711,479  
Cohort count: n=135,619

Define Cohort | Add Variables | Build Analysis

Index Event: Afib Dx

Criteria:

- Age at Index Date  $\geq$  40
- Gender Code is "Male"
- Hypertension
- NSAIDS
  - Relationships: Time to Hypertension: On or after: 0
  - Codes: NSAIDS (based on "Prescriptions")
- Occurrences:  $\geq$  1

Attrition:

- Total Population: 100%
- Study Period: 90%
- Afib Dx: 46%
- Age ge 40: 45%
- Males: 19%

Setting up a time series

Code	Description	Type	Vocabulary
<input type="checkbox"/> 50845013701	Anti Inflammatory oral ...	RX	NDC
<input type="checkbox"/> 52904048602	Pain Relief Anti inflam...	RX	NDC
<input type="checkbox"/> 70529011801	Cortaren Corticosteroi...	RX	NDC
<input type="checkbox"/> 70529011201	Neuromaquel Neurom...	RX	NDC

# Definition of additional variables (e.g., used when there is data that could be a risk factor)

SAS Health: Cohort Builder - Build Cohorts

Project: Stroke in Afib Patients

Cohort: [redacted]  
Data source: 800K\_ICD10 N=711,479  
Cohort count: n=135,619

Filter

- ethnicity
- Gender Code
- GFR test
- GI Bleed
- HbA1c
- HBP
- History of Obstructive CVD
- HTN
- Hyperlipidemia
- Hypert3
- Hypertension
- Stroke
- StrokeSCA
- T2DM
- Test
- Testing for CVD
- Valsartan
- Victoza
- ZIP Code

Name	Type	Results
<input type="checkbox"/> <a href="#">Antiarrhythmic</a>	Custom Variable - Prescriptions	
<input type="checkbox"/> <a href="#">Anticoagulants</a>	Custom Variable - Prescriptions	
<input type="checkbox"/> <a href="#">Cardioversion</a>	Custom Variable - Procedures	
<input type="checkbox"/> <a href="#">Charlson Risk Factor Set</a>	Risk Factor Set	

- Predefined items such as risk scores, etc.
- Define and share others

# Utilization of Analysis Templates (pre-created template selection)

SAS Health: Cohort Builder - Build Cohorts

Project: Stroke in Afib Patients

Define Cohort | Add Variables | **Build Analysis**

Cohort: [Cohort ID]  
Data source: 800K\_ICD10 N=711,479  
Cohort count: n=135,619

### Add-in Templates

Filter [ ] Select all

**Advanced Filter**

- Category (no filter) [Reset](#)
  - Select all (13)
  - Add-in templates (9)
  - Patient-level Outcomes Predictive Model (3)
  - Project-level Template (1)
- Owner (no filter) [Reset](#)
  - Select all (13)
  - [Owner] (5)
  - [Owner] (4)
  - sas (1)
  - sas.healthcohortservice (1)
  - [Owner] (1)[More](#)
- Visibility (no filter) [Reset](#)
  - Select all (13)
  - Private (7)
  - Public (6)

**Add-in Templates (13)**

- Cohort Characterization Add-in templates  
Owner: sas.healthcohortser...  
Runs: 5  
Last modified: 5/22/21, 3:39 AM
- Health Outcomes Binary Target Patient-level Outcomes Predictive...  
Health Outcomes Prediction Model with cohort output tables as input data  
Owner: daolal  
Runs: 4  
Last modified: 4/29/21, 2:21 AM
- Identify Top Codes (By Memb... Add-in templates  
Add-in to create datasets with highest diagnosis, procedure and prescription  
Owner: sas  
Runs: 3  
Last modified: 6/18/21, 11:23 PM
- Incidence Prevalence Measur... Patient-level Outcomes Predictive...  
New Add-in t

- No coding needed to apply
- Transparent methodology
- SAS programmers can create

# Adjustment of Output Variables

The screenshot displays the SAS Health Cohort Builder interface. At the top, the navigation bar includes 'Home', 'Project', and 'Build Analysis' buttons. The project name is 'Stroke in Afib Patients'. The cohort is identified as '800K\_ICD10' with N=711,479 and a count of n=135,619.

The main workspace is titled 'DemoPurpose\_BinaryRegressionPredictiveModel'. It features a 'Prediction Model Setting' section where the target outcome variable is 'Stroke' and the target variable value is '1'. Below this, a 'Predictors' section lists available variables: Stroke, Census State, Race Code, and age\_at\_study\_end. A 'Selected items (7):' list includes Antiarrhythmic, Anticoagulants, and Cardioversion.

The 'Output Data' tab is active, showing a 'Logistic Regression' model. The model information table is as follows:

Model Information	
Data Source	CB_AV_STROKE_FLG
Response Variable	AV_STROKE_FLG
Distribution	Binary
Link Function	Logit
Optimization Technique	Newton-Raphson with Ridging

The 'Number of Observations' table is:

Description	Total	Training	Validation
Number of Observations Read	135619	101714	33905
Number of Observations Used	135619	101714	33905

The 'Response Profile' table is:

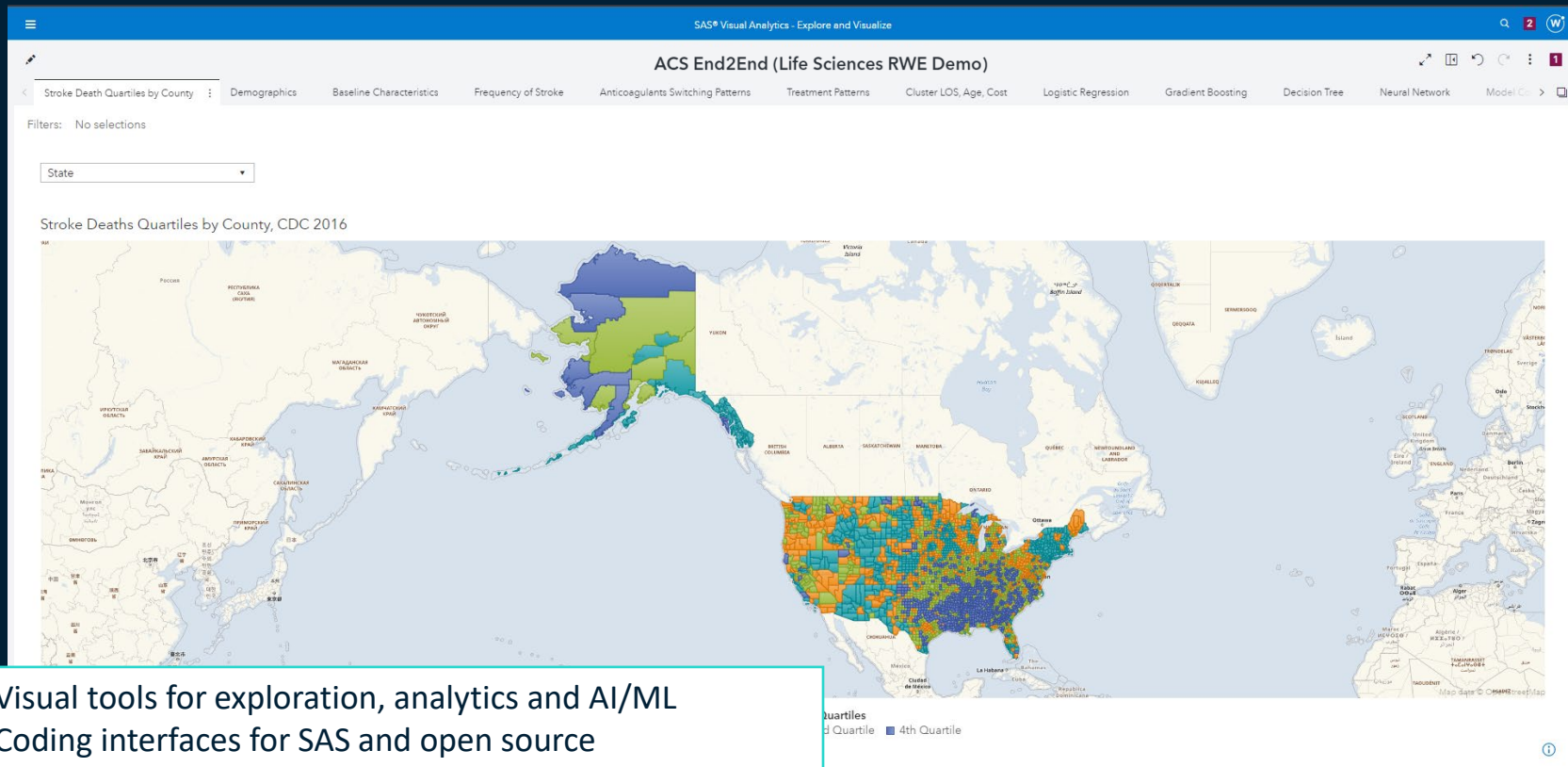
Ordered Value	AV_STROKE_FLG	Total Frequency	Training	Validation
1	0	111092	83214	27878
2	1	24527	18500	6027

The text below the table states: 'Probability modeled is AV\_STROKE\_FLG = 1.'

The 'Selection Information' table is:

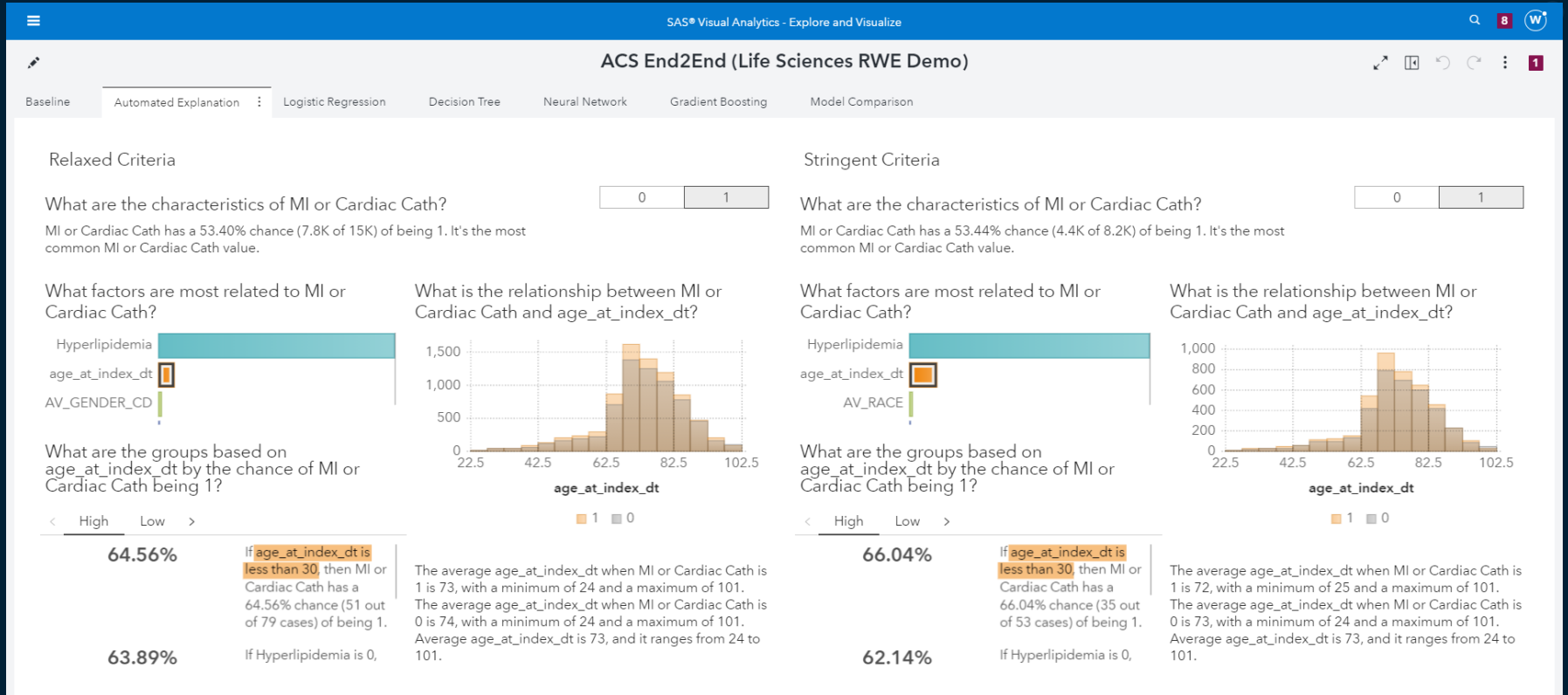
Selection Information	
Selection Method	Stepwise
Select Criterion	SBC

# Visualization of Analysis Outputs (VA function)

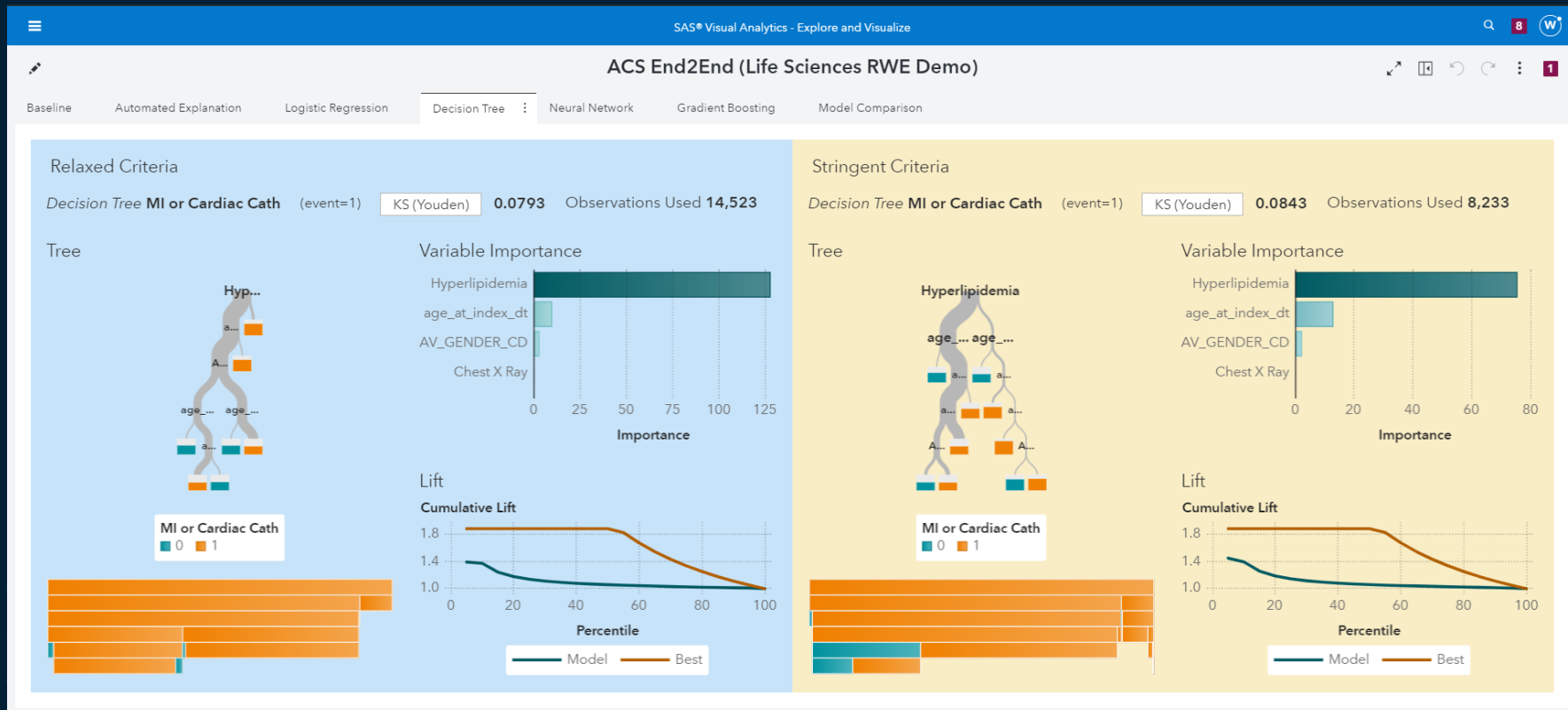


- Visual tools for exploration, analytics and AI/ML
- Coding interfaces for SAS and open source

# Utilization of analysis results. Automatic analysis function for output (AI function)



# Utilization of analysis results. Automatic analysis function for output (AI function)



# SAS® Clinical Enrollment Simulation (Overview)



Most clinical trials fail to meet enrollment timelines.



Approximately **90%** of clinical trials do not meet patient enrollment targets on time.



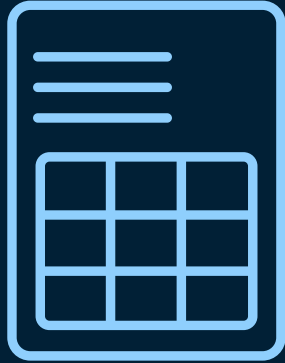
**Every day** of delay in clinical trial completion can cost sponsors **millions**.



Missing clinical subject recruitment milestones can **negatively impact the reputation** of a Clinical Research Organization (CRO).

# Traditional Approach to Projecting Enrollment

Using Excel to Calculate Future



- Typically work backwards from goal
- Unrealistic assumptions from static estimates with no variability in inputs
- Doesn't account for enrollment decays over time
- Approximately 20% of sites never enroll anyone
- Costs are not accounted for
- Difficult to track “what if” scenarios

# SAS Helps You Optimize Study Start-up and Progress

Identify the right sites and optimal clinical strategy.

---



Ensure clinical trials get up and running quickly, efficiently and in keeping with your milestones for completion.



# Optimize Study Start-up and Progress

How SAS Delivers



- Powerful simulation analytics to quickly obtain insights
- Simulate different enrollment scenarios, costs and global regulatory timelines
- Better insights to inform decision making and trial strategy
- Easy data integration and real-time visualizations

# Our Approach

## Discrete Event Simulation Yields Better Enrollment Projections



### Develop Enrollment Projections

- Enter all assumptions and simulate multiple replications of the enrollment
- Show likelihood to achieve specific targets



### Capture Variability

- Get a better understanding of likely outcomes
- Produce more accurate predictions for achieving targets



### Realistic Enrollment Estimates

- Obtain peak and lower enrolment period windows to account for enrolment decay



### Include Costs

- Include startup costs for countries and sites and per-patient costs
- See the cost impact of different scenarios







### Scenario Tracking

- Compare and track multiple scenarios with varying inputs
- Test the impact of factors such as adding new sites and/or countries or changing baseline assumptions

# Simulating Enrollment output例

## Calculate Probabilities of Success

			Success Rate Range		
Target Name	Target Date	Success Rate	Best	Median	Worst
<b>Dates</b>					
First Country Approval	Not provided	N/A	July 26, 2022	August 6, 2022	August 25, 2022
First Site Startup	August 31, 2022	 46%	August 19, 2022	September 1, 2022	September 21, 2022
First Subject Screening	Not provided	N/A	August 23, 2022	September 5, 2022	September 25, 2022
First Subject Enrollment	September 15, 2022	 89%	August 23, 2022	September 6, 2022	September 25, 2022
First Subject's First Visit	Not provided	N/A	August 30, 2022	September 13, 2022	October 2, 2022
First Subject's Last Visit	Not provided	N/A	February 19, 2023	March 5, 2023	March 24, 2023
End of Enrollment	December 31, 2023	 69%	November 2, 2023	December 1, 2023	March 25, 2024
Last Subject's First Visit	Not provided	N/A	November 9, 2023	December 8, 2023	April 1, 2024
Last Subject's Last Visit	Not provided	N/A	April 30, 2024	May 29, 2024	September 21, 2024
<b>Cost</b>					
Estimated Cost	\$2,900,000	 61%	\$2,802,916	\$2,885,824	\$2,974,546

# Site Setup Values

SAS® Clinical Enrollment Simulation Cloud 12 D

Home  
BUILDER  
Scenarios  
ELEMENTS  
Studies  
Countries  
Sites

Sites > Site: Twin Pines Medical Center

## Twin Pines Medical Center

### Details

Name: \*

Description:

Created:  / December 09, 2022 at 09:12:23 AM

Last modified:  / December 09, 2022 at 09:12:23 AM

Country: United States

Number of available sites: \*

### Therapeutic Areas

<input type="checkbox"/> Therapeutic Area *	Startup Cost * <input type="text"/>	Best Case * Startup Time (Days) <input type="text"/>	Typical * Startup Time (Days) <input type="text"/>	Worst Case * Startup Time (Days) <input type="text"/>	Startup Success... <input type="text"/>	Probability of E... <input type="text"/>	Months High R... <input type="text"/>	High Recruitment * Subjects per Month <input type="text"/>	Low Recruitment * Subjects per Month <input type="text"/>
<input type="checkbox"/> Cardiology/carc	<input type="text" value="\$7,000"/>	<input type="text" value="21"/>	<input type="text" value="30"/>	<input type="text" value="45"/>	<input type="text" value="99%"/>	<input type="text" value="2%"/>	<input type="text" value="9"/>	<input type="text" value="6"/>	<input type="text" value="4"/>

# Setting of simulation parameters for the entire clinical trial

SAS® Clinical Enrollment Simulation Cloud

Home | BUILDER | Scenarios | ELEMENTS | Studies | Countries | Sites

Scenarios > Scenario: Cardio 13 with 10 US Sites

Cardio 13 with 10 US Sites

Scenario and Study Details | Study Discontinuation Estimates | Visit Schedule | Additional Targets | Countries and Sites

Country Settings

<input type="checkbox"/>	Name	Startup Cost *	Best Case Startup Time (Days) *	Typical Startup Time (Days) *	Worst Case Startup Time (Days) *	Startup Success Probabi...	Concurrent Sites Possible *	Screen Failure Rate
<input type="checkbox"/>	United States	\$12,000	24	30	60	100%	100	22%

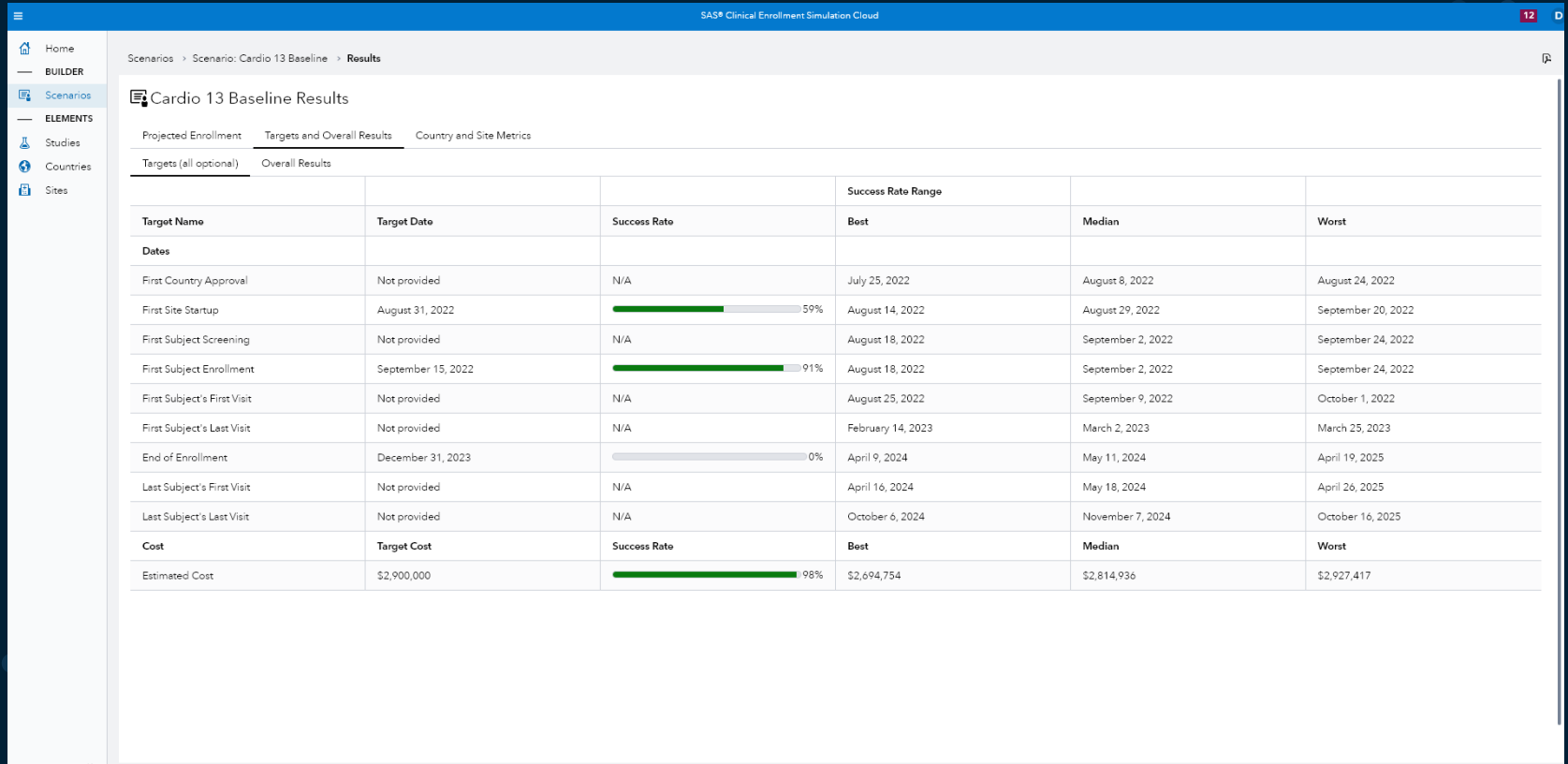
Site Settings

United States Sites

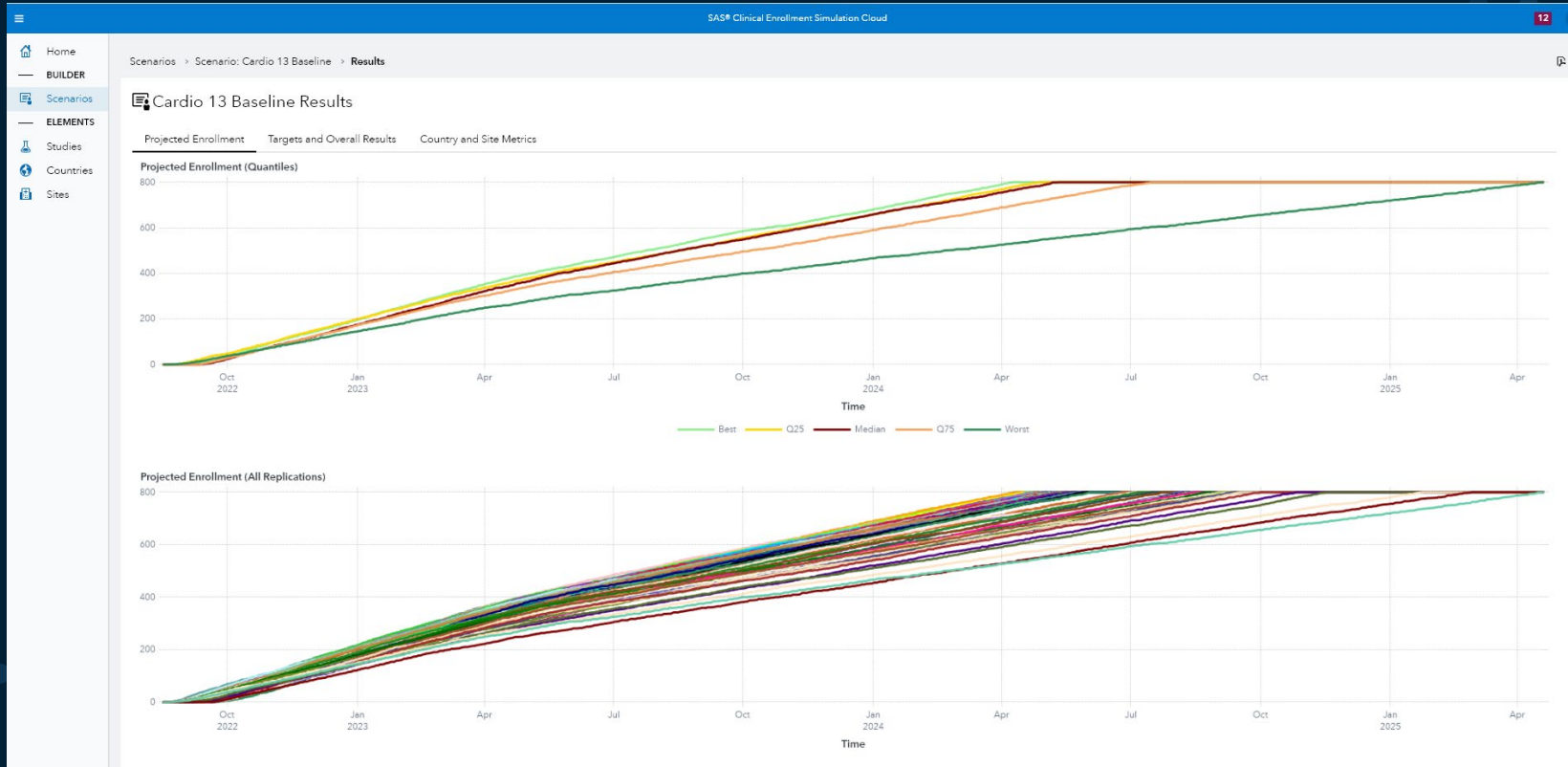
Site	Number of Sites *	Startup Cost *	Best Case Startup Time (Days) *	Typical Startup Time (Days) *	Worst Case Startup Time (Days) *	Startup Success...	Probability of E...	Months High R...	High Recruitment Subjects per Mont
Chenango Me...	1	\$5,500	14	21	30	99%	2%	8	10
Derry Medical...	1	\$4,500	21	30	45	95%	2%	8	9
Lassen Gener...	1	\$6,000	21	30	45	97%	2%	9	8
Sacred Heart ...	1	\$5,000	21	30	45	99%	5%	6	10
Shasta Region...	1	\$5,500	21	30	45	99%	2%	6	9
St Eligus Hosp...	1	\$7,500	21	30	45	99%	2%	8	7
Tower Medica...	1	\$6,550	21	30	45	99%	2%	7	6
Twin Pines Me...	1	\$7,000	21	30	45	99%	2%	9	6
Western Regi...	1	\$8,000	21	30	45	99%	2%	5	8
Wexler Medic...	1	\$4,250	21	30	45	99%	2%	8	8



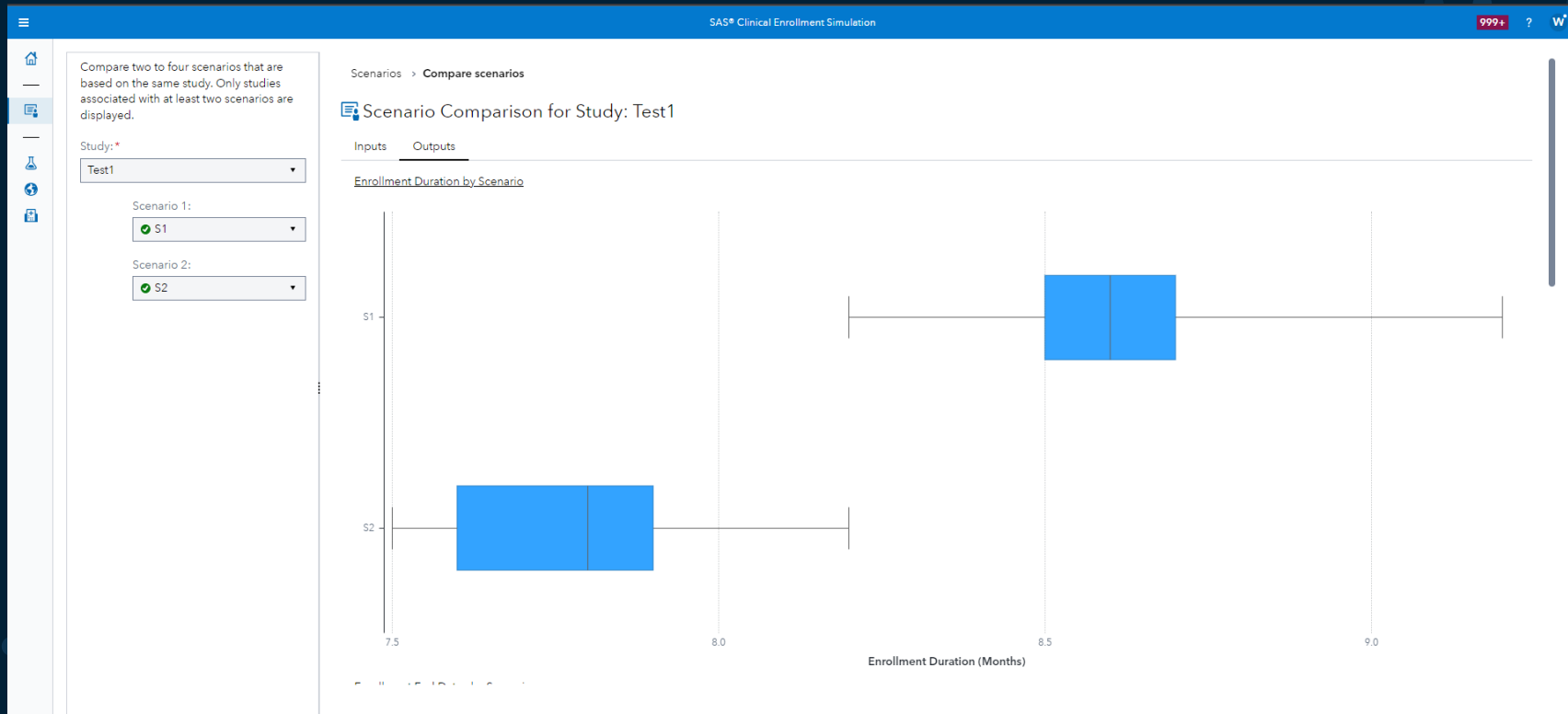
# Simulation Result Screen



# Simulation Result Screen



# Simulation Result Screen Recruitment Period



# Thank you!!

- Today's presentation may seem to be a bit busy, as we have explained several products in a short time. If you are interested, please feel free to contact us.

- Contact :

SAS Japan Consulting Service Supervisory Division

Life Science Team

Toshiaki Habu

Mail: [toshiaki.habu@sas.com](mailto:toshiaki.habu@sas.com)