REAL WORLD PATIENTS: RISK ASSESSMENT AND PREDICTION MODELS FOR HOSPITALIZATION AND READMISSION

David Olaleye, PhD
SAS Institute, Cary, NC

Future forward with SAS.

Digitize the world of pharma.
Shape the future of analytics.
Accelerate your career.

This information is confidential and covered under the terms of any SAS agreements as executed by customer and SAS Institute Inc.

About the Author

David Olaleye, PhD

David is Senior Manager and Principal Research Statistician at SAS Institute in Cary, NC. He received his postgraduate training in statistics, demography, and clinical epidemiology from the University of Pennsylvania School of Arts and Sciences, and School of Medicine, Philadelphia. His areas of specialization include episode-of-care analytics, real world evidence, pharmacovigilance and drug safety, data science, and developing statistical algorithms for mining healthcare and clinical trials databases. David has authored and published articles in peer-reviewed journals.

This information is confidential and covered under the terms of any SAS agreements as executed by customer and SAS Institute Inc.

AGENDA

01 Introduction

02 Real World Data
- Solving complex problems

03 Real World Patients
- SAS® Episode Analytics
- SAS® Real World Evidence

04 Real World Questions
- Readmissions
- Cost prediction

05 Conclusion

rmation is confidential and covered under the terms of any SAS
reements as executed by customer and SAS Institute Inc..

Solving Complex Problems

WHAT

Real World Data

Volume
Velocity
Variety
Veracity

HOW

Real World Patients

Patient safety and drug surveillance

Treatment pathways

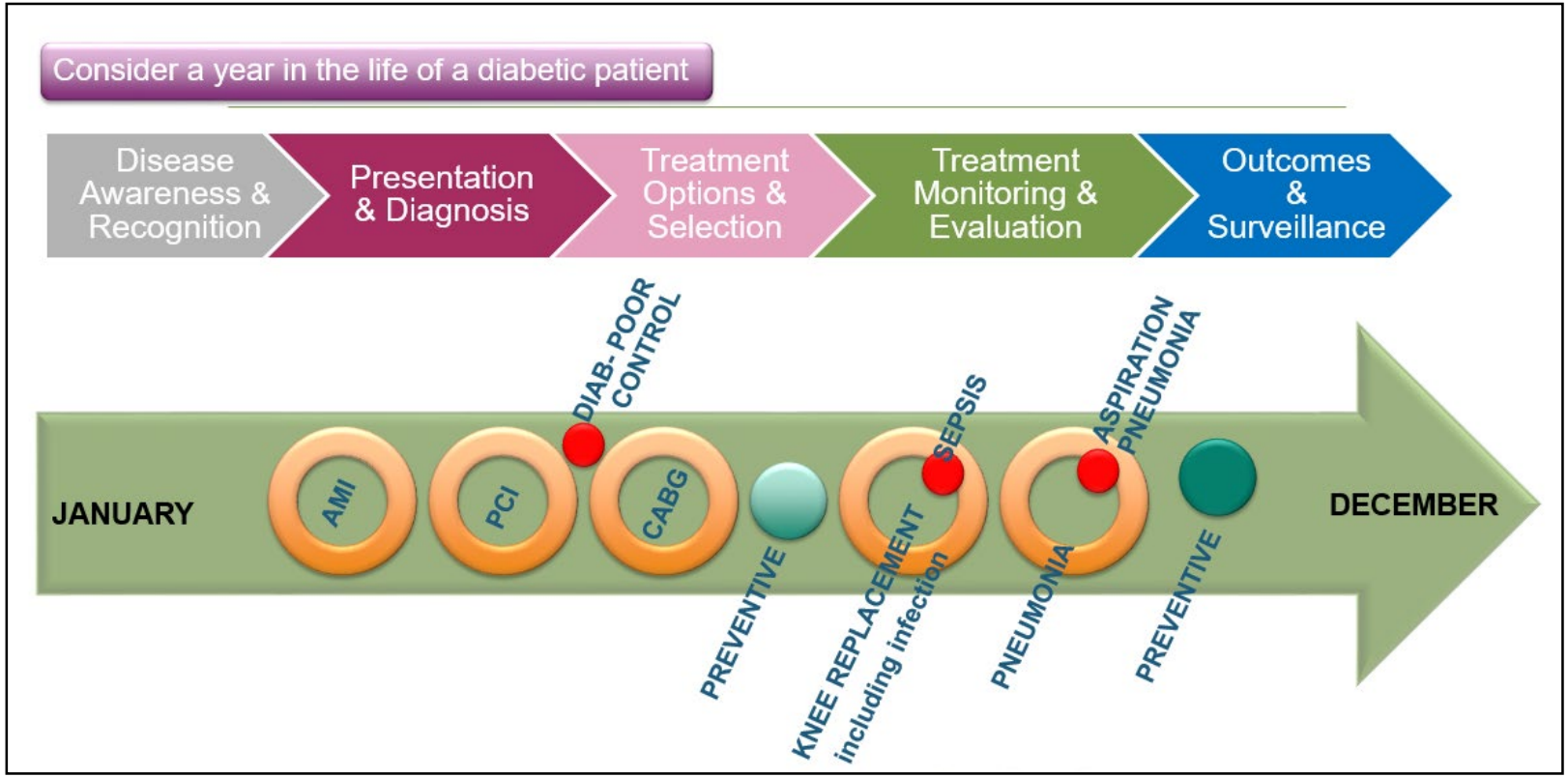
EXECUTE

Real World Questions

Market safe and effective products

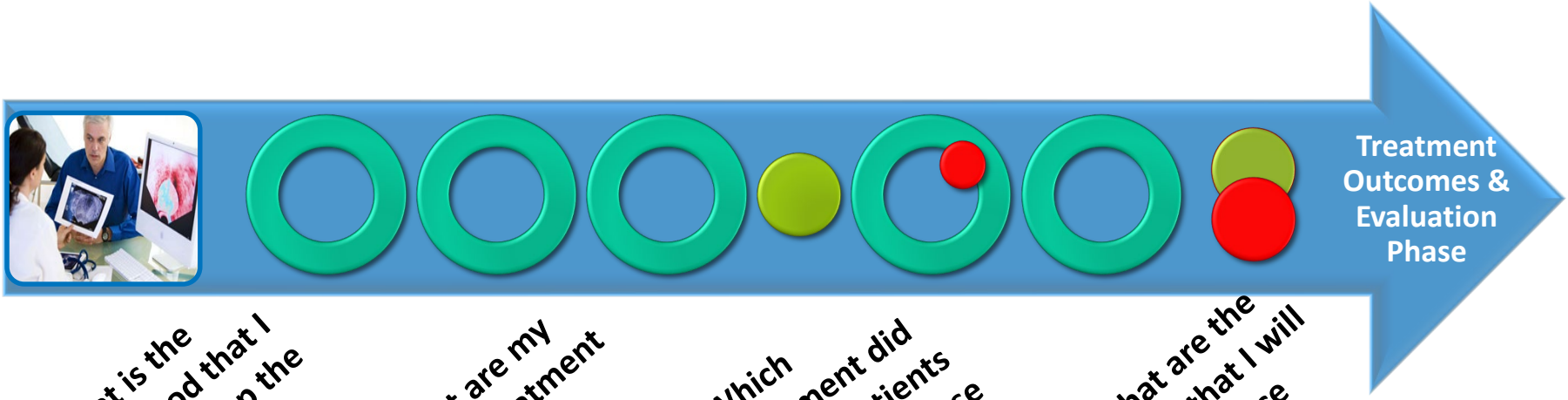
Increase shareholder value

A diabetic patient's episode journey



This information is confidential and covered under the terms of any SAS agreements as executed by customer and SAS Institute Inc..

A diabetic patient's episode journey



What is the likelihood that I will develop the disease?

What are my best treatment options?

Which treatment did other 'patients like me' choose after diagnosis?

What are the odds that I will experience adverse outcomes?

This information is confidential and covered under the terms of any SAS agreements as executed by customer and SAS Institute Inc..

Key Challenges

- Q: leverage claims, clinical, survey, and care management data to perform prospective and retrospective analysis on my patient population?
- Q: uncover insights, patterns and determine which episodes are clinically related?
  - **Intelligent episode-of-care data aggregation**

- Q: proactively identify high risk members for targeted care outreach?
- Q: predict patients at risk of index- or re-admission based on their clinical profiles?
  - **Predictive analytics**

- Q: uncover relationships between provider practice patterns and quality of care?
- Q: identify gaps in care and minimize adverse health outcomes
  - **Prescriptive analytics**

This information is confidential and covered under the terms of any SAS agreements as executed by customer and SAS Institute Inc..

Episode Analytics Life Cycle Flow

Data Preparation
- Profile
- Cleanse
- Episode Definition

1

Route: Journey
- Index Cohort
- Episode Construction
- Care Pathways

2 3

Exploration
- Frequencies
- Correlations
- Categories

4

Analytics
- Machine Learning
- Logistic Regression
- Risk Segmentation

5

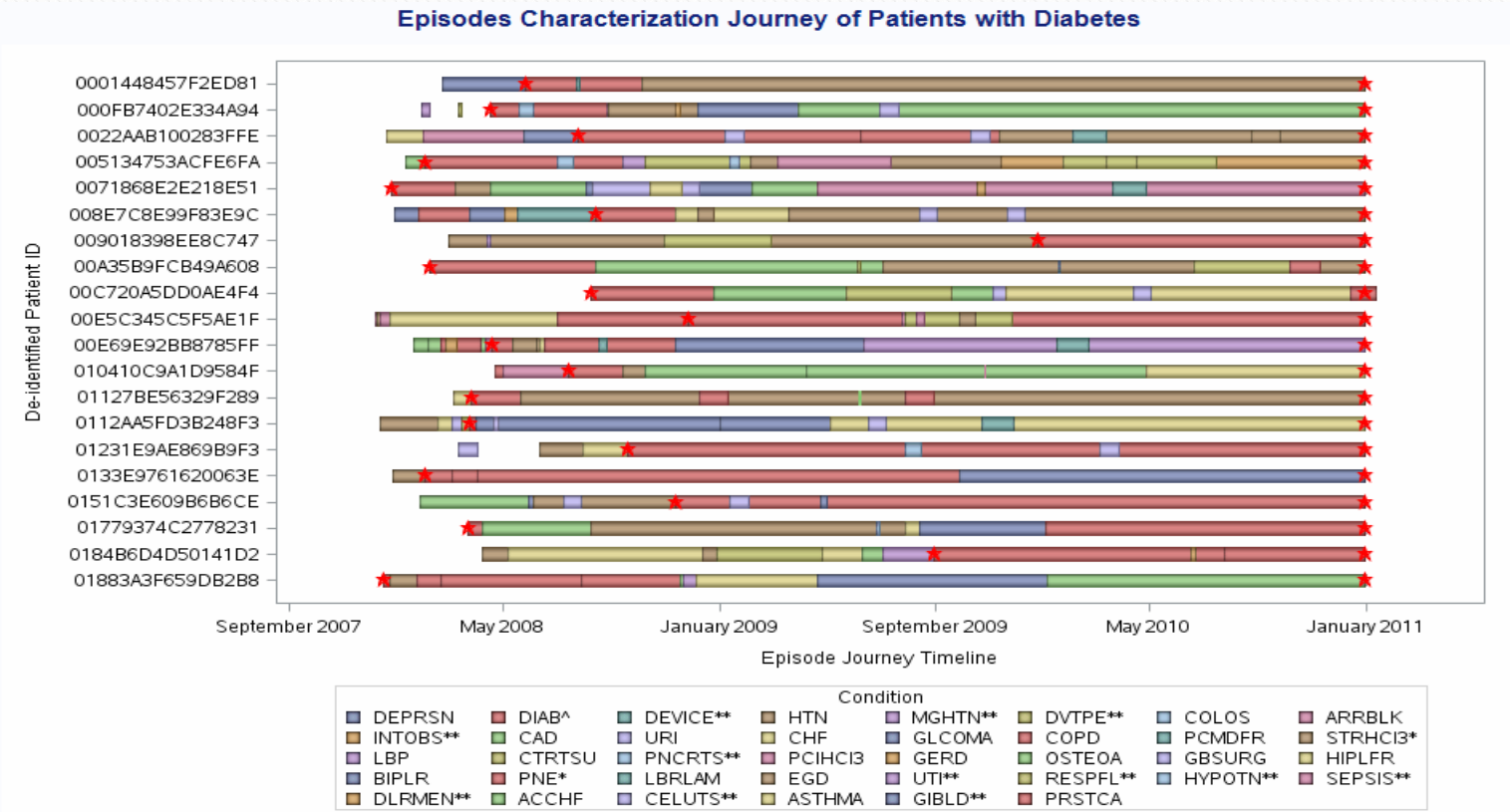
Meaning
- Readmission/Cost
- Mortality/Survival
- Efficiency/Value

Data Scientist

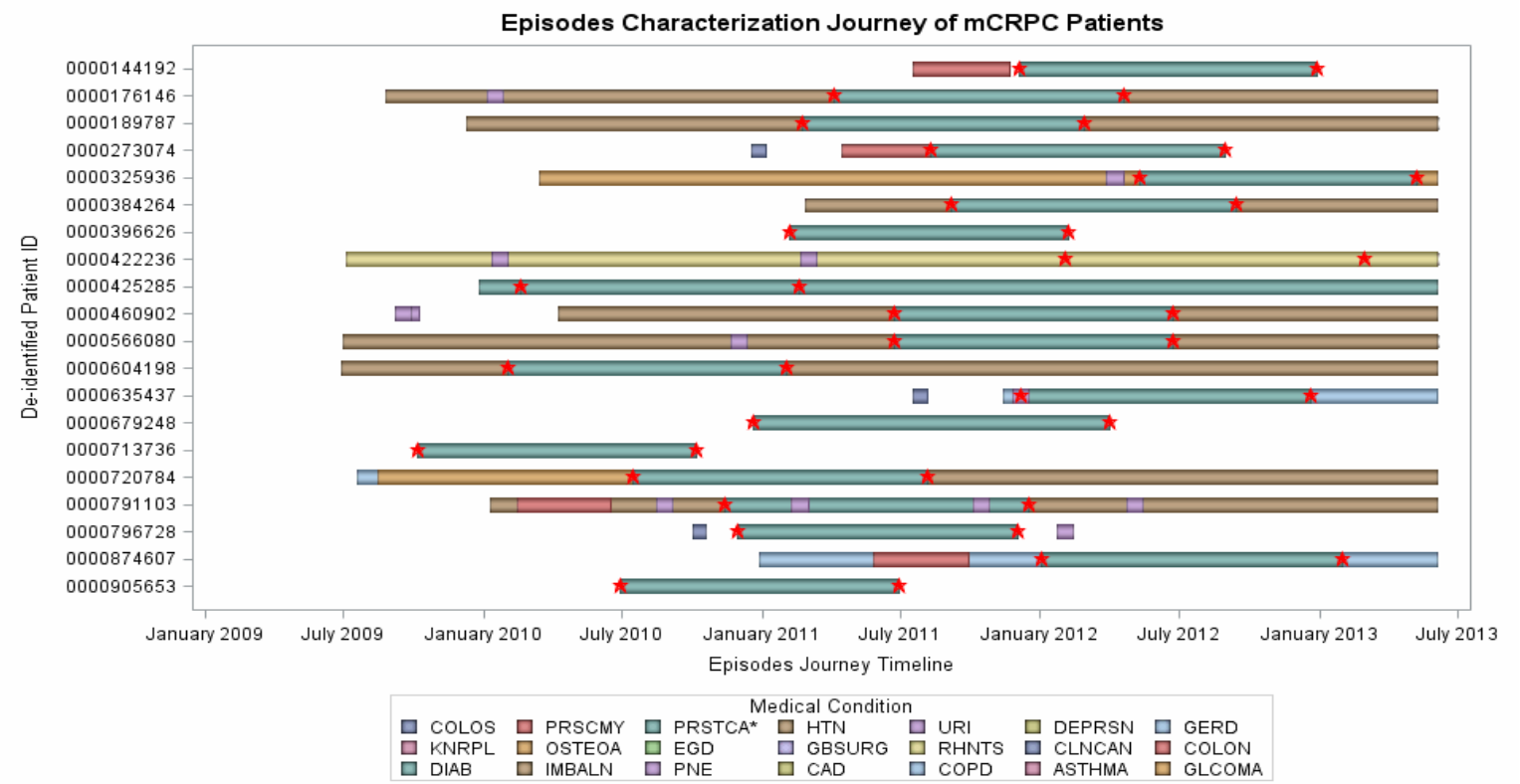
Patient / Payer / Provider

This information is confidential and covered under the terms of any SAS agreements as executed by customer and SAS Institute Inc..

Episode of Care Journey Profile



Episode of Care Journey Profile



This information is confidential and covered under the terms of any SAS agreements as executed by customer and SAS Institute Inc..

Cohort Builder: SAS® Real World Evidence

- SAS® Visual Analytics platform
  - Manage, analyze, and visualize real world data
  - Cohort creation and discovery
    - Process cohorts based on simple/complex query logic and rules
    - Temporal relationships across different event codes – diagnoses, procedures, prescriptions, labs, etc.
    - Create user-defined and pre-defined analysis variables
  - Pre-built analytical add-in models
  - Integrated with SAS Visual Analytics for visualization and reporting
  - Advanced analytics using SAS® Viya® machine learning algorithms

This information is confidential and covered under the terms of any SAS agreements as executed by customer and SAS Institute Inc..

Analytic Cohort Discovery and Creation

SAS® Real World Evidence

David Olaleye

Home › dave project › Diabetes Rx Surveillance

Name: Diabetes Rx Surveillance
Description: Diabetes Rx Surveillance with CMS 800K
Date Range: Jan 1, 2008 - Dec 31, 2010

Data source: 800K_CMS_ICD10_RWE43_20180511 *n=711,479*
Data source description: 800K_CMS_ICD10_RWE43_20180511
Enrollment: None *n=636,980 ; 89.53%*

Type: Index Event Cohort

Actions

Search

- Demographics
- Diagnoses - Dx
- Diagnosis Related Groups
- Laboratory Tests
- Prescriptions - Rx
- Procedures - Px
- Visits
- Vitals

- Index Events

1 Cohort — 2 Analysis Variables — 3 Add-ins

Index Event | DM RXs IED | n=163,582 23%

Inclusion/Exclusion Criteria | n=103,802 15%

AND | OR | NOT | Apply

AGE_GE18	*n=163,582*
AMI	*n=13,189*
Biguanides	*n=86,506*
Dopamine agonist	*n=18,841*
DPP-4 inhibitors	*n=4,817*
Glucagon	*n=912*
Meglitinides	*n=5,740*
SGLT2 inhibitors	*n=0*
Sulfonylureas	*n=82,081*
Thiazolidinediones	*n=16,234*
Diabetes_ECOM	*n=103,802*

Cohort Variables

Filters

Name	Count / %
AGE_GE18	163,582 / 100%
AMI	13,189 / 8%
Biguanides	86,506 / 53%
Dopamine agonist	18,841 / 12%
DPP-4 inhibitors	4,817 / 3%
Glucagon	912 / 1%
Meglitinides	5,740 / 4%

User-defined Analysis Variables

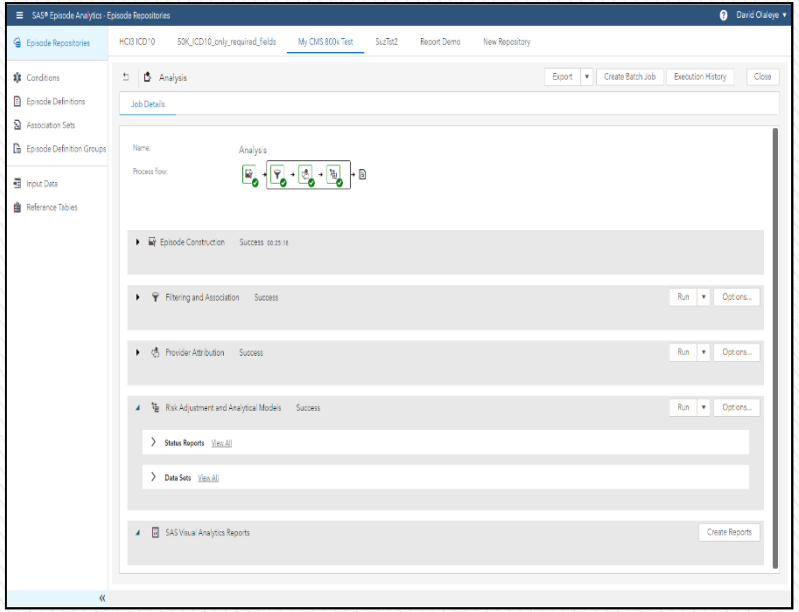
Name	Count / %
Candida Infection	8,418 / 5%
C-Difficile Infection	3,143 / 2%
rhabdomyolysis	2,108 / 1%
Diabetes poor control	38,392 / 23%
Overdose Poisonings	182 / 0%
Complications from Amputation	779 / 0%
Severe Sepsis	5,269 / 3%

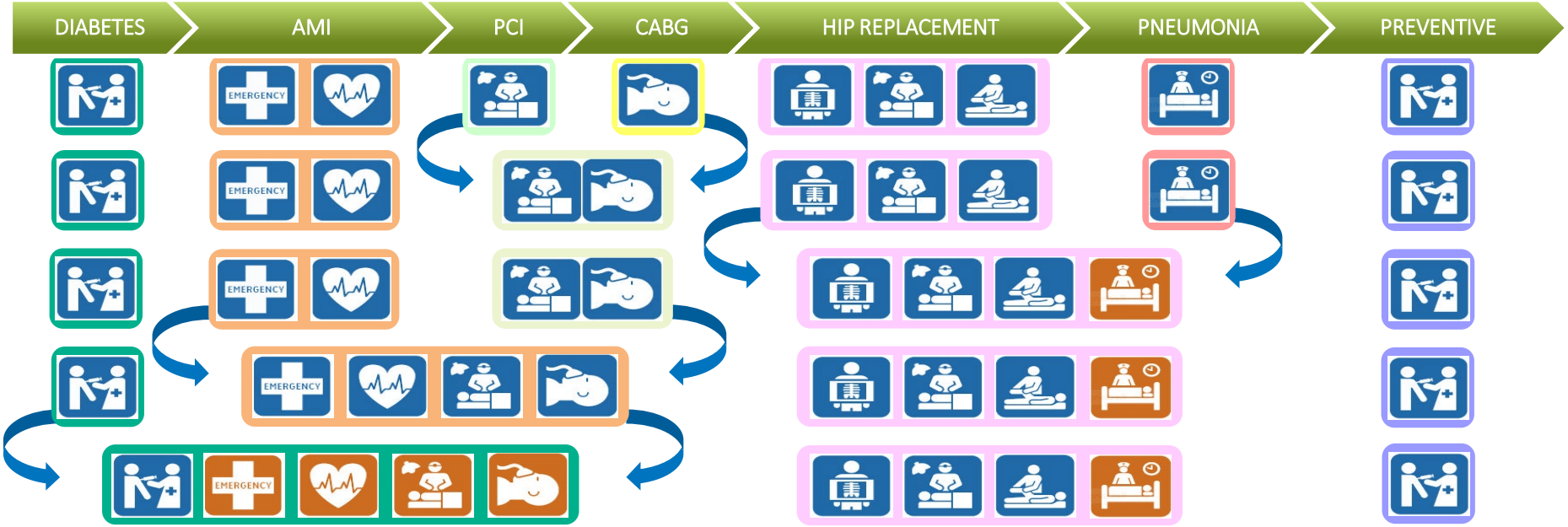
This information is confidential and covered under the terms of any SAS agreements as executed by customer and SAS Institute Inc..

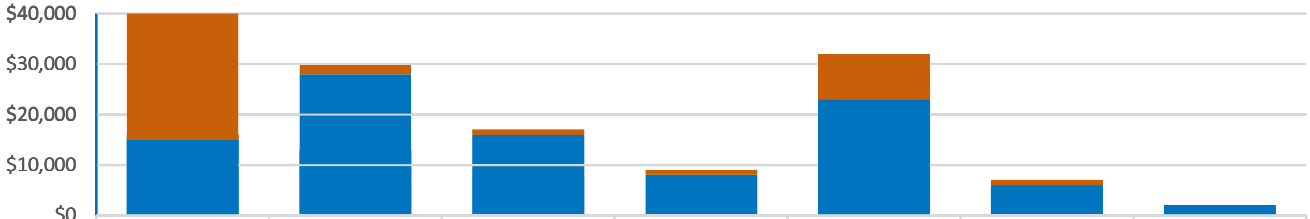
Episodes Builder: SAS® Episode Analytics

Episode bundling solution for payers and providers

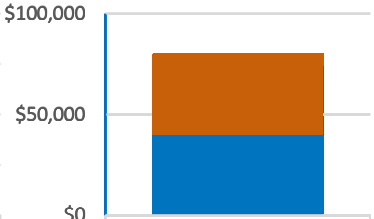
- automated construction of episodes for chronic, acute, surgical conditions, etc.
- industry-standard and user-customized episode definitions
- cost allocation, filtering and episode association logic
- provider attributed episodes assignment







	DIABETIC	AMI	PCI	CABG	KNEE REPL	PNEUMONIA	PREVENTIVE
PAC	$1,000	$1,000	$1,000	$1,000	$9,000	$1,000	
TYPICAL	$15,000	$12,000	$16,000	$8,000	$23,000	$6,000	$2,000



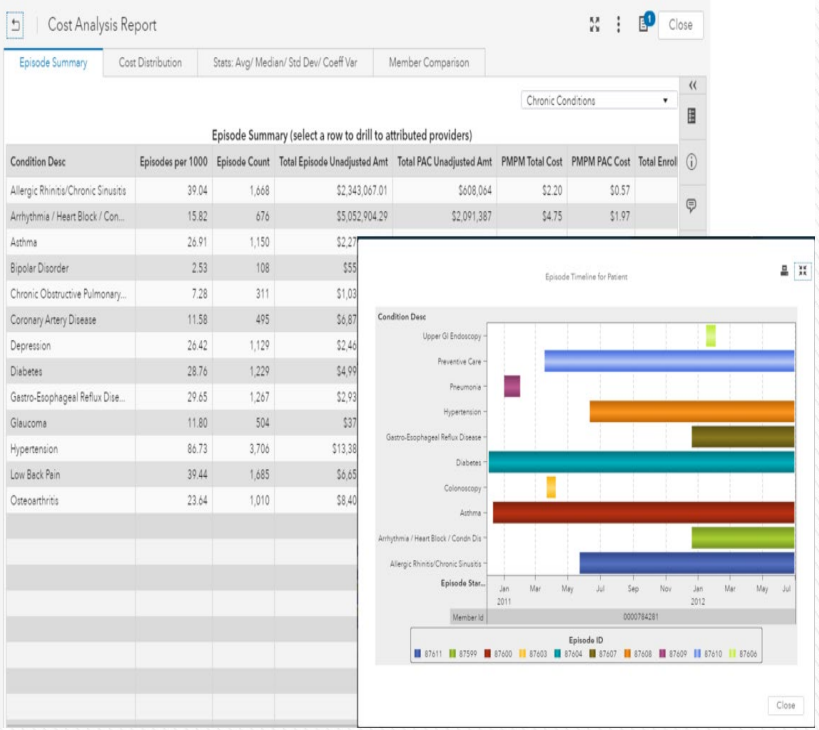
	PREVENTIVE
PAC	$12,000
TYPICAL	$74,000

Copyright © 2015, SAS Institute Inc. All rights reserved.

Models

Risk adjustment models

- analytic model plug-in engine and creation of analytical-ready data sets
- choice of comorbidity risk factors
  - CMS-HCC, Charlson, Elixhauser
- choice of cost prediction models
  - one-part/two-part models
  - Tweedie, normal, gamma, etc.
- population health focused models
  - readmissions, provider-profiling
- interactive visualization and reporting
  - patient’s journey, cost analysis



Model Add-in Builder and Templates

SWW | SAS Wide Web
SAS® Real World Evider
seamiddev.unx.sas.com:7980/SASRealWorldEvidence/?view=JOB_LIST
SAS® Real World Evidence 4.2
Lina Clover
Navigation › Add-ins
Filter
Download Add-ins...
Upload Add-in

Name	Description
Comparative Effectiveness OR-Model-New	Model for per effectiveness Uploaded 11/
Comparative Nested TE Outcome Model_U	Model for per effectiveness 11/14/2017
Comparative NonNested TE Outcome Model	Model for per effectiveness 11/14/2017
Comparative Time To Event Analysis	Model for per time-to-event
Comparative Treatment Outcomes Model-F	Model for per effectiveness analyses.
Cost Analysis Model	Model for clu costs associat procedures.
Cost Analysis Model_FMM	Model for clu costs associat procedures.
ENH Length of Stay Analysis	Model for pro length of stay
ENH Readmission Analysis	Model for pre for a study co
For i18n test	Add-in crea 2:24:55 AM
FTP Chronic Disease Progression Model	Modeling dis world evidenc
FTP Chronic Risk Factor	Exploring and

seamiddev.unx.sas.com:7980/SASRealWorldEvidence/?view=JOB_LIST

Add-in Builder: Test

Properties | Code | User Interface

Name: Comparative NonNested TE Outcome Model

Description: Model for performing comparative effectiveness and outcomes research 11/14/2017

Test with data from: You can test with any cohort regardless of which cohort type is selected for the add-in.

Workspace: MY Test 800k WKSP Cohort: LAMA Cohort Study

Add-in Cohort Type
- [ ] Population Cohorts
- [x] Index Event Cohorts

Add-in Settings
- [ ] Public
- [ ] Example (for use within the Add-in Builder only)

Download Add-in

Names | Results:

Libnames | Workspace Data | Cohort Data | Macro Variables

Libnames generated for this Add-in

Name	Path	Description
wsdata	/rwe/warehouse/rwe42/filesets/adfe15a3-d637-4d46-99ee-512cbc185f0e	folder containing workspace data
odsout	&ods_temp_folder	temp folder to write ODS output and any JSON files
dataout	&output_folder	permanent folder to write SAS datasets
cohort1	/rwe/warehouse/rwe42/rwe/workspaces/WOKD9YCU/cohorts/CCTLNDKC	Cohort libref

Save | Save and Close | Cancel

agreements as executed by customer and SAS Institute Inc..

Use Cases and Demonstration

- Case Study: Episode-of care based hospital readmissions analytics
  - Find members with higher likelihood of readmission to an acute hospital
  - Find members who have higher costs associated with preventable readmissions
  - Model built at member-level and episode-level

This information is confidential and covered under the terms of any SAS agreements as executed by customer and SAS Institute Inc.

Real World Application

Data Domain
- Profile
- Cleanse
- Episode Definition

- Data Source
  - Random sample of 800K patients from publicly available 2008-2010 CMS Synthetic PUF Medicare data
- Patient Population
  - Created episodes of care data for two medical conditions
  - COPD (n=6,094) and Diabetes (n=13,503)
- Analytic Models
  - Episode cost risk adjustment model
  - Readmission model
  - Predictor variables include patient demographic characteristics and Charlson and Elixhauser comorbidity factors

READMISSIONS ANALYTICS

- Readmission definition
  - readmission to a hospital within 30/45/60 days of a discharge from the same or another hospital
  - target readmission time interval: 30/45/60-day
- Identifying episode-related readmissions
  - anchor claim: IP claim during the study period
  - anchor condition: index condition on the anchor claim
  - subsequent IP claims at an acute facility that occur 2 or more days after the initial stay (or at the same acute facility regardless of the number of days) - assigned to the episode as PAC
  - inpatient stay after an episode is signaled by an outpatient or professional services claim - assigned to the episode as a complication
  - readmissions to an acute care facility within a related open episode - assigned to the episode as PAC
  - transfer to another acute care facility within the same day or the next day following the discharge from the first acute care facility - assigned to the episode as typical
  - transfer to a post-acute (rehabilitation, long-term, or skilled nursing) facility - assigned as typical or typical w/complication based on the diagnoses on the stay.

This information is confidential and covered under the terms of any SAS agreements as executed by customer and SAS Institute Inc..

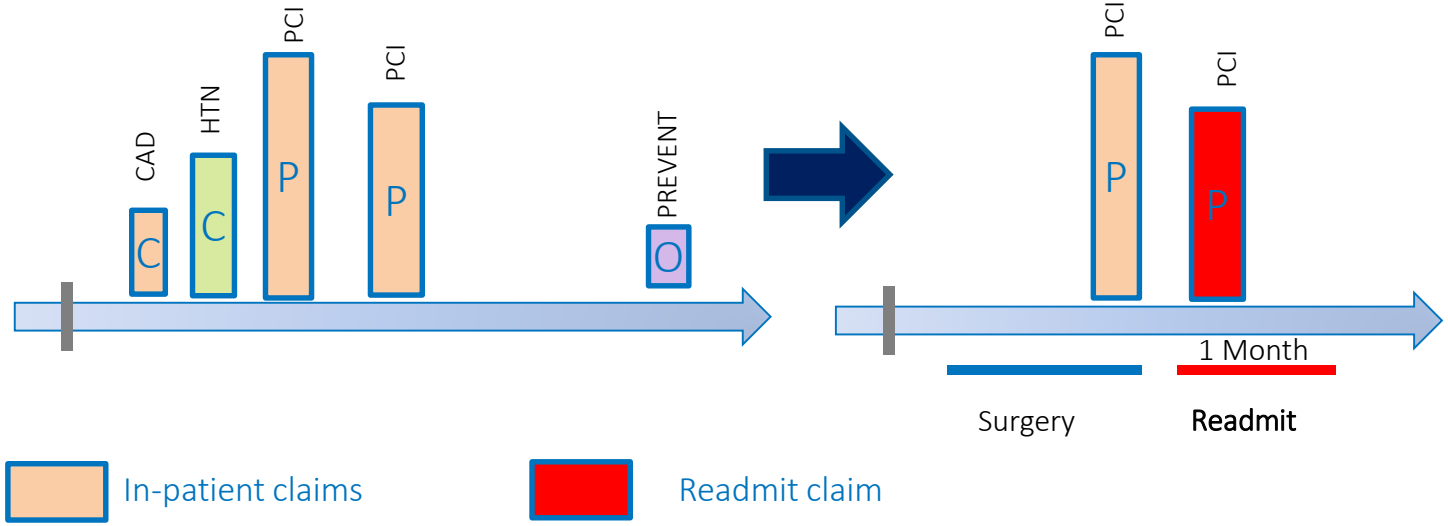
EXAMPLE: A MEMBER READMISSION PROFILE

Route: Journey

- Index Cohort
- Episode Construction
- Care Pathways

Member ID=0000377379 Episodes during Study Period

Member ID	Episode ID	Condition	Episode Start Date	Episode End Date
0000377379	38811	LBP	September 01, 2009	March 16, 2012
0000377379	38814	URI	November 05, 2009	November 26, 2009
0000377379	38809	DEPRSN	January 29, 2010	March 16, 2012
0000377379	38812	PREVNT	September 06, 2011	March 16, 2012
0000377379	38813	RESPFL	September 23, 2011	October 24, 2011
0000377379	38808	COPD	November 01, 2011	March 16, 2012
0000377379	38810	DIAB	November 01, 2011	March 16, 2012
0000377379	38807	CDIFF	November 07, 2011	November 09, 2011



This information is confidential and covered under the terms of any SAS agreements as executed by customer and SAS Institute Inc..

Exploration: Risk Factors

Exploration
- Frequencies
- Correlations
- Categories

- Frequency and percentage distribution of demographic and comorbidity risk factors

Characteristics		Diabetes		COPD	
		N	%	N	%
Number of patients / episodes		13503	100	6094	100
Gender	Female	7078	52.4	3259	53.5
	Male	6425	47.8	2835	46.5
Age Group	55+	7258	53.8	3222	52.9
	<55	6245	46.2	2872	47.1
History of:					
Cancer	Yes	886	6.6	466	7.7
Cerebrovascular disease	Yes	592	4.3	365	6.0
Congestive heart failure	Yes	646	4.8	411	6.7
Chronic pulmonary disease	Yes	2306	17.1	979	16.1
Dementia	Yes	26	0.2	19	0.3
AIDS/HIV	Yes	206	1.5	124	2.0
Myocardial infarction	Yes	422	3.1	242	4.0
Mild liver disease	Yes	620	4.6	330	5.4
Moderate/severe liver disease	Yes	132	1.0	78	1.3
Paraplegia and hemiplegia	Yes	68	0.5	65	1.1
Periphral vascular disease	Yes	1081	8.0	609	10.0
Renal disease	Yes	905	6.7	472	7.8
Connective tissue/rheumatic disease	Yes	1212	9.0	607	10.0

Actual and Risk-adjusted Episode Costs

Medical Condition	Label	N	Mean	Median	Std Dev
COPD	Episode Cost	6094	2325.621495	1015.000000	4869.380205
COPD	Risk-adjusted Episode Cost	6094	2360.600392	1240.289186	3990.483269
DIAB	Episode Cost	13503	3201.752768	1475.000000	6089.271173
DIAB	Risk-adjusted Episode Cost	13503	3253.273150	1788.974488	5384.328070

- Among patients with COPD episodes, unadjusted and adjusted mean episode costs were $2325 and $2360, respectively.
- For Diabetes, an unadjusted and adjusted cost of $3201 and cost of $3253.

Hospitalization and Readmission

Condition	Gender	Age Group	Number of Episodes	Number of Admissions	Number of Readmissions	Percent of Episodes	Percent of Admissions	Percent of Readmissions	Percent of Admissions in Group	Percent of Readmissions in Group
DIAB	F	55+	3939	611	7	29.17	27.72	21.88	15.51	1.15
DIAB	M	55+	3319	548	12	24.58	24.86	37.50	16.51	2.19
DIAB	F	<55	3139	536	7	23.25	24.32	21.88	17.08	1.31
DIAB	M	<55	3106	509	6	23.00	23.09	18.75	16.39	1.18
DIAB			13503	2204	32	100.00	100.00	100.00	16.32	1.45
COPD	F	55+	1789	327	1	29.36	28.07	8.33	18.28	0.31
COPD	M	55+	1433	288	4	23.51	24.72	33.33	20.10	1.39
COPD	F	<55	1470	285	4	24.12	24.46	33.33	19.39	1.40
COPD	M	<55	1402	265	3	23.01	22.75	25.00	18.90	1.13
COPD			6094	1165	12	100.00	100.00	100.00	19.12	1.03

- Total admissions were 2204 (16.3%) and 1165 (19.1%) in Diabetes and COPD cohorts, respectively.
- About 1.5% and 1% resulted in 30-day readmissions based on total admissions
- Higher percentages of readmissions were observed among males aged 55+

Readmission Model

Analytics
- Machine Learning
- Logistic Regression
- Risk Segmentation

- Fitted logistic regression model
- Notable findings include:
- For Diabetes cohort - male gender, older age group, myocardial infarction, pulmonary disease, renal disease but not significant
- For COPD cohort – renal disease, cancer elevated risk of readmission, but not significant.

Logistic Regression Model Estimation

Condition	Risk Factor Type	Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Odds Ratio Estimate	Lower 95% Confidence Limit for Odds Ratio	Upper 95% Confidence Limit for Odds Ratio	RISK_FACTOR_GROUP_DESC
DIAB	CHAR	CHAR_CANCER	-0.8207	1.0359	0.6276	0.4282	0.440	0.058	3.352	Cancer
DIAB	CHAR	CHAR_CHF	-0.0596	0.7650	0.0061	0.9379	0.942	0.210	4.220	Congestive Heart Failure
DIAB	CHAR	CHAR_COPD	0.4553	0.4338	1.1017	0.2939	1.577	0.674	3.690	Chronic Pulmonary Disease
DIAB	CHAR	CHAR_HIV	0.7455	1.0504	0.5036	0.4779	2.107	0.269	16.515	AIDS/HIV
DIAB	CHAR	CHAR_MI	0.5294	0.7663	0.4774	0.4896	1.698	0.378	7.624	Myocardial Infarction
DIAB	CHAR	CHAR_PVD	0.3144	0.5645	0.3102	0.5776	1.369	0.453	4.140	Peripheral Vascular Disease
DIAB	CHAR	CHAR_RD	0.0277	0.5603	0.0024	0.9606	1.028	0.343	3.083	Renal Disease
DIAB	CHAR	CHAR_Rheum	-1.0741	1.0316	1.0841	0.2978	0.342	0.045	2.580	Connective Tissue Disease/Rheumatic Disease
DIAB	CHAR	Intercept	-4.3259	0.2237	374.0908	<.0001	.	.		
DIAB	CHAR	agegp 55+	0.1487	0.1818	0.6693	0.4133	1.346	0.660	2.746	Age Group
DIAB	CHAR	sex F	-0.1661	0.1804	0.8475	0.3573	0.717	0.354	1.455	Gender

Logistic Regression Model Estimation

Condition	Risk Factor Type	Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Odds Ratio Estimate	Lower 95% Confidence Limit for Odds Ratio	Upper 95% Confidence Limit for Odds Ratio	RISK_FACTOR_GROUP_DESC
COPD	CHAR	CHAR_CANCER	0.3872	1.0651	0.1321	0.7162	1.473	0.183	11.880	Cancer
COPD	CHAR	CHAR_COPD	0.0542	0.7870	0.0047	0.9451	1.056	0.226	4.937	Chronic Pulmonary Disease
COPD	CHAR	CHAR_RD	-0.2409	1.0587	0.0518	0.8200	0.786	0.099	6.260	Renal Disease
COPD	CHAR	Intercept	-4.6044	0.3394	184.0936	<.0001	.	.	.	
COPD	CHAR	agegp 55+	-0.2293	0.2954	0.6027	0.4376	0.632	0.199	2.012	Age Group
COPD	CHAR	sex F	-0.2145	0.2947	0.5300	0.4666	0.651	0.205	2.067	Gender

Cost Model

- Fitted generalized gamma regression model
- Notable predictors of cost include:
  - lower age group
  - length of episode duration
  - patients with episode complications
  - congestive heart failure
  - chronic pulmonary disease
  - renal disease

Medical Condition	Model Parameter	Model Estimate	Lower Confidence Limit	Upper Confidence Limit	Pr > Chi-Square
DIAB	Intercept	4.76988	4.64708	4.89269	0.00000
	Sex (Female)	-0.02788	-0.06571	0.00995	0.14857
	age_group 55+	-0.06117	-0.09941	-0.02292	0.00172
	age_group <55	0.00000	.	.	.
	died_ind	0.97329	0.67127	1.27532	0.00000
	Enrolment status > 180 days	0.15124	0.09643	0.20604	0.00000
	Episode duration (months)	0.05669	0.05285	0.06053	0.00000
	Episode complication status (Y)	0.61949	0.50833	0.73066	0.00000
	Care setting (institutional/non-institutional)	0.89133	0.82995	0.95270	0.00000
	Cancer	-0.00072	-0.07792	0.07649	0.98544
	Cerebrovascular disease	0.01950	-0.07612	0.11512	0.68939
	Congestive heart failure	0.18259	0.09067	0.27452	0.00010
	Chronic pulmonary disease	0.11207	0.05787	0.16627	0.00005
	Dementia	0.10474	-0.30980	0.51928	0.62046
	AIDS/HIV	0.30805	0.15277	0.46334	0.00010
	Myocardial infarction	0.03894	-0.07283	0.15070	0.49473
	Mild liver disease	0.37278	0.27606	0.46950	0.00000
	Moderate/severe liver disease	-0.03149	-0.22325	0.16028	0.74760
	Paraplegia and hemiplegia	0.34370	0.07400	0.61340	0.01250
	Periphral vascular disease	0.24064	0.16808	0.31319	0.00000
	Renal disease	0.78797	0.70851	0.86743	0.00000
	Connective tissue/rheumatic disease	-0.00413	-0.07331	0.06506	0.90693
	Dispersion	0.87591	0.85400	0.89838	.
	Power	2.00000	.	.	.

Conclusion

WHAT

Observational Data

Early detection of drug safety issues

Patient healthcare journeys

HOW

Cohort and Episode Builder

Investigate and answer real world questions

Discover and visualize patient cohort

EXECUTE

Risk Assessment & Prediction

Proactively detect at-risk patients for hospitalization readmission, and length of stay

Future forward with SAS.

Digitize the world of pharma.
Shape the future of analytics.
Accelerate your career.

This information is confidential and covered under the terms of any SAS agreements as executed by customer and SAS Institute Inc.