Paper SAS2051-2018

Cost Analysis in Population Health Using SAS® Real World Evidence

Jay Paulson, SAS Institute Inc.

ABSTRACT

Health care organizations are faced with the challenge of analyzing high volumes of population-level health data to gain insights into ways to reduce cost while improving health care quality and efficiency. This paper explains how to develop statistical models using SAS® Real World Evidence to analyze cost and understand the key factors associated with health care costs. Using SAS Real World Evidence, a population cohort based on a research question is created. Statistical modeling is performed on this cohort using the powerful Add-in Builder available in SAS Real World Evidence. The results from statistical models are then analyzed to understand characteristics associated with cost.

INTRODUCTION

SAS Real World Evidence provides a platform for management, analysis, and visualization of real-world population-level health data. This platform provides you with the capability to rapidly query observational data, apply built-in analytics, and find insights that demonstrate value. The three pillars of SAS Real World Evidence are:

1. Management of data assets
2. Identifying patient cohorts of interest
3. Knowledge and insights via analytics

This paper is focused mainly on the third pillar of SAS Real World Evidence, that is, deriving insights using analytics.

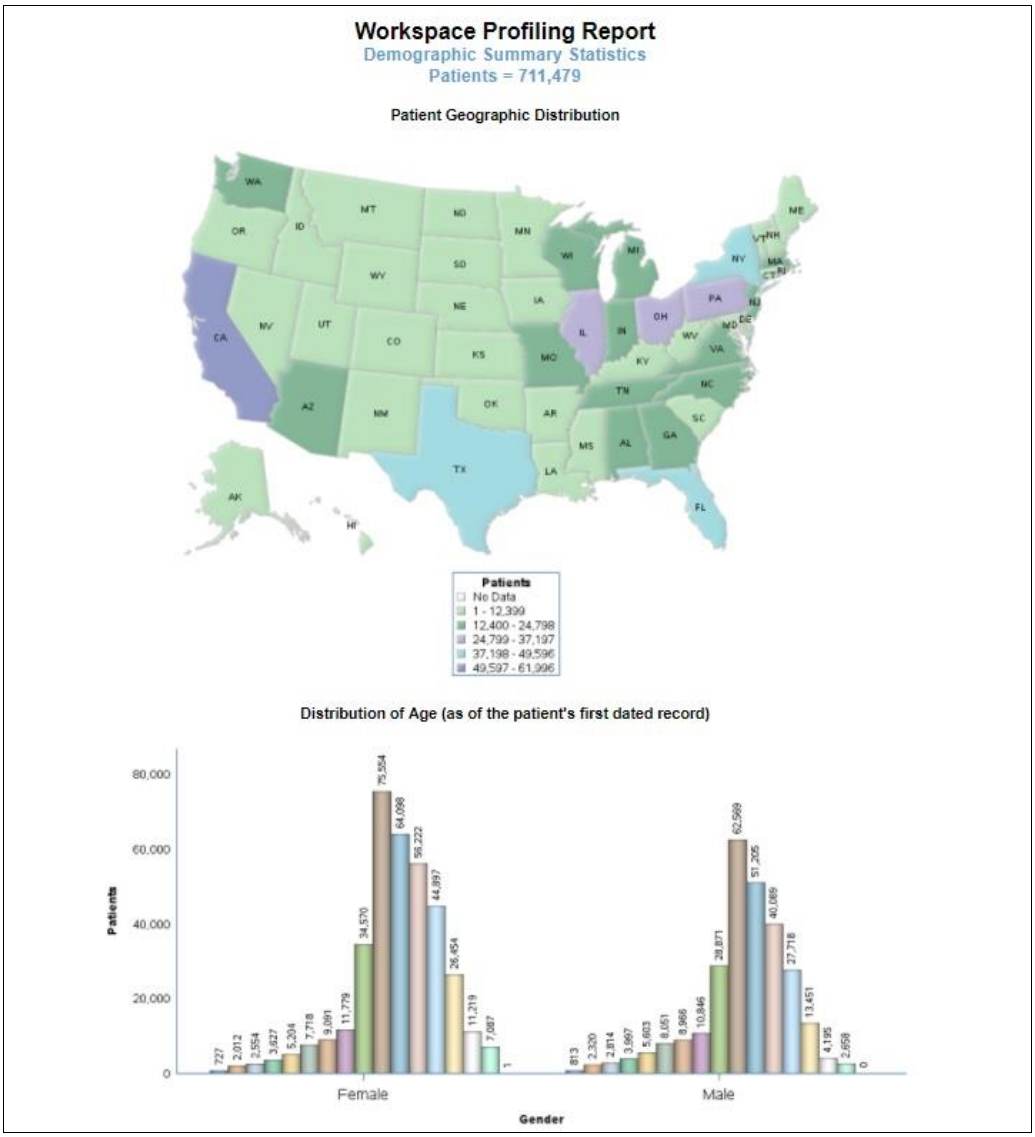
USING SAS REAL WORLD EVIDENCE

As an analyst, you are interested in gaining insights into ways to reduce cost while improving health care quality and efficiency. Let us look at the example of analyzing the medical cost of treating patients who have been diagnosed with diabetes and have been administered insulin. After the data is collected and made available to SAS Real World Evidence, you must use a workspace for creating and analyzing cohorts.

WORKSPACE DATA

After creating a workspace, the built-in workspace profiling report is a great place to start your analysis as this gives a clear idea about the population you are looking at. For example, Figure 1 below shows the demographic distribution of the workspace called “Medicare Population.” This workspace has a cohort of 711,479 patients.

Figure 1: Workspace Profiling - Demographics



POPULATION COHORT ANALYSIS

All SAS Real World Evidence analyses are performed on cohorts, which are typically subsets of the workspace data that are created based on the population of interest (based on your research question). For example, our population of interest is patients who have been diagnosed with diabetes and have been administered insulin.

A population cohort called “Diabetes Cohort” is created. You have instant access to the cohort profile reports. For example, Figure 2 shows the demographic distribution of the cohort. The population size has been reduced to 12,097 patients with a mean age of 70 at the start of study.

Figure 2: Cohort Profile - Demographic Distribution

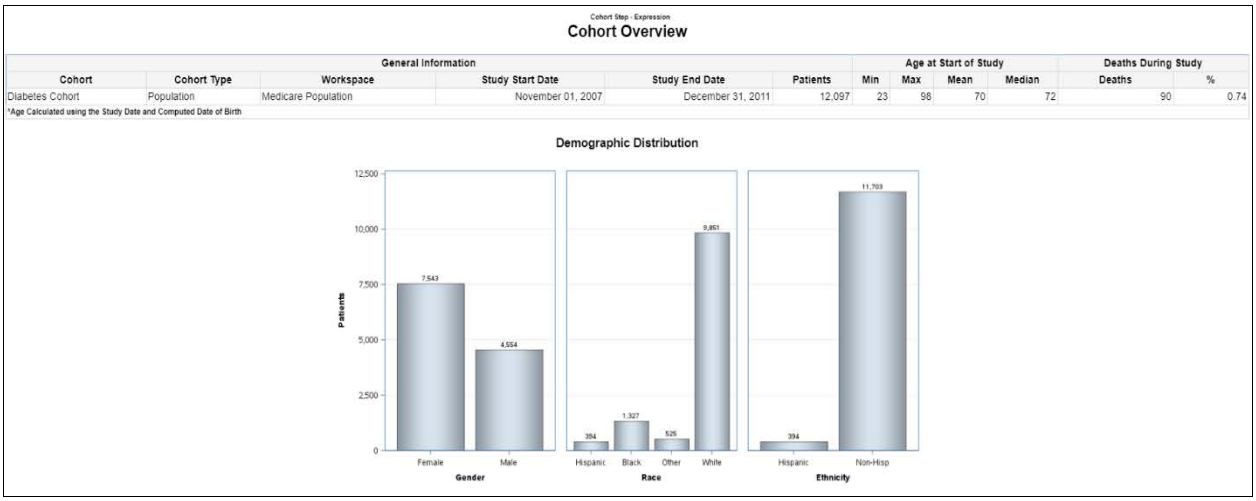


Figure 3 shows the age profile of the cohort at the beginning of study, grouped by gender, race, and ethnicity.

Figure 3: Cohort Profile - Age at Beginning of Study

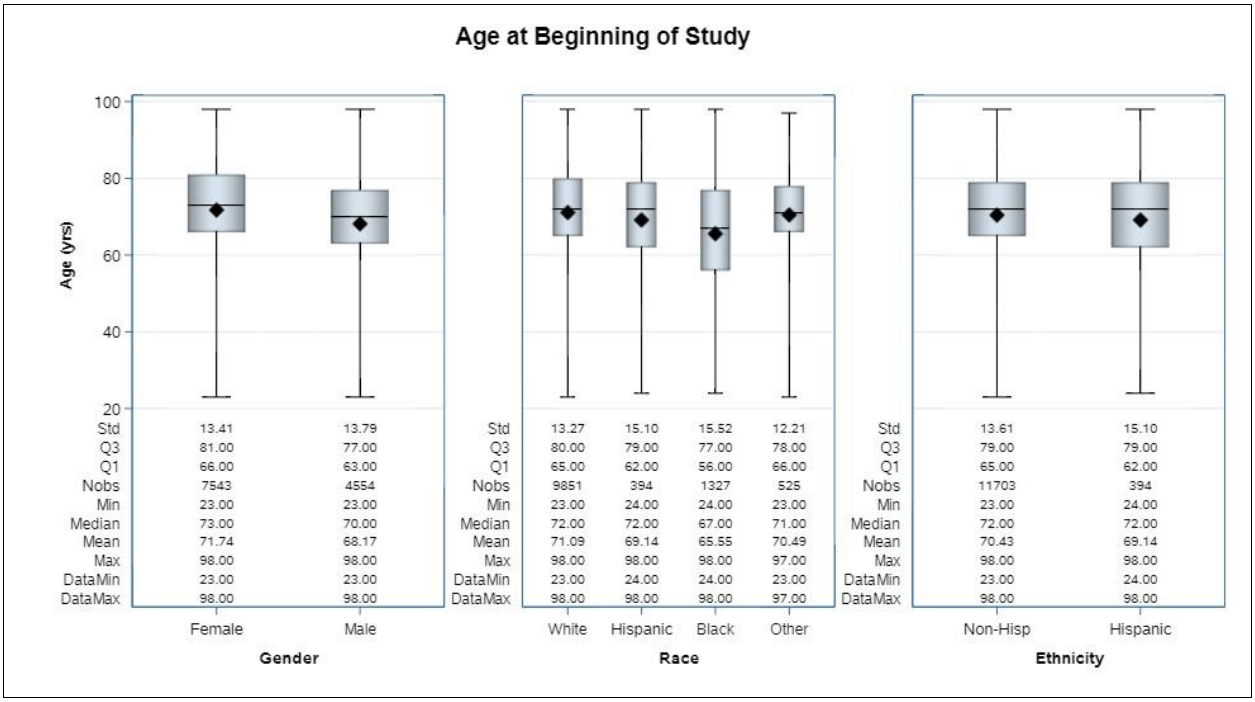
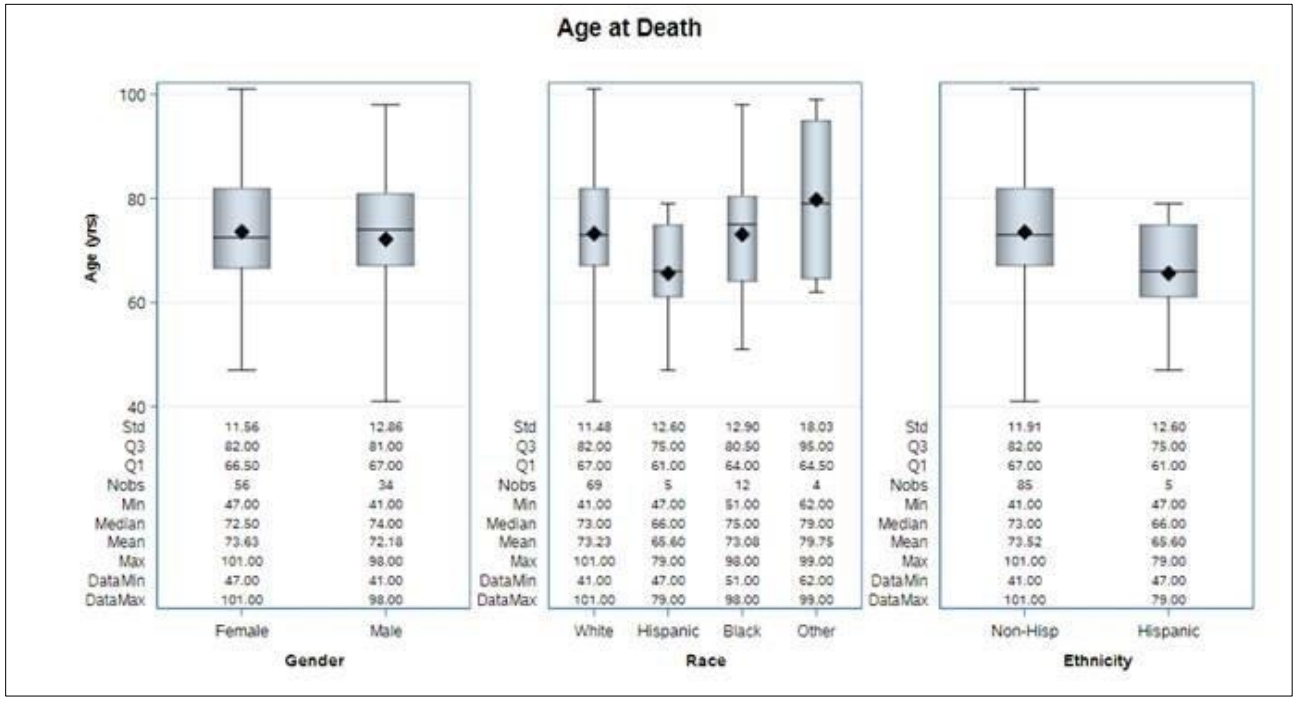


Figure 4 shows the age profile of the cohort at death, grouped by gender, race, and ethnicity.

Figure 4: Cohort Profile - Age at Death



After creating the population cohort, you are ready to use the Cost Analysis Model add-in, one of several statistical models that are available with SAS Real World Evidence to provide you with in-depth insights and answers for your research questions.

COST ANALYSIS STATISTICAL MODEL

The purpose of the Cost Analysis Model add-in is to perform cost analysis for a cohort of patients. The following questions and objectives are addressed by this add-in:

- Identify clusters of patients based on cost
- Compare characteristics of “high cost” patient clusters with other clusters
- Predict risk-adjusted expected costs for patients
- Compare expected costs with actual costs
- Analyze characteristics of patients with high cost variations from the expected costs

The following covariates and predictor variables are used by the cost model:

- Demographic variables
- Risk factors

PROCEDURE USED FOR COST ANALYSIS

Cost analysis is performed using the HPFMM statistical procedure. HPFMM is a high-performance counterpart to the FMM (Finite Mixture Models) procedure. The FMM procedure fits statistical models to data for which the distribution of the response is a finite mixture of univariate distributions - that is, each response comes from one of several random univariate distributions that have unknown probabilities. The HPFMM procedure is designed to fit finite mixtures of regression models or finite mixtures of generalized linear models in which the covariates and regression structure can be the same across components or can be different.

DISCLAIMER ON DATA USED FOR COST ANALYSIS

This paper uses CMS 2008-2010 Data Entrepreneurs' Synthetic Public Use File (DE-SynPUF) data for cost analysis. Because of the synthetic process used to generate DE-SynPUF data to eliminate disclosure risk, the results from the output tables given in this paper must not be used to draw inferences or conclusions. The objective of this paper is to highlight the statistical modeling features of SAS Real World Evidence and not to highlight insights gained from DE-SynPUF data.

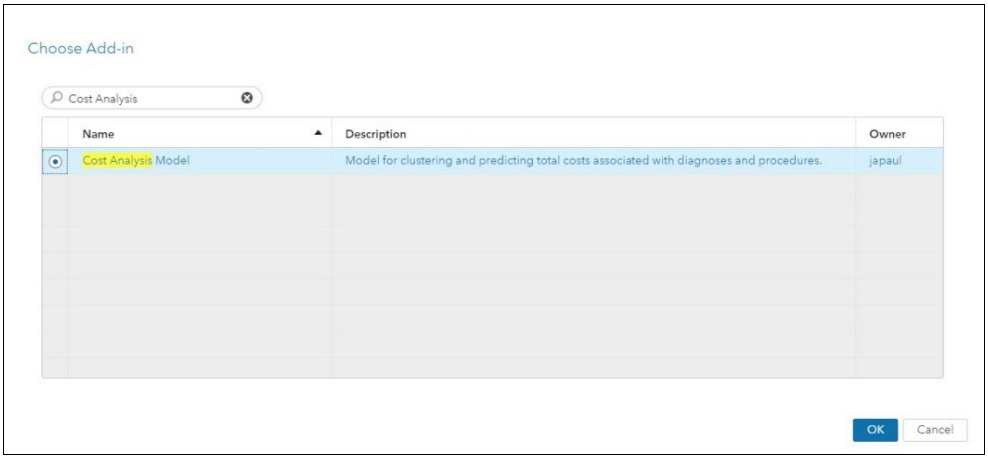
DEFINITION OF COST

This paper uses 'Total Paid Amount' for cost analysis. This is typically the total of the amount paid by payers, the deductible paid by the patient, and the amount paid by co-insurance (if any).

RUNNING COST ANALYSIS MODEL FROM A COHORT

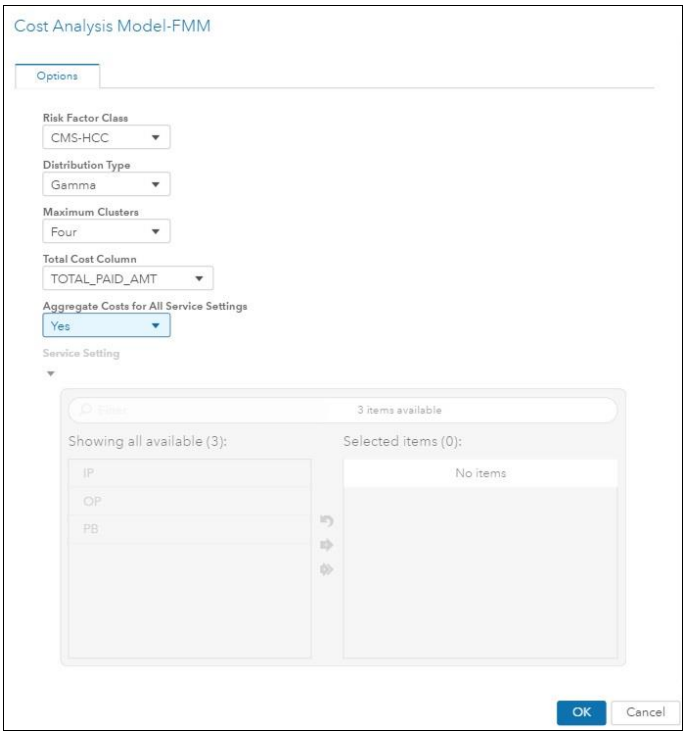
The Cost Analysis Model can be added to your cohort by using the "Add-in Jobs" functionality as shown in Figure 5. After the model is added to your cohort, you can customize the model using the options window.

Figure 5: Add Cost Analysis Model



There are six input parameter values you must specify in the options window before running the model. These parameters are described in detail in Table 1.

Figure 6: Specifying Input Parameters for the Cost Analysis Model Add-in



DESCRIPTION OF INPUT PARAMETERS

Table 1 lists the input parameters you must specify in the options window for running the Cost Analysis Model add-in. The values for two parameters (TOTAL_COST_VARIABLE and SERVICE_SETTING_ALL, which contain values that are dynamically populated using the Services table) are required to run the model. The other four parameters use default values if you do not specify values.

Table 1: Input Parameters for the Cost Analysis Model Add-in

Parameter	Description
RISK_FACTOR_CLASS	Specifies the Risk/Comorbidity Factors Group Type. Possible values (different types of risk factor types) are: • **STANDARD**: Risk factor based on HCI3 metadata definitions • **CMS-HCC**: Risk factor based on CMS-HCC risk factor class • **CCS**: Risk factor based on AHRQ-Clinical Condition Code Set • **CHARLSON**: Risk factor based on Charlson comorbidity factors • **LACE**: Risk factor based on Lace comorbidity factors • **ELIXHAUSER**: Risk factor based on Elixhauser comorbidity factors
DISTRIBUTION_TYPE	Specifies the distribution type. Possible values are: • NORMAL • GAMMA
MAX_CLUSTERS	Specifies the maximum number of clusters to be used in the clustering analysis.
TOTAL_COST_VARIABLE	Specifies the cost column, which is obtained dynamically from the Service table. This column is used for cost calculations.
AGGREGATE_COSTS	Indicates whether costs should be aggregated for all service setting types. If this is 'Yes', then the parameter SERVICE_SETTING_ALL is not applicable. If this is 'No', then you must specify one or more values for SERVICE_SETTING_ALL parameter.
SERVICE_SETTING_ALL	Contains values that are selected dynamically from the SERVICE_SETTING column in the Services table. Some examples of values are IP (inpatient) and OP (outpatient). This is required only if AGGREGATE_COSTS is set to 'No'.

COST ANALYSIS OUTPUTS

The Cost Analysis Model add-in generates reports and SAS data sets as output. You can use the output data sets generated by the model to perform further analysis using SAS® Visual Analytics.

The output discussed in this paper are built-in ODS reports that are generated by the Cost Analysis Model add-in. Figure 7 shows the cost analysis output window. There are two types of output you can view in the output window: reports and output data sets.

In the next section we will look at some of this output in detail.

Figure 7: Cost Analysis Output

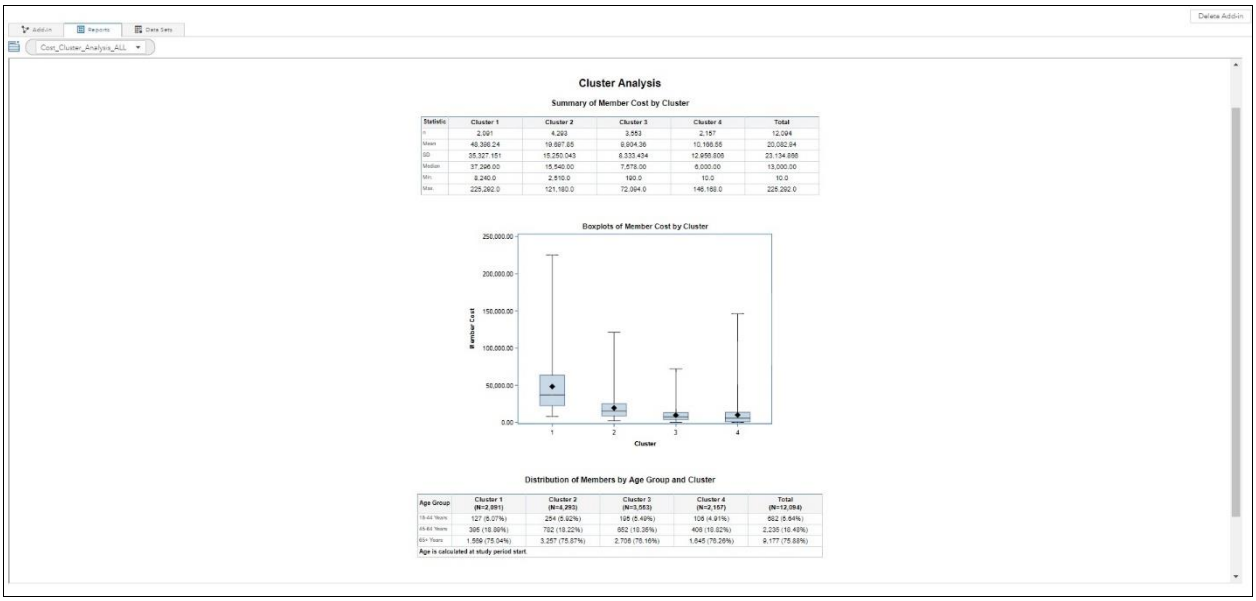
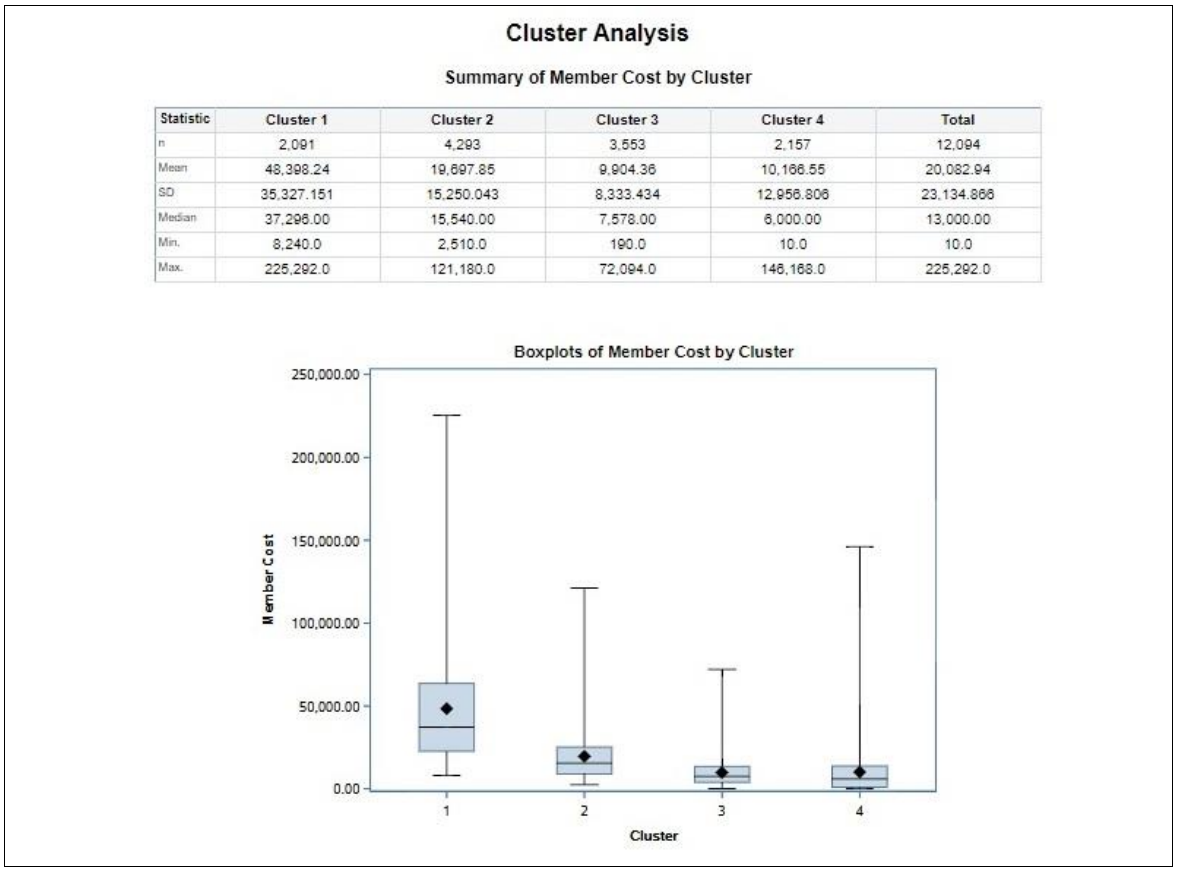


Figure 8 shows the cost clusters that are generated by the model. The model has identified four patient clusters based on cost, demographics, and risk factors. Cluster 1, with the highest mean cost of $48,398.24, probably needs further investigation.

Figure 8: Cost Cluster Information



Cluster Analysis

Summary of Member Cost by Cluster

Statistic	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Total
n	2,091	4,293	3,553	2,157	12,094
Mean	48,398.24	19,697.85	9,904.36	10,166.55	20,082.94
SD	35,327.151	15,250.043	8,333.434	12,956.806	23,134.866
Median	37,296.00	15,540.00	7,578.00	6,000.00	13,000.00
Min.	8,240.0	2,510.0	190.0	10.0	10.0
Max.	225,292.0	121,180.0	72,094.0	146,168.0	225,292.0

Figure 9 shows the demographic details of all four clusters. This cohort is demographically very similar.

Figure 9: Demographic Information by Cluster

Distribution of Members by Age Group and Cluster

Age Group	Cluster 1 (N=2,091)	Cluster 2 (N=4,293)	Cluster 3 (N=3,553)	Cluster 4 (N=2,157)	Total (N=12,094)
18-44 Years	127 (6.07%)	254 (5.92%)	195 (5.49%)	106 (4.91%)	682 (5.64%)
45-64 Years	395 (18.89%)	782 (18.22%)	652 (18.35%)	406 (18.82%)	2,235 (18.48%)
65+ Years	1,569 (75.04%)	3,257 (75.87%)	2,706 (76.16%)	1,645 (76.26%)	9,177 (75.88%)
Age is calculated at study period start.					

Distribution of members by gender and cluster

Gender	Cluster 1 (N=2,091)	Cluster 2 (N=4,293)	Cluster 3 (N=3,553)	Cluster 4 (N=2,157)	Total (N=12,094)
Female	1,290 (61.69%)	2,693 (62.73%)	2,218 (62.43%)	1,339 (62.08%)	7,540 (62.34%)
Male	801 (38.31%)	1,600 (37.27%)	1,335 (37.57%)	818 (37.92%)	4,554 (37.66%)

Figure 10 shows the most frequent risk/comorbidity factors across all 4 clusters. These are the 38 risk factors identified by the model to be significant.

Figure 10: Risk/Comorbidity Factor Groups across Clusters

Obs	Risk/Comorbidity Factor Group	Description	Cluster_1	Cluster_2	Cluster_3	Cluster_4
1	HCC1	HIV/AIDS	Y	Y	Y	Y
2	HCC10	LYMPHOMA AND OTHER CANCERS	Y	Y	Y	Y
3	HCC100	ISCHEMIC OR UNSPECIFIED STROKE	Y	Y	Y	Y
4	HCC103	HEMIPLEGIA/HEMIPARESIS	Y	Y	Y	Y
5	HCC104	MONOPLEGIA, OTHER PARALYTIC SYNDROMES	Y	Y	Y	Y
6	HCC106	ATHEROSCLEROSIS OF THE EXTREMITIES WITH ULCERATION OR GANGRENE	Y	Y	Y	Y
7	HCC107	VASCULAR DISEASE WITH COMPLICATIONS	Y	Y	Y	Y
8	HCC11	COLORECTAL, BLADDER, AND OTHER CANCERS	Y	Y	Y	Y
9	HCC110	CYSTIC FIBROSIS	Y	Y	Y	Y
10	HCC114	ASPIRATION AND SPECIFIED BACTERIAL PNEUMONIAS	Y	Y	Y	Y
11	HCC135	ACUTE RENAL FAILURE	Y	Y	Y	Y
12	HCC136	CHRONIC KIDNEY DISEASE, STAGE 5	Y	Y	Y	Y
13	HCC137	CHRONIC KIDNEY DISEASE, SEVERE (STAGE 4)	Y	Y	Y	Y
14	HCC161	CHRONIC ULCER OF SKIN, EXCEPT PRESSURE	Y	Y	Y	Y
15	HCC170	HIP FRACTURE/DISLOCATION	Y	Y	Y	Y
16	HCC173	TRAUMATIC AMPUTATIONS AND COMPLICATIONS	Y	Y	Y	Y
17	HCC176	COMPLICATIONS OF SPECIFIED IMPLANTED DEVICE OR GRAFT	Y	Y	Y	Y
18	HCC186	MAJOR ORGAN TRANSPLANT OR REPLACEMENT STATUS	Y	Y	Y	Y
19	HCC189	AMPUTATION STATUS, LOWER LIMB/AMPUTATION COMPLICATIONS	Y	Y	Y	Y
20	HCC2	SEPTICEMIA, SEPSIS, SYSTEMIC INFLAMMATORY RESPONSE SYNDROME/SHOCK	Y	Y	Y	Y
21	HCC21	PROTEIN-CALORIE MALNUTRITION	Y	Y	Y	Y
22	HCC22	MORBID OBESITY	Y	Y	Y	Y
23	HCC23	OTHER SIGNIFICANT ENDOCRINE AND METABOLIC DISORDERS	Y	Y	Y	Y
24	HCC27	END-STAGE LIVER DISEASE	Y	Y	Y	Y
25	HCC29	CHRONIC HEPATITIS	Y	Y	Y	Y
26	HCC33	INTESTINAL OBSTRUCTION/PERFORATION	Y	Y	Y	Y
27	HCC34	CHRONIC PANCREATITIS	Y	Y	Y	Y
28	HCC39	BONE/JOINT/MUSCLE INFECTIONS/NECROSIS	Y	Y	Y	Y
29	HCC47	DISORDERS OF IMMUNITY	Y	Y	Y	Y
30	HCC55	DRUG/ALCOHOL DEPENDENCE	Y	Y	Y	Y
31	HCC70	QUADRIPLEGIA	Y	Y	Y	Y
32	HCC8	METASTATIC CANCER AND ACUTE LEUKEMIA	Y	Y	Y	Y
33	HCC82	RESPIRATOR DEPENDENCE/TRACHEOSTOMY STATUS	Y	Y	Y	Y
34	HCC83	RESPIRATORY ARREST	Y	Y	Y	Y
35	HCC84	CARDIO-RESPIRATORY FAILURE AND SHOCK	Y	Y	Y	Y
36	HCC86	ACUTE MYOCARDIAL INFARCTION	Y	Y	Y	Y
37	HCC88	ANGINA PECTORIS	Y	Y	Y	Y
38	HCC9	LUNG AND OTHER SEVERE CANCERS	Y	Y	Y	Y

Figure 11 shows the top 25 high risk patients (that is,, patients with highest actual cost to expected cost ratio). The patient in row 1 has an actual cost to expected cost ratio of 12.92, which means the patient's actual cost was 12.92 times the risk-adjusted expected cost.

Figure 11:Top 25 High Risk Patients

Obs	Unique Member ID	Cluster	Age at Study Start	Gender	Actual Cost	Expected Cost (Risk Adjusted)	Actual to Expected Cost Ratio
1	0C94C6010B497603	1	69	Female	93,850.00	7,262.47	12.92
2	3A7417B66CDFDD54	1	84	Male	55,600.00	4,675.23	11.89
3	639C9688BD70010F	1	81	Female	119,304.00	13,091.89	9.11
4	91268DFD712F034A	1	48	Male	54,230.00	6,173.59	8.78
5	6D2C2256205E11F7	1	88	Female	62,000.00	7,887.89	7.86
6	C6B6283581CB3E63	1	71	Female	66,070.00	8,625.12	7.66
7	A1B165335CB8642F	1	64	Male	56,758.00	7,448.84	7.62
8	928069B4BA392A65	1	64	Male	66,768.00	8,900.40	7.50
9	A640DABD0F90D097	1	71	Female	52,620.00	7,047.18	7.47
10	A45E4AFA3FD9BFE1	1	54	Female	44,806.00	6,001.69	7.47
11	F54B0C8B7D5C0A75	1	62	Female	63,068.00	8,561.83	7.37
12	EC81376D6E5A381E	1	72	Female	69,168.00	9,430.01	7.33
13	21E96AA6FC7B8F74	1	72	Female	60,496.00	8,257.37	7.33
14	2F3D708B6CFE9B5F	1	67	Male	59,222.00	8,134.80	7.28
15	9B42396B6A72FBBF	1	37	Female	64,378.00	9,025.47	7.13
16	E382FDDB301A20E7	1	76	Male	44,400.00	6,427.17	6.91
17	51865CB79303106E	1	65	Female	59,550.00	8,644.77	6.89
18	976274E4711515BC	1	65	Female	39,838.00	5,793.63	6.88
19	16A6DAE55FC1EDC9	1	47	Female	101,420.00	14,750.12	6.88
20	4BC7CDC3216242A2	1	51	Female	54,960.00	8,047.82	6.83
21	C9B3EA289449D187	1	81	Male	71,566.00	10,490.81	6.82
22	DFB12C19E92AAC02	1	78	Female	39,730.00	5,857.28	6.78
23	1E6A4E7B1A165D0E	1	74	Male	31,604.00	4,675.23	6.76
24	6C1E46EEDE804504	1	73	Male	45,714.00	6,881.69	6.64
25	79CDE919CC5C24CF	1	74	Male	67,438.00	10,177.97	6.63

Figure 12 shows the most frequent diagnosis codes of the top 10% of high risk patients. There are conditions like Hyperlipidemia, Hypertension, and Heart Disease contributing to the high risk in cluster 1.

Figure 12: Most Frequent Diagnosis Codes of High Risk Patients

Obs	Diagnosis	Code Type	Code	Cluster_1	Cluster_3	Cluster_4
1	ANEMIA, UNSPECIFIED	ICD-10 Diagnosis	D649	766 (6.33%)	8 (0.07%)	4 (0.03%)
2	HYPOTHYROIDISM, UNSPECIFIED	ICD-10 Diagnosis	E039	886 (7.32%)	8 (0.07%)	N/A
3	TYPE 2 DIABETES MELLITUS WITHOUT COMPLICATIONS	ICD-10 Diagnosis	E119	1090 (9.01%)	10 (0.08%)	N/A
4	PURE HYPERCHOLESTEROLEMIA	ICD-10 Diagnosis	E780	960 (7.94%)	10 (0.08%)	4 (0.03%)
5	MIXED HYPERLIPIDEMIA	ICD-10 Diagnosis	E782	702 (5.80%)	N/A	4 (0.03%)
6	OTHER HYPERLIPIDEMIA	ICD-10 Diagnosis	E784	1123 (9.28%)	10 (0.08%)	4 (0.03%)
7	HYPERLIPIDEMIA, UNSPECIFIED	ICD-10 Diagnosis	E785	1123 (9.28%)	10 (0.08%)	4 (0.03%)
8	HYPOKALEMIA	ICD-10 Diagnosis	E876	N/A	N/A	4 (0.03%)
9	POSTVIRAL FATIGUE SYNDROME	ICD-10 Diagnosis	G933	787 (6.51%)	9 (0.07%)	N/A
10	ESSENTIAL (PRIMARY) HYPERTENSION	ICD-10 Diagnosis	I10	1183 (9.78%)	10 (0.08%)	4 (0.03%)
11	ATHEROSCLEROTIC HEART DISEASE OF NATIVE CORONARY ARTERY WITHOUT ANGINA PECTORIS	ICD-10 Diagnosis	I2510	1001 (8.27%)	10 (0.08%)	4 (0.03%)
12	UNSPECIFIED ATRIAL FIBRILLATION	ICD-10 Diagnosis	I4891	881 (7.28%)	10 (0.08%)	4 (0.03%)
13	HEART FAILURE, UNSPECIFIED	ICD-10 Diagnosis	I509	794 (6.56%)	N/A	N/A
14	PNEUMONIA, UNSPECIFIED ORGANISM	ICD-10 Diagnosis	J189	N/A	8 (0.07%)	N/A
15	CHRONIC OBSTRUCTIVE PULMONARY DISEASE, UNSPECIFIED	ICD-10 Diagnosis	J449	809 (6.69%)	8 (0.07%)	4 (0.03%)
16	GASTRO-ESOPHAGEAL REFLUX DISEASE WITHOUT ESOPHAGITIS	ICD-10 Diagnosis	K219	743 (6.14%)	N/A	N/A
17	OSTEOARTHRITIS OF KNEE, UNSPECIFIED	ICD-10 Diagnosis	M179	N/A	N/A	3 (0.02%)
18	UNSPECIFIED OSTEOARTHRITIS, UNSPECIFIED SITE	ICD-10 Diagnosis	M1990	706 (5.84%)	N/A	3 (0.02%)
19	PAIN IN UNSPECIFIED KNEE	ICD-10 Diagnosis	M25569	N/A	N/A	3 (0.02%)
20	OTHER CERVICAL DISC DEGENERATION, UNSPECIFIED CERVICAL REGION	ICD-10 Diagnosis	M5030	N/A	N/A	3 (0.02%)
21	RADICULOPATHY, THORACIC REGION	ICD-10 Diagnosis	M5414	N/A	N/A	3 (0.02%)
22	RADICULOPATHY, THORACOLUMBAR REGION	ICD-10 Diagnosis	M5415	N/A	N/A	3 (0.02%)
23	RADICULOPATHY, LUMBAR REGION	ICD-10 Diagnosis	M5416	N/A	N/A	3 (0.02%)
24	RADICULOPATHY, LUMBOSACRAL REGION	ICD-10 Diagnosis	M5417	N/A	N/A	3 (0.02%)
25	PAIN IN UNSPECIFIED LIMB	ICD-10 Diagnosis	M79609	713 (5.89%)	N/A	N/A
26	URINARY TRACT INFECTION, SITE NOT SPECIFIED	ICD-10 Diagnosis	N390	669 (5.53%)	N/A	4 (0.03%)
27	OTHER ABNORMALITIES OF BREATHING	ICD-10 Diagnosis	R0589	N/A	8 (0.07%)	N/A
28	CHEST PAIN, UNSPECIFIED	ICD-10 Diagnosis	R079	696 (5.75%)	9 (0.07%)	N/A
29	RESPIRATORY ARREST	ICD-10 Diagnosis	R092	N/A	10 (0.08%)	N/A
30	UNSPECIFIED ABDOMINAL PAIN	ICD-10 Diagnosis	R109	N/A	N/A	3 (0.02%)
31	WEAKNESS	ICD-10 Diagnosis	R531	787 (6.51%)	9 (0.07%)	3 (0.02%)
32	OTHER MALAISE	ICD-10 Diagnosis	R5381	808 (6.68%)	10 (0.08%)	3 (0.02%)
33	OTHER FATIGUE	ICD-10 Diagnosis	R5383	787 (6.51%)	9 (0.07%)	3 (0.02%)
34	SYNCOPE AND COLLAPSE	ICD-10 Diagnosis	R55	N/A	7 (0.06%)	N/A
35	LOCALIZED EDEMA	ICD-10 Diagnosis	R600	N/A	7 (0.06%)	N/A
36	GENERALIZED EDEMA	ICD-10 Diagnosis	R601	N/A	7 (0.06%)	N/A
37	ENCOUNTER FOR FOLLOW-UP EXAMINATION AFTER COMPLETED TREATMENT FOR CONDITIONS OTHER THAN MALIGNANT NEOPLASM	ICD-10 Diagnosis	Z09	N/A	N/A	3 (0.02%)
38	ENCOUNTER FOR IMMUNIZATION	ICD-10 Diagnosis	Z23	709 (5.86%)	7 (0.06%)	N/A
39	ENCOUNTER FOR OTHER SPECIFIED AFTERCARE	ICD-10 Diagnosis	Z5189	N/A	9 (0.07%)	N/A
40	LONG TERM (CURRENT) USE OF ANTICOAGULANTS	ICD-10 Diagnosis	Z7901	680 (5.62%)	N/A	N/A
41	LONG TERM (CURRENT) USE OF OPIATE ANALGESIC	ICD-10 Diagnosis	Z79891	903 (7.46%)	10 (0.08%)	3 (0.02%)
42	OTHER LONG TERM (CURRENT) DRUG THERAPY	ICD-10 Diagnosis	Z79899	903 (7.46%)	10 (0.08%)	N/A

Figure 13 shows the most frequent procedure codes of the top 10% of high risk patients. The high risk cluster 1 has several procedures related to outpatient office visits, blood collection, comprehensive metabolic panels, blood count, and so on.

Figure 13: Most Frequent Procedure Codes of High Risk Patients

Obs	Procedure	Code Type	Code	Cluster_1	Cluster_3	Cluster_4
1	COLLECTION VENOUS BLOOD VENIPUNCTURE	CPT Procedure	36415	1153 (9.53%)	10 (0.08%)	4 (0.03%)
2	RADIOLOGIC EXAMINATION CHEST SINGLE VIEW FRONTAL	CPT Procedure	71010	832 (6.88%)	10 (0.08%)	4 (0.03%)
3	RADIOLOGIC EXAM CHEST 2 VIEWS FRONTAL&LATERAL	CPT Procedure	71020	851 (7.03%)	10 (0.08%)	3 (0.02%)
4	BASIC METABOLIC PANEL CALCIUM TOTAL	CPT Procedure	80048	875 (7.23%)	8 (0.07%)	N/A
5	COMPREHENSIVE METABOLIC PANEL	CPT Procedure	80053	1064 (8.80%)	10 (0.08%)	4 (0.03%)
6	LIPID PANEL	CPT Procedure	80061	973 (8.04%)	8 (0.07%)	4 (0.03%)
7	URNLS DIP STICK/TABLET REAGENT AUTO MICROSCOPY	CPT Procedure	81001	682 (5.64%)	7 (0.06%)	N/A
8	HEMOGLOBIN GLYCOSYLATED A1C	CPT Procedure	83036	824 (6.81%)	8 (0.07%)	4 (0.03%)
9	ASSAY OF MAGNESIUM	CPT Procedure	83735	N/A	N/A	3 (0.02%)
10	ASSAY OF THYROID STIMULATING HORMONE TSH	CPT Procedure	84443	847 (7.00%)	7 (0.06%)	4 (0.03%)
11	BLOOD COUNT COMPLETE AUTO&AUTO DIFRNTL WBC	CPT Procedure	85025	1097 (9.07%)	7 (0.06%)	4 (0.03%)
12	BLOOD COUNT COMPLETE AUTOMATED	CPT Procedure	85027	N/A	N/A	3 (0.02%)
13	PROTHROMBIN TIME	CPT Procedure	85610	978 (8.08%)	7 (0.06%)	4 (0.03%)
14	LEVEL IV SURG PATHOLOGY GROSS&MICROSCOPIC EXAM	CPT Procedure	88305	637 (5.27%)	8 (0.07%)	N/A
15	IIV3 VACCINE 3 YRS & OLDER FOR IM USE	CPT Procedure	90658	630 (5.21%)	N/A	N/A
16	OPHTH MEDICAL XM&EVAL COMPRHNSV ESTAB PT 1/>	CPT Procedure	92014	616 (5.09%)	N/A	N/A
17	ECG ROUTINE ECG W/LEAST 12 LDS I&R ONLY	CPT Procedure	93010	792 (6.55%)	9 (0.07%)	3 (0.02%)
18	THERAPEUTIC PX 1/> AREAS EACH 15 MIN EXERCISES	CPT Procedure	97110	853 (7.05%)	8 (0.07%)	3 (0.02%)
19	MANUAL THERAPY TQS 1/> REGIONS EACH 15 MINUTES	CPT Procedure	97140	592 (4.89%)	N/A	4 (0.03%)
20	THERAPEUT ACTVITY DIRECT PT CONTACT EACH 15 MIN	CPT Procedure	97530	N/A	N/A	3 (0.02%)
21	CHIROPRACTIC MANIPULATIVE TX SPINAL 3-4 REGIONS	CPT Procedure	98941	N/A	N/A	3 (0.02%)
22	OFFICE OUTPATIENT VISIT 5 MINUTES	CPT Procedure	99211	678 (5.60%)	N/A	4 (0.03%)
23	OFFICE OUTPATIENT VISIT 10 MINUTES	CPT Procedure	99212	931 (7.70%)	8 (0.07%)	3 (0.02%)
24	OFFICE OUTPATIENT VISIT 15 MINUTES	CPT Procedure	99213	1164 (9.62%)	10 (0.08%)	4 (0.03%)
25	OFFICE OUTPATIENT VISIT 25 MINUTES	CPT Procedure	99214	1137 (9.40%)	10 (0.08%)	4 (0.03%)
26	OFFICE OUTPATIENT VISIT 40 MINUTES	CPT Procedure	99215	N/A	8 (0.07%)	3 (0.02%)
27	INITIAL HOSPITAL CARE/DAY 70 MINUTES	CPT Procedure	99223	N/A	N/A	3 (0.02%)
28	SBSQ HOSPITAL CARE/DAY 25 MINUTES	CPT Procedure	99232	903 (7.46%)	9 (0.07%)	3 (0.02%)
29	SBSQ HOSPITAL CARE/DAY 35 MINUTES	CPT Procedure	99233	659 (5.45%)	7 (0.06%)	N/A
30	HOSPITAL DISCHARGE DAY MANAGEMENT 30 MIN/<	CPT Procedure	99238	N/A	8 (0.07%)	N/A
31	INITIAL INPATIENT CONSULT NEW/ESTAB PT 55 MIN	CPT Procedure	99253	N/A	7 (0.06%)	N/A
32	EMERGENCY DEPARTMENT VISIT HIGH/URGENT SEVERITY	CPT Procedure	99284	N/A	N/A	3 (0.02%)
33	EMERGENCY DEPT VISIT HIGH SEVERITY&THREAT FUNCJ	CPT Procedure	99285	642 (5.31%)	N/A	N/A
34	SBSQ NURSING FACIL CARE/DAY MINOR COMPLJ 15 MIN	CPT Procedure	99308	N/A	9 (0.07%)	N/A
35	GROUND MILEAGE, PER STATUTE MILE	HCPCS Procedure	A0425	N/A	8 (0.07%)	N/A
36	AMBULANCE SERVICE, ADVANCED LIFE SUPPORT, EMERGENCY TRANSPORT, LEVEL 1 (ALS 1 - EMERGENCY)	HCPCS Procedure	A0427	N/A	7 (0.06%)	N/A
37	ADMINISTRATION OF INFLUENZA VIRUS VACCINE	HCPCS Procedure	G0008	618 (5.11%)	N/A	3 (0.02%)

These reports are expected to give analysts an in-depth look at the key factors contributing to high health care costs. This model provides the capability to dig deeper using the Service Setting option so that you can analyze inpatient costs, outpatient costs, and so on, separately. It also enables you to analyze costs using different risk factors (Standard, CMS-HCC, and so on).

CONCLUSION

SAS Real World Evidence uses the power of SAS technology to provide a platform for the management, analysis, and visualization of real-world population-level health data. This paper focused on deriving insights using analytics. Using a cohort of patients who have been diagnosed with diabetes and have been administered insulin, we identified clusters of high cost patients to analyze their characteristics. After predicting expected costs for patients, we analyzed patients with very high variations from their expected cost. This was done using the Cost Analysis Model add-in and the outputs generated by the model.

ACKNOWLEDGMENTS

The author is grateful to David Olaleye and Lina Clover of the Health & Life Sciences R&D Division at SAS Institute Inc. for their contributions to the preparation of this paper.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Jay Paulson
SAS Institute Inc.
SAS Campus Drive
Cary, NC 27513
jay.paulson@sas.com