CrossMark

Subgroups of High-Cost Medicare Advantage Patients: an Observational Study

Brian W. Powers, MD, MBA[1,2,3,4], Jiali Yan, MS[5], Jingsan Zhu, MS, MBA[6], Kristin A. Linn, PhD[7], Sachin H. Jain, MD, MBA[3], Jennifer L. Kowalski, MS[8], and Amol S. Navathe, MD, PhD[6,9]

[1]Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA; [2]Department of Population Medicine, Harvard Medical School/Harvard Pilgrim Health Care Institute, Boston, MA, USA; [3]CareMore Health System, Cerritos, CA, USA; [4]Atrius Health, Boston, MA, USA; [5]Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA; [6]Department of Medical Ethics and Health Policy, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA; [7]Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA; [8]Anthem Public Policy Institute, Washington, DC, USA; [9]Corporal Michael J. Cresencz VA Medical Center, Philadelphia, PA, USA.

BACKGROUND: There is a growing focus on improving the quality and value of health care delivery for high-cost patients. Compared to fee-for-service Medicare, less is known about the clinical composition of high-cost Medicare Advantage populations.

OBJECTIVE: To describe a high-cost Medicare Advantage population and identify clinically and operationally significant subgroups of patients.

DESIGN: We used a density-based clustering algorithm to group high-cost patients (top 10% of spending) according to 161 distinct demographic, clinical, and claims-based variables. We then examined rates of utilization, spending, and mortality among subgroups.

PARTICIPANTS: Sixty-one thousand five hundred forty-six Medicare Advantage beneficiaries.

MAIN MEASURES: Spending, utilization, and mortality.

KEY RESULTS: High-cost patients ($n = 6154$) accounted for 55% of total spending. High-cost patients were more likely to be younger, male, and have higher rates of comorbid illnesses. We identified ten subgroups of high-cost patients: acute exacerbations of chronic disease (mixed); end-stage renal disease (ESRD); recurrent gastrointestinal bleed (GIB); orthopedic trauma (trauma); vascular disease (vascular); surgical infections and other complications (complications); cirrhosis with hepatitis C (liver); ESRD with increased medical and behavioral comorbidity (ESRD+); cancer with high-cost imaging and radiation therapy (oncology); and neurologic disorders (neurologic). The average number of inpatient days ranged from 3.25 (oncology) to 26.09 (trauma). Preventable spending (as a percentage of total spending) ranged from 0.8% (oncology) to 9.5% (complications) and the percentage of spending attributable to prescription medications ranged from 7.9% (trauma and oncology) to 77.0% (liver). The percentage of patients who were persistently high-cost ranged from 11.8% (trauma) to 100.0% (ESRD+). One-year mortality ranged from 0.0% (liver) to 25.8% (ESRD+).

CONCLUSIONS: We identified clinically distinct subgroups of patients within a heterogeneous high-cost Medicare Advantage population using cluster analysis. These subgroups, defined by condition-specific profiles and illness trajectories, had markedly different patterns of utilization, spending, and mortality, holding important implications for clinical strategy.

KEY WORDS: high-cost patients; care management; medicare advantage.

INTRODUCTION

The concentration of health care spending—5% of patients incur roughly half of total spending[1–3]—has prompted an intense focus on improving the quality and value of health care delivery for high-cost patients.[3–5] Current efforts to describe and segment high-cost patients have focused predominately on fee-for-service (FFS) Medicare populations.[3,6,7] Despite the fact that one in three Medicare beneficiaries is now enrolled in a Medicare Advantage (MA) plan,[8] less is known about high-cost MA patients.

Better understanding the clinical composition of high-cost patient populations has the potential to improve care management program design. Researchers and policymakers have suggested that suboptimal patient targeting—singular interventions are frequently applied to diverse patient populations—may be driving the limited effectiveness of current care management approaches.[3,7,9,10]

Descriptive analyses have shown that within high-cost patient populations, there is substantial variation in demographics, functional status, diagnoses, and disease burden.[6,7,11–13] Improving care management program effectiveness, therefore, may require identifying distinct subgroups of high-cost patients, and tailoring interventions to meet the unique needs of each group. Existing frameworks for identifying subgroups of high-cost patients are derived from expert opinion.[3,6,7,14] There may be an opportunity to supplement these approaches

by identifying subgroups exclusively based on the analysis of variation within patient data.

In this study, we aimed to (1) describe the demographic and clinical characteristics of a high-cost MA population, (2) use cluster analysis to derive high-cost patient subgroups from large volumes of clinical and claims data, and (3) explore whether these subgroups were meaningfully associated with patterns of utilization, spending, and mortality.

METHODS

Study Population

We obtained data for patients enrolled in MA plans offered by CareMore Health System in 2014 (n = 93,047). CareMore, a subsidiary of Anthem, Inc., offered MA plans in California, Arizona, and Virginia in 2014. We excluded patients who were not continuously enrolled for the entire year (n = 27,163), those who died during 2014 (n = 3649), those who did not have any approved medical claims during 2014 (n = 1867), and those enrolled in an institutional special needs plan (n = 1981). The final study population consisted of 61,546 patients.

Data

We extracted demographic, clinical, utilization, spending, and mortality data for the study population from CareMore's electronic data warehouse (EDW). The EDW aggregates data from electronic medical records (EMR) and administrative sources. Data were obtained for the years 2013–2015.

A full description of study variables is provided in Appendix 1. Briefly, we grouped variables into the following categories: Demographics, chronic conditions, active diagnoses, procedures, laboratory, pharmacy, mortality, utilization, and spending. Demographic variables included age and gender. Chronic conditions were assessed individually according to the Elixhauser[15] classification using prior year (2013) ICD-9 diagnosis codes. Active diagnoses were determined using 2014 ICD-9 codes, and grouped according to Agency for Healthcare Research and Quality (AHRQ) Clinical Conditions Software (CCS) categories.[16] Procedures were determined using 2014 CPT codes, and grouped according to AHRQ CSS categories.[17] Laboratory variables were assessed directly from the EMR. Pharmacy variables were calculated using pharmacy fill data, and included the number of unique medications as well as adherence, which was defined as the proportion of days covered (PDC)[18] among a limited set of common outpatient medications. Subsequent year (2015) mortality was assessed directly from the EDW. Utilization and spending variables were assessed directly from paid claims. Preventable spending was calculated using the AHRQ Prevention Quality Indicators algorithm[19] for inpatient spending, and the algorithm created by Billings et al.[20] for emergency department spending, both of which have been validated and used in prior work segmenting high-cost Medicare populations.[21,22]

Descriptive Analyses

We defined "high-cost" patients as those in the top decile of spending in 2014 (n = 6154). First, we described demographic, chronic condition, pharmacy, utilization, and spending variables across the entire study population. We then compared these variables between high-cost and non-high-cost subgroups using t tests for continuous variables and χ^2 tests for categorical variables.

Identifying Subgroups

Clustering is an unsupervised machine learning technique that groups observations (e.g., patients) according to similarities among measured characteristics. Clustering algorithms iteratively group observations into subgroups until finding the allocation that maximizes both intra-group similarity and inter-group differences (see an accompanying article[23] for more information).

The dataset used for cluster analysis included demographics, chronic conditions, active diagnoses, procedures, laboratory, and pharmacy variables for the subset of high-cost patients (n = 6154). Utilization and spending variables were not used for clustering, allowing for comparison of utilization and spending across clusters (see below).

To perform cluster analysis, we began by analytically reducing the number of variables in the dataset—a task known in computer science as dimension reduction. We removed variables with extremely low variance, and those that were highly correlated. A total of 161 variables remained (full list provided in Appendix 2). We then utilized a non-linear dimension reduction algorithm[24,25] to create a low-dimension representation of the dataset. Additional detail regarding dimension reduction is provided in Appendix 3 and an accompanying article.[23] Finally, we applied a density-based clustering algorithm—Ordering Points To Identify the Clustering Structure (OPTICS)[26,27]—to the low-dimension dataset. We restricted the minimum number of patients per subgroup to be at least 62 (or 1% of the high-cost population) in order to ensure that the subgroups were operationally meaningful. Our rationale for choosing the OPTICS algorithm and information on tuning parameters is described in an accompanying article.[23]

Subgroup Analysis

To describe the clinical composition of the resultant subgroups, we first calculated high-cost population means and subgroup-specific means for each variable used in clustering. We then calculated standardized ratios of subgroup means to population means, such that larger numbers represented variables for which the subgroup deviated most from the broader high-cost population. We assigned a clinical descriptive label to each subgroup based on the variables with the highest standardized ratios as well as variables for which the ratios varied most among subgroups. Given the numerous variables, we chose to present the ten variables with the largest standardized ratios (labeled as "distinguishing factors") for each subgroup. Appendix 4 contains a complete list of standardized ratios for all subgroups.

Next, we compared spending, utilization, and mortality across subgroups. We calculated 2014 utilization rates, average spending, composition of spending, and rates of preventable spending among each subgroup. To better understand the trajectory of spending for subgroups, we calculated average spending, preventable spending, and the prevalence of persistent high-cost status in 2015. We defined persistent high-cost status as remaining in the top decile of total spending in 2015. Finally, we calculated 2015 mortality rates among each subgroup. We excluded patients not continuously enrolled in a CareMore MA plan from the 2015 analyses ($n = 1430$).

Data preparations were done in SAS version 9.4 (SAS Institute, Cary, NC). R version 3.2.5 was used for all other analyses.

RESULTS

High-Cost Patient Characteristics

In a national Medicare Advantage population ($n = 61,546$), the highest cost 10% of patients ($n = 6154$) accounted for 55% of total population spending in 2014, with average annual spending of $55,696 per patient. Among high-cost patients in 2014, 64% were persistently high-cost (remained in top 10% of spending in 2015).

Table 1 describes utilization patterns for high-cost and non-high-cost patients. Compared to non-high-cost patients, high-cost patients had higher average rates of inpatient (IP) admissions (1.7 vs. 0.1), IP days (12.9 vs. 0.9), and emergency department (ED) visits (2.4 vs. 0.4). The total annual spending was roughly ten times higher among high-cost patients ($55,696 vs. $5071), and the rate of preventable spending was also notably higher (7.1% vs. 3.6%).

High-cost patients also varied substantially from non-high-cost patients across demographic and clinical characteristics (Table 1). High-cost patients were younger (average age 70.9 vs. 73.5), and more likely to be male (49.6% vs. 42.9%). High-cost patient had higher rates of co-occurring chronic conditions (9.0 vs. 4.4) as well as significantly higher rates of all individually assessed chronic conditions, including congestive heart failure (44% vs. 11%), vascular disease (60% vs. 28%), and renal failure (69% vs. 39%). Among high-cost patients, rates of polypharmacy were higher (14.6 vs. 7.9 average prescriptions), and medication adherence was lower (0.7 vs. 0.8 PDC).

High-Cost Patient Subgroups

Cluster analysis identified ten subgroups of high-cost patients. The number of patients in each subgroup ranged from 56 to 3686, and 382 patients were not assigned to any subgroup. Differentiation across subgroups was driven predominately by comorbidities and procedures; laboratory values and demographics were less important (Table 2). There were divergent patterns of index year utilization and spending (Table 3) as well as spending and mortality trajectories (Table 4) across

subgroups. Each subgroup and associated patterns of utilization, spending, and mortality are described below. Comparative statements are relative to other high-cost patients, not the entire study population.

Acute Exacerbations of Chronic Disease (Mixed). Distinguished by procedures associated with acute hospitalizations. These patients also had higher-than-average rates of cerebrovascular disease, chronic obstructive pulmonary disease, ischemic heart disease, congestive heart failure, and behavioral health disorders. Although rates of utilization and spending were close to average, this subgroup was characterized by higher-than-average rates of preventable spending in 2014 (7.9%) and 2015 (4.3%).

End-Stage Renal Disease. Distinguished by active diagnoses, chronic conditions, and procedures related to end-stage renal disease (ESRD) and dialysis. These patients were among the highest cost (average 2014 spending $74,385), were among the most likely to be persistently high-cost (97.3%), and had a high rate of mortality in 2015 (18.5%). Inpatient and ED utilization were substantially below average.

Recurrent Gastrointestinal Bleed. Distinguished by active diagnoses and procedures related to recurrent gastrointestinal bleed (GIB), as well as chronic conditions representing the sequelae of recurrent GIB. Rates of ED and inpatient utilization were among the highest (2.74 inpatient admissions, 3.25 ED admissions, and 18.96 inpatient days) in 2014, but these patients were among the least likely to remain persistently high-cost (34.6%).

Orthopedic Trauma (Trauma). Distinguished by active diagnoses and procedures related to fractures and other traumatic events. These patients had the highest rates of inpatient utilization (26.09 inpatient days), but were the least likely to be persistently high-cost (11.8%).

Vascular Disease (Vascular). Distinguished by active diagnoses and procedures related to peripheral vascular disease. Rates of utilization and spending in 2014 were among the lowest; rates of persistently high-cost status were close to average.

Surgical Infections and Other Complications (Complications). Distinguished by active diagnoses and procedures related to surgical wounds, infections, and other iatrogenic complications. Patients in this subgroup had higher-than-average rates of inpatient utilization and total spending (19.69 inpatient days and total spending of $60,103), but were less likely to be persistently high-cost (31.1%). The rate of mortality in 2015 was among the highest (19.7%).

Cirrhosis with Hepatitis C (Liver). Distinguished by active diagnoses and chronic conditions related to the diagnosis,

TABLE 1 Characteristics of the Study Population, by High-Cost Status

	Overall population	High-cost subgroup	Non-high-cost subgroup	*p* value
Number of patients (no.)	61,546	6155	55,391	
Demographics				
Age, mean (SD)	73.2 (9.5)	70.9 (11.4)	73.5 (9.2)	< 0.001
Male (%)	43.5	49.6	42.9	< 0.001
Chronic conditions				
Total no., mean (SD)	4.9 (3.2)	9 (3.4)	4.4 (2.8)	< 0.001
Hypertension (%)	80.4	94.4	78.9	< 0.001
Renal failure (%)	42.3	69.0	39.4	< 0.001
Diabetes uncomplicated (%)	41.5	63.1	39.1	< 0.001
Diabetes complicated (%)	38.4	59.3	36.0	< 0.001
Obesity (%)	31.9	44.6	30.5	< 0.001
Peripheral vascular disorders (%)	31.2	60.4	28.0	< 0.001
Depression (%)	30.2	46.1	28.4	< 0.001
Chronic pulmonary disease (%)	28.8	48.4	26.6	< 0.001
Hypothyroidism (%)	21.3	27.7	20.6	< 0.001
Drug abuse (%)	19.7	37.0	17.7	< 0.001
Fluid and electrolyte disorders (%)	16.9	54.8	12.7	< 0.001
Congestive heart failure (%)	14.3	43.9	11.0	< 0.001
Weight loss (%)	12.4	37.5	9.6	< 0.001
Valvular disease (%)	11.5	28.2	9.6	< 0.001
Solid tumor without metastasis (%)	9.1	19.0	8.0	< 0.001
Rheumatoid arthritis/collagen (%)	9.0	15.4	8.3	< 0.001
Other neurological disorders (%)	8.6	23.0	7.1	< 0.001
Deficiency anemia (%)	8.0	33.1	5.2	< 0.001
Liver disease (%)	7.6	17.1	6.5	< 0.001
Coagulopathy (%)	5.5	19.7	3.9	< 0.001
Alcohol abuse (%)	5.3	9.7	4.8	< 0.001
Psychoses (%)	3.4	9.2	2.8	< 0.001
Pulmonary circulation disorders (%)	3.4	11.3	2.5	< 0.001
Peptic ulcer disease (%)	2.1	5.6	1.8	< 0.001
Blood loss anemia (%)	1.6	6.7	1.1	< 0.001
Paralysis (%)	1.6	5.3	1.2	< 0.001
Metastatic cancer (%)	1.4	6.4	0.9	< 0.001
Lymphoma (%)	0.8	2.6	0.6	<0.001
AIDS/HIV (%)	0.1	0.7	0.0	< 0.001
Prescription medications				
No. of prescriptions, mean (SD)	8.5 (5.5)	14.6 (6.3)	7.9 (4.9)	< 0.001
Proportion of days covered, mean (SD)	0.8 (0.2)	0.7 (0.2)	0.8 (0.2)	< 0.001
Utilization				
Inpatient admissions, mean (SD)	0.3 (0.8)	1.7 (1.8)	0.1 (0.4)	< 0.001
Inpatient days, mean (SD)	2.1 (100.4)	12.9 (20.8)	0.9 (105.6)	< 0.001
Emergency department visits, mean (SD)	0.6 (1.6)	2.4 (3.7)	0.4 (0.9)	< 0.001
Total spending ($), mean (SD)	10,133 (19,842)	55,696 (37,217)	5071 (5229)	< 0.001
Preventable spending ($), mean (SD)	558 (4114)	3930 (12,168)	184 (980)	< 0.001

SD, standard deviation; AIDS/HIV, acquired immunodeficiency syndrome/human immunodeficiency virus

management, and sequelae of hepatitis C infection. These patients had the highest rates of average spending in 2014 ($78,706), which was driven predominately by prescription drug costs (77.0% of total spending). Patients in this subgroup were more likely than average to remain high-cost (58.6%) and had the lowest rate of mortality in 2015 (0.0%).

ESRD with Increased Medical and Behavioral Comorbidity (ESRD+). Distinguished by diagnoses, chronic conditions, and procedures related to ESRD and dialysis. Compared to the ESRD subgroup, there were higher rates of congestive heart failure, behavioral health disorders, liver failure, and cerebrovascular disease. Patients in this subgroup also had higher rates of preventable spending (4.8% vs. 3.0%) than those in the ESRD subgroup. These patients were the most likely to remain high-cost in 2015 (100.0%) and had the highest rate of mortality in 2015 (25.8%).

Cancer with High-Cost Imaging and Radiation Therapy (Oncology). Distinguished by metastatic and non-metastatic

cancer diagnoses, imaging procedures related to disease staging and surveillance, and brachytherapy. Patients were predominately male (88%) and the most common oncologic diagnosis was prostate cancer. Rates of 2014 utilization (3.25 inpatient days), spending ($46,240), persistent high-cost status (20.0%), and mortality in 2015 (5.5%) were among the lowest.

Neurologic Disorders (Neurologic). Distinguished by active diagnoses and chronic conditions encompassing neurologic disorders (most notably multiple sclerosis) and neurologic diagnostic procedures. These patients had among the lowest rates of 2014 inpatient utilization, but were among the most likely to remain high-cost (75.6%), with spending driven by prescription drugs (51.7% of total spending).

DISCUSSION

We found that health care spending was highly concentrated in a national Medicare Advantage population. The highest cost

TABLE 2 Description of High-Cost Patient Subgroups

Description	No. of patients (%)	Distinguishing factors	
		Variable*	**Category**
Acute exacerbations of chronic disease (mixed)	3686 (59.9%)	Arterial blood gases	Procedures
		Bacterial infection	Active diagnoses
		Metastatic cancer	Chronic conditions
		Other vascular catheterization; not heart	Procedures
		CT scan chest	Procedures
		Routine chest X-ray	Procedures
		Electroencephalogram	Procedures
		Other respiratory therapy	Procedures
		Cerebrovascular disease	Active diagnoses
		CT scan abdomen	Procedures
End-stage renal disease (ESRD)	1222 (19.9%)	Peritoneal dialysis	Procedures
		Diseases of the urinary system	Active diagnoses
		Complications	Active diagnoses
		Iron deficiency anemia	Chronic conditions
		Other vascular catheterization; not heart	Procedures
		Creatinine	Laboratory
		Bacterial infection	Active diagnoses
		Fluid and electrolyte disorders	Chronic conditions
		Blood loss anemia	Chronic conditions
		Congestive heart failure	Chronic conditions
Recurrent GI bleed (GIB)	221 (3.6%)	Blood loss anemia	Chronic conditions
		Gastrointestinal hemorrhage	Active diagnoses
		Peptic ulcer disease excluding bleeding	Chronic conditions
		Upper gastrointestinal endoscopy; biopsy	Procedures
		Anemia	Active diagnoses
		Other vascular catheterization; not heart	Procedures
		Iron deficiency anemia	Chronic conditions
		Coagulopathy	Chronic conditions
		Routine chest X-ray	Procedures
		CT scan abdomen	Procedures
Orthopedic trauma (trauma)	201 (3.3%)	Fractures	Active diagnoses
		Headache	Active diagnoses
		Traction; splints; and other wound care	Procedures
		Other diagnostic radiology	Procedures
		Other CT scan	Procedures
		Other injuries/conditions due to external causes	Active diagnoses
		Other non-OR nervous system procedures	Procedures
		Diagnostic physical therapy	Procedures
		Electroencephalogram	Procedures
		CT scan head	Procedures
Vascular disease with interventions (vascular)	86 (1.4%)	Diseases of arteries; arterioles; and capillaries	Active diagnoses
		Other vascular catheterization; not heart	Procedures
		Other diagnostic cardiovascular procedures	Procedures
		Complications	Active diagnoses
		Diseases of veins and lymphatics	Active diagnoses
		History of mental health/substance abuse	Active diagnoses
		Diagnostic US	Procedures
		Open wounds	Active diagnoses
		Anesthesia	Procedures
		Cerebrovascular disease	Active diagnoses
Surgical infections and other complications (complications)	81 (1.3%)	Complications	Active diagnoses
		Traction; splints; and other wound care	Procedures
		Open wounds	Active diagnoses
		Ancillary services	Procedures
		Other vascular catheterization; not heart	Procedures
		Skin and subcutaneous tissue infections	Active diagnoses
		Blood loss anemia	Chronic conditions
		Anesthesia	Procedures
		Bacterial infection	Active diagnoses
		Electroencephalogram	Procedures
Cirrhosis with hepatitis C (liver)	81 (1.3%)	Viral infection	Active diagnoses
		Liver disease	Active diagnoses
		Liver disease	Chronic conditions
		Upper gastrointestinal endoscopy; biopsy	Procedures
		Other bone diseases and musculoskeletal deformities	Active diagnoses
		Radioisotope scan and function studies	Procedures
		Immunizations and screening for infectious disease	Active diagnoses
		Anxiety disorders	Active diagnoses
		Coagulopathy	Chronic conditions
		Other injuries and conditions due to external causes	Active diagnoses

Description	No. of patients (%)	Distinguishing factors	
		Variable*	Category
ESRD with increased medical and behavioral comorbidity (ESRD+)	73 (1.2%)	Peritoneal dialysis	Procedures
		Diseases of the urinary system	Active diagnoses
		Complications	Active diagnoses
		Iron deficiency anemia	Chronic conditions
		Other vascular catheterization; not heart	Procedures
		Blood loss anemia	Chronic conditions
		Creatinine	Laboratory
		Congestive heart failure	Chronic conditions
		Coagulopathy	Chronic conditions
		Fluid and electrolyte disorders	Chronic conditions
Cancer with high-cost Imaging and radiation therapy (oncology)	65 (1.1%)	Other CT scan	Procedures
		Metastatic cancer	Chronic conditions
		Solid tumor without metastasis	Chronic conditions
		Diagnostic procedures on nose; mouth; and pharynx	Procedures
		Radioisotope scan and function studies	Procedures
		Fluid and electrolyte disorders	Active diagnoses
		Diseases of male genital organs	Active diagnoses
		Magnetic resonance imaging	Procedures
		CT scan abdomen	Procedures
		Noninfectious gastroenteritis	Active diagnoses
Neurologic disorders(neurologic)	56 (0.9%)	Hereditary/degenerative nervous system conditions	Active diagnoses
		Other neurological disorders	Chronic conditions
		Electroencephalogram	Procedures
		CT scan head	Procedures
		Paralysis	Chronic conditions
		Magnetic resonance imaging	Procedures
		Noninfectious gastroenteritis	Active diagnoses
		Bacterial infection	Active diagnoses
		Diagnostic physical therapy	Procedures
		Psychoses	Chronic conditions

ESRD, end-stage renal disease, GIB, gastrointestinal bleed; CT, computed tomography; OR, operating room
*As described in the "Methods" section, these variables represent those with the highest standardized ratios (subgroup means to high-cost population means), represented in descending order (see Appendix 4 for a complete list of all standardized ratios for all variables among all subgroups)

10% of patients accounted for 55% of total spending, a level of spending concentration similar to that of FFS Medicare.[2,28] The majority of high-cost patients were persistently high-cost—65% remained in the highest cost decile the following year. This is in contrast to FFS Medicare, where rates of persistently high-cost status range from 25 to 45%.[2,21,28]

High-cost patients in this study had roughly twice as many co-occurring chronic conditions as non-high-cost patients, which is consistent with recent research in a FFS Medicare population.[7] High-cost patients had higher rates of all comorbid conditions assessed, including diabetes, congestive heart failure, chronic obstructive pulmonary disease, hypertension, depression, and renal failure. Rates of comorbid conditions were higher than those previously described for high-cost FFS Medicare beneficiaries.[2,21,28,29]

To better understand the composition of this high-cost population, we used cluster analysis to identify subgroups of patients according to similarities across 161 demographic and clinical variables. We identified ten subgroups: acute exacerbations of chronic disease (mixed); end-stage renal disease (ESRD); recurrent gastrointestinal bleed (GIB); orthopedic trauma (trauma); vascular disease (vascular); surgical infections and other complications (complications); cirrhosis with hepatitis C (liver); ESRD with increased medical and behavioral comorbidity (ESRD+); cancer with high-cost imaging and radiation therapy (oncology); and neurologic

TABLE 3 Utilization and Spending for High-Cost Patient Subgroups, 2014

	Mixed	ESRD	GIB	Trauma	Vascular	Complications	Liver	ESRD+	Oncology	Neurologic
Inpatient admissions (mean)	1.87	1.01	2.74	2.42	1.49	2.46	0.52	1.58	0.4	1.39
Inpatient days (mean)	14.4	6.07	18.96	26.09	9.33	19.69	3.98	8.47	3.25	14.5
ED visits (mean)	2.59	1.94	3.25	2.65	1.28	2.1	0.91	2.3	0.75	1.77
Total spending, mean ($)	47,936	74,385	52,259	40,175	45,829	60,103	78,706	79,724	46,240	66,417
Preventable (%)	7.9%	3.0%	8.0%	5.7%	6.8%	9.5%	0.9%	4.8%	0.8%	3.8%
Inpatient (%)	42.0%	12.8%	57.2%	44.2%	42.7%	40.2%	9.4%	13.0%	4.4%	20.4%
Medication (%)	18.4%	11.1%	10.9%	7.9%	14.2%	11.3%	77.0%	8.1%	7.9%	51.7%
Other (%)	39.6%	76.1%	31.9%	47.9%	43.1%	48.5%	13.6%	78.8%	87.7%	27.9%

ESRD, end-stage renal disease; GIB, gastrointestinal bleed; ED, emergency department

TABLE 4 Utilization and Spending Trajectories for High-Cost Patient Subgroups, 2014–2015

	Total spending, mean ($)		Preventable spending (%)		Top decile of total spending (%)		Mortality* (%)	
	2014	2015	2014	2015	2014	2015	2014	2015
Mixed	$47,936	$27,513	7.9%	4.3%	100.0%	35.8%	–	13.5%
ESRD	$74,385	$80,245	3.0%	2.1%	100.0%	97.3%	–	18.5%
GIB	$52,259	$25,985	8.0%	4.1%	100.0%	34.6%	–	17.0%
Trauma	$40,175	$13,685	5.7%	4.1%	100.0%	11.8%	–	10.4%
Vascular	$45,829	$29,044	6.8%	2.1%	100.0%	35.7%	–	8.6%
Complications	$60,103	$20,361	9.5%	2.4%	100.0%	31.1%	–	19.7%
Liver	$78,706	$41,852	0.9%	1.2%	100.0%	58.6%	–	0.0%
ESRD+	$79,724	$84,996	4.8%	1.6%	100.0%	100.0%	–	25.8%
Oncology	$46,240	$17,052	0.8%	1.7%	100.0%	20.0%	–	5.5%
Neurologic	$66,417	$67,295	3.8%	1.9%	100.0%	75.6%	–	9.8%

Mortality not reported for 2014 since patients who died in 2014 were excluded from the analysis (see the "Methods" section). Patients that died in 2015 were included in calculations for other values reported in this table

disorders (neurologic). We found that these subgroups, while identified using only clinical and demographic data, had markedly different patterns of utilization, spending, and mortality.

Taken together, our findings hold important implications for the design and implementation of care management programs. First, these results add to a growing awareness of the heterogeneity of high-cost populations.[6,7,11,22,30,31] Traditional narratives describe high-cost patient populations as being comprised of individuals with multiple, poorly controlled chronic conditions, often with coincident frailty and behavioral health disorders.[4] Though we found this description to be true in aggregate (i.e., higher-than-average average rates of comorbidity among high-cost patients), it obscures substantial heterogeneity within the high-cost population. The mixed subgroup (roughly 60% of patients) more closely resembled the aforementioned narrative, but the remaining 40% of patients had disparate diagnoses and clinical compositions, often dominated by either a single condition or a single acute event.

The subgroups we identified share some similarities with high-cost patient subgroups identified in other populations. For example, research in FFS Medicare and single-health system populations have described subgroups of high-cost patients similar to the acute/mixed,[6,7,30–32] ESRD/ESRD+,[6,7] vascular,[6] GIB,[32] complications,[33] trauma,[30] and neurologic[30] subgroups in this study. Incomplete overlap suggests that certain high-cost patient subgroups can be generalized across populations, but also that there is significant variability among different populations.

Second, we found disparate patterns of utilization, spending, and mortality across subgroups, suggesting that uniform care management strategies and interventions are likely to be insufficient. Traditional care management approaches—in which nurse care managers or other allied health professionals assist patients with disease management and medication adherence to reduce the risk of destabilization and inpatient utilization[4,30]—hold promise for subgroups with multi-morbidity and persistently high spending (e.g., ESRD and ESRD+). Interestingly, patients in the acute/mixed subgroup were

relatively unlikely to remain high-cost (35.8%), casting doubt on the efficacy of traditional care management programs in reducing spending in this subgroup, and pointing to the importance of better identifying patients at risk of being persistently high-cost.

Among the other subgroups identified, traditional care management approaches are unlikely to be effective. For example, the neurologic and liver subgroups had among the highest rates of spending and persistently high spending. Within these subgroups, spending was driven predominately by prescription drug costs, indicating that the rational use and pricing of specialty pharmaceuticals may be the most effective strategies for reducing spending. For subgroups defined by acute events (trauma, GIB, complications), there may be limited opportunities to improve care and reduce spending. High rates of mortality among the ESRD, ESRD+, and GIB subgroups should prompt a focus on recognizing, and addressing, life-limiting illness with palliative care and other interventions.

This study has several limitations. First, our study population consisted of patients enrolled in MA plans offered by a single health insurer. As such, the subgroups we identified may not be generalizable to other populations, including other MA or FFS Medicare populations. However, less is known about MA populations, so our analysis begins to fill an important gap. Second, patterns of spending, utilization, and mortality among the clinical phenotypes we identified could be impacted by existing care management programs at CareMore[34] and, therefore, may not be generalizable to other populations. Third, we did not have access to patient-level data on important social determinants of health (e.g., income, education, social isolation), despite a growing appreciation for the impact of these factors on spending and outcomes, especially among high-cost patients. Finally, we used the OPTICS algorithm for cluster analysis. Different clustering algorithms are likely to produce different results. However, as discussed in an accompanying article,[23] we believe OPTICS is the optimal algorithm for clustering high-cost patient populations.

Compliance with Ethical Standards:

This study was approved by the Institutional Review Board of the University of Pennsylvania.

Conflict of Interest: *Dr. Navathe reports that he has received grant support from Hawaii Medical Service Association and Oscar Health; personal fees from Navvis and Co., Navigant Inc., Lynx Medical, Indegene Inc., and Sutherland Global Services; personal fees and equity from NavaHealth; serves on the board without compensation for Integrated Services, Inc., speaking fees from the Cleveland Clinic, and honoraria from Elsevier Press. Dr. Linn reports that she has received grant support from Hawaii Medical Service Association. Dr. Jain reports employment by Anthem, Inc.; stock ownership in Anthem, Inc., and honoraria from Elsevier Press. Ms. Kowalski reports employment by Anthem, Inc. and stock ownership in Anthem, Inc. and Amazon. Dr. Powers reports employment by Anthem, Inc. All other authors declare no conflicts of interest.*

Publisher's Note *Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.*

REFERENCES

1. National Institute of Health Care Management. The Concentration of U.S. Health Care Spending. 2017; https://www.nihcm.org/topics/cost-quality/concentration-of-us-health-care-spending. Accessed October 19, 2018.
2. **Riley GF**. Long-term trends in the concentration of Medicare spending. Health Aff (Millwood). 2007;26(3):808–816.
3. National Academy of Medicine. Effective care for high-need patients. Washington, DC: National Academy of Medicine; 2017.
4. **Hong CS**, **Abrams MK**, **Ferris TG**. Toward increased adoption of complex care management. N Engl J Med. 2014;371(6):491–493.
5. **McWilliams JM**. Cost Containment and the Tale of Care Coordination. N Engl J Med. 2016;375(23):2218–2220.
6. **Clough JD**, **Riley GF**, **Cohen M**, et al. Patterns of care for clinically distinct segments of high cost Medicare beneficiaries. Healthc (Amst). 2016;4(3):160–165.
7. **Joynt KE**, **Figueroa JF**, **Beaulieu N**, **Wild RC**, **Orav EJ**, **Jha AK**. Segmenting high-cost Medicare patients into potentially actionable cohorts. Healthc (Amst). 2017;5(1–2):62–67.
8. **Jacobson G**, **Damico A**, **Neuman T**, **Gold M**. Medicare Advantage 2017 Spotlight: Enrollment Market Update. 2017; http://files.kff.org/attachment/Issue-Brief-Medicare-Advantage-2017-Spotlight-Enrollment-Market-Update. Accessed October 19, 2018.
9. **Hong CS**, **Siegel AL**, **Ferris TG**. Caring for High-Need, High-Cost Patients: What Makes for a Successful Care Management Program? 2014; https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_issue_brief_2014_aug_1764_hong_caring_for_high_need_high_cost_patients_ccm_ib.pdf. Accessed October 19, 2018.
10. **Blumenthal D**, **Abrams MK**. Tailoring complex care management for high-need, high-cost patients. JAMA. 2016;316(16):1657–1658.
11. **Hayes SL**, **Salzberg CA**, **McCarthy D, et al.** High-Need, High-Cost Patients: Who Are They and How Do They Use Health Care—A Population-Based Comparison of Demographics, Health Care Use, and Expenditures. 2016; https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_issue_brief_2016_aug_1897_hayes_who_are_high_need_high_cost_patients_v2.pdf. Accessed October 19, 2018.
12. **Cohen S, Uberoi N.** Differentials in the concentration in the level of health expenditures across population subgroups in the U.S., 2010. *Statistical Brief #421* 2013; https://meps.ahrq.gov/data_files/publications/st421/stat421.shtml. Accessed October 19, 2018.
13. **Zodet M.** Characteristics of Persons with High Health Care Expenditures in the U.S. Civilian Noninstitutionalized Population, 2014. *Statistical Brief #496* 2016; https://meps.ahrq.gov/data_files/publications/st496/stat496.pdf. Accessed October 19, 2018.
14. **Lynn J**, **Straube BM**, **Bell KM**, **Jencks SF**, **Kambic RT**. Using population segmentation to provide better health care for all: the "Bridges to Health" model. Milbank Q. 2007;85(2):185–208; discussion 209-112.
15. **Quan H**, **Sundararajan V**, **Halfon P**, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005;43(11):1130–1139.
16. Agency for Healthcare Research and Quality. Clinical Classifications Software (CCS) for ICD-9-CM. 2017; https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp. Accessed October 19, 2018.
17. Agency for Healthcare Research and Quality. Clinical Classifications Software for Services and Procedures. 2017; https://www.hcup-us.ahrq.gov/toolssoftware/ccs_svcproc/ccssvcproc.jsp. Accessed October 19, 2018.
18. **Choudhry NK**, **Shrank WH**, **Levin RL**, et al. Measuring concurrent adherence to multiple related medications. Am J Manag Care. 2009;15(7):457–464.
19. Agency for Healthcare Research and Quality. Prevention Quality Indicators Overview. 2017; http://www.qualityindicators.ahrq.gov/modules/pqi_resources.aspx. Accessed October 19, 2018.
20. New York University. Faculty & Research. 2017; https://wagner.nyu.edu/faculty/billings/nyued-background. Accessed October 19, 2018.
21. **Joynt KE**, **Gawande AA**, **Orav EJ**, **Jha AK.** Contribution of preventable acute care spending to total spending for high-cost Medicare patients. JAMA. 2013;309(24):2572–2578.
22. **Figueroa JF**, **Joynt Maddox KE**, **Beaulieu N**, **Wild RC**, **Jha AK**. Concentration of Potentially Preventable Spending Among High-Cost Medicare Subpopulations: An Observational Study. Ann Intern Med. 2017;167(10):706–713.
23. **Yan J**, **Linn KA**, **Powers BW**, et al. Applying Machine Learning Algorithms to Segment High-Cost Patient Populations. J Gen Intern Med. 2018.
24. **Van Der Maaten L.** Accelerating t-SNE using tree-based algorithms. J Mach Learn Res. 2014;15(1):3221–3245.
25. **Maaten Lvd, Hinton G.** Visualizing data using t-SNE. J Mach Learn Res. 2008;9(Nov):2579–2605.
26. **Ester M**, **Kriegel H-P**, **Sander J**, **Xu X**. A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining; 1996; Portland, Oregon.
27. **Ankerst M**, **Breunig MM**, **Kriegel H-P**, #246, Sander r. OPTICS: ordering points to identify the clustering structure. SIGMOD Rec. 1999;28(2):49–60.
28. Congressional Budget Office. High-Cost Medicare Beneficiaries. Washington, DC: Congressional Budget Office;2005.
29. The Lewin Group. Individuals Living in the Community with Chronic Conditions and Functional Limitations: A Closer Look. 2010; https://aspe.hhs.gov/system/files/pdf/75961/closerlook.pdf. Accessed August 21, 2018.
30. **Powers BW**, **Chaguturu SK**, **Ferris TG**. Optimizing high-risk care management. JAMA. 2015;313(8):795–796.
31. **Figueroa JF, Jha AK.** Approach for achieving effective care for high-need patients. JAMA Intern Med. 2018.
32. **Newcomer SR**, **Steiner JF**, **Bayliss EA**. Identifying subgroups of complex patients with cluster analysis. Am J Manag Care. 2011;17(8):e324–332.
33. **Lee NS**, **Whitman N**, **Vakharia N**, **Ph DG**, **Rothberg MB**. High-cost patients: hot-spotters don't explain the half of it. J Gen Intern Med. 2017;32(1):28–34.
34. **Hostetter M, Klein S, McCarthy D.** CareMore: Improving Outcomes and Controlling Health Care Spending for High-Needs Patients. 2017; https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_case_study_2017_mar_1937_hostetter_caremore_case_study_v2.pdf. Accessed October 19, 2018.