The American Journal of Surgery 220 (2020) 714–720

Contents lists available at ScienceDirect

The American Journal of Surgery

journal homepage: www.americanjournalofsurgery.com

Assessment of utilization efficiency using machine learning techniques: A study of heterogeneity in preoperative healthcare utilization among super-utilizers

J. Madison Hyer [a], Anghela Z. Paredes [a], Susan White [b], Aslam Ejaz [a], Timothy M. Pawlik [a, *]

[a] Department of Surgery, Division of Surgical Oncology, The Ohio State University Wexner Medical Center and James Cancer Hospital and Solove Research Institute, Columbus, OH, USA
[b] Department of Financial Services, The Ohio State University Wexner Medical Center and James Cancer Hospital and Solove Research Institute, Columbus, OH, USA

ARTICLE INFO

Article history:
Received 14 August 2019
Received in revised form
22 January 2020
Accepted 22 January 2020

Keywords:
Morbidity
Complexity
Risk factors
Machine learning

ABSTRACT

Introduction: In the United States, 5% of patients represent up to 55% of all health care costs. This study sought to define healthcare utilization patterns among super-utilizers, as well as assess possible variation in patient outcomes.
Methods: Medicare super-utilizers undergoing either a total hip or knee arthroplasty were identified and entered into a cluster analysis using annual preoperative charges to identify distinct patterns of utilization.
Results: Among 19,522 super-utilizers who underwent THA or TKA, there was a marked heterogeneity in overall utilization with 5 distinct clusters of utilization patterns. Of note, comorbidity burden was similar among the 5 clusters. Patient outcomes also varied by Cluster type, ranging from 6.9% to 16.5% experiencing complications and 1.0%–3.2% experiencing 90-day mortality.
Conclusion: While previous studies have suggested that super-utilizers are a homogenous group of patients, the current study demonstrated a large degree of heterogeneity within super-utilizers. Variations in utilization patterns were associated with postoperative outcomes and subsequent health care costs.
© 2020 Elsevier Inc. All rights reserved.

Introduction

The United States spends more on healthcare per capita than any other country in the world. In fact, healthcare care expenditures comprised nearly 18% of the United States annual gross domestic product in 2017.[1,2] As such, there has been increased efforts to reduce wasteful healthcare utilization and spending, while concurrently improving patient outcomes and quality of life. Traditionally, most studies have characterized high healthcare utilization as a one-dimensional proxy for comorbidity burden, poor medical management, or an undesired outcome.[3–7] More recently, however, some investigators have questioned whether specific patterns of healthcare of under- or over-utilization may be associated with worse outcomes.[3–7] To this point, Gawande recently proposed the idea that *better* care – not less – could in fact reduce total healthcare expenditures, while improving patient quality of life.[7]

Within the surgical population, there is a subgroup of patients who consume many more healthcare resources compared with other individuals. These high-resource utilizers – or “super-utilizers” – represent the most extreme cases of high healthcare utilization. Recent estimates have suggested that the top 5% of healthcare utilizers may be responsible for more than 40–55% of all healthcare costs.[2,8–10] The majority of studies analyzing super-utilizers have largely studied this group as a homogenous population relative to other low-utilizer patient populations.[10–13] Significant heterogeneity may exist, however, within the super-utilizer population itself. Given that super-utilizers represent the group of patients most associated with extreme health care costs, further examination of variations in “outlier” healthcare utilization among super-utilizers may allow for more targeted efforts at cost-savings.

One methodology to identify variations in the super-utilizer population is through the use of machine learning. Machine

* Corresponding author. Department of Surgery, The Urban Meyer III and Shelley Meyer Chair for Cancer Research, The Ohio State University, Wexner Medical Center, 395 W. 12th Ave, Suite 670, USA.
E-mail address: tim.pawlik@osumc.edu (T.M. Pawlik).

https://doi.org/10.1016/j.amjsurg.2020.01.043
0002-9610/© 2020 Elsevier Inc. All rights reserved.

learning, a branch of artificial intelligence, is a statistical approach that offers a number of benefits compared with conventional statistical techniques.[10,14–21] One of the principal benefits to machine learning is the ability to minimize human input, error, and user bias, thereby maximizing objectivity and minimizing subjectivity.[22] The objective of the current study was to define distinct patterns of preoperative utilization within the super-utilizer orthopedic surgical population using machine learning techniques. Specifically, we sought to characterize variations in clinical outcomes among various subgroups of patients undergoing either a total hip arthroplasty (THA) or total knee arthroplasty (TKA) identified as super-utilizers.

Methods

Data source and sample selection

Data were derived from 100% Medicare Inpatient and Outpatient Standard Analytic Files (SAFs) between 2012 and 2016. The SAFs are maintained by the Centers for Medicare and Medicaid Services (CMS) and include encounter-level (i.e. diagnoses, procedures, expenditures, etc.), as well as patient-level (i.e. race, gender, age, etc.) data. Patients aged 65 or older at the time of surgery who underwent either a THA or TKA between 2013 and 2015 were included. International Classification of Disease, Ninth Revision, (ICD-9) codes THA (8151) and TKA (8154) were used to identify the study cohort. For patients who underwent both a THA and TKA between 2013 and 2015, only the first procedure was included in the analysis. To be included in the analytic cohort, patients had to 1) be enrolled in Medicare parts A & B at the time of surgery, 2) received no additional payments from a health maintenance organization (HMO), and 3) had at least one recorded inpatient or outpatient encounter in the year prior to surgery and the year following surgery. In order to examine a relatively homogenous group in regards to patient characteristics, only patients undergoing THA or TKA were included in the final cohort.

To estimate Medicare expenditures, payments were price standardized and adjusted by wage index, Disproportionate Share Hospital (DSH) and Indirect Medical Education (IME).[10,23] Expenditure data from years 2012–2016 were included to ensure expenditure data from an entire year preceding and following surgery were evaluated. Super-utilizers, as determined by annual preoperative expenditure, were identified using a bisecting k-means clustering method with bin sorting by median to computer cluster seed that has been previously described and utilized.[10,11,24] Super-utilizers were identified as the group with the highest median pre-operative expenditure. Length of stay was measured as the number of days between admission to discharge from the index inpatient encounter. The Charlson and Elixhauser comorbidity indices and Centers for Medicare and Medicaid Service's Hierarchical Condition Category (CMS-HCC) were calculated.[25–27]

Statistical analysis

Descriptive statistics were presented as median (interquartile range [IQR]) and frequency (%: relative frequency) for continuous and categorical variables, respectively. Prior to classifying patients into different preoperative utilization groups, the preoperative inpatient and outpatient charges 365 days prior to the date of surgery were compiled and categorized on the basis of charge type. These categorizations have been previously described and short descriptions are provided in Supplemental Table 1.[28] To cluster patients into distinct preoperative utilization groups based on patterns of preoperative healthcare utilization, a hierarchical cluster analysis using Ward's minimum variance method was utilized.[24] The cubic clustering criterion, pseudo *F* and the dendrogram were utilized to determine the proper number of clusters for the study.

Following the cluster analysis, a backward stepwise discriminant analysis was performed on all variables included in the cluster analysis in addition to adjusted total inpatient and total outpatient expenditure. From the discriminant analysis, a subset of variables noted to best classify patients into respective utilization clusters were entered into a classification tree analysis using cost-complexity pruning to avoid over fitting and set a maximum of 10 leaves to maximize clinical applicability.[29] For the classification tree analysis, data were split into a 70/30 training/validation cohorts. All cut points were rounded to the nearest thousand United States Dollar (USD) and an analysis to identify concordant classifications was performed. All analyses were conducted using SAS v9.4.

Results

Among a total of 682,114 patients who underwent a THA or TKA between 2013 and 2015, 19,522 (2.9%) were identified as super-utilizers and included in the final analytical cohort (**Table 1**). For reference, the median (IQR) preoperative utilization among super-utilizers included in this study was $38,200 ($31,700-$50,500), whereas $1000 ($300-$2900), compared to all other Medicare beneficiaries. Median age was 74 (IQR: 69–79) years and the majority of patients were female (n = 11,145; 57.1%). The distribution of surgical procedures was approximately even as 51.6% (n = 10,069) of patients underwent a TKA and 48.4% (n = 9453) underwent a THA. The median (IQR) number of preoperative annual inpatient and outpatient encounters among all super-utilizers were 2[1–3] and 12,[6–20] respectively; the median per patient preoperative expenditure was $38,166 (IQR: $31,708-$50,459). Following surgery, median LOS was 3 days (IQR: 3–4), 1079 (16.%) experienced a complication during the index hospitalization, and the incidence of 30-day hospital readmission was 12.8% (n = 835) of patients. Of note, median (IQR) postoperative annual inpatient and outpatient encounters decreased to 1 (0–2) and 9[4–17] following surgery, respectively. In addition, median annual per patient postoperative expenditure decreased markedly to $12,697 (IQR: $2,124-$28,888) (Table 2).

Cluster analysis: defining differences among super-utilizers

In order to define potential differences among super-utilizers, machine learning was utilized to best classify patients into respective utilization clusters. Of note, cluster analysis identified five distinct clusters of patients among all super-utilizers. Compared with other clusters, Cluster 1 (n = 6,527, 33.4%) was comprised of patients who were more likely to be older (median age: 75; IQR: 70–81) and non-white (n = 533, 8.2%). Annual median preoperative expenditure was the lowest ($34,660, IQR: $30,511-$42,296) in Cluster 1 patients compared with all other super-utilizer cluster groups. Of note, Cluster 1 patients were characterized by comparatively low preoperative utilization across most utilization categories (Table 3). Cluster 2 was the largest cohort with 8589 (44.0%) patients (median age: 74, IQR 69–79; 44.5% male). Median per-patient preoperative health expenditure among Cluster 2 patients was slightly higher than Cluster 1 patients ($39,030, IQR: $32,010-$51,580). Of note, preoperative expenditures for patients in Cluster 2 was largely driven by charges associated with medical/surgical supplies ($31,520 IQR $18,070-$47,620) and surgical care that occurred prior to the THA/TKA admission ($29,310 IQR $16,110-$45,380).

Super-utilizer Cluster 3 was comprised of 1158 (5.9%) patients and had the highest proportion of female patients (n = 730, 63.0%), as well as individuals who underwent TKA (n = 635, 54.8%). Median

Table 1
Patient-level characteristics by utilization group are presented as N (%) and median (IQR) for categorical and continuous variables, respectively.

Variable	Cluster 1 N = 6527	Cluster 2 N = 8589	Cluster 3 N = 1158	Cluster 4 N = 2270	Cluster 5 N = 978
Age (years)	75 (70, 81)	74 (69, 79)	72 (68, 77)	73 (69, 78)	72 (68, 76)
Male	2644 (40.5%)	3823 (44.5%)	428 (37.0%)	1023 (45.1%)	459 (46.9%)
Race					
White	5994 (91.8%)	8083 (94.1%)	1072 (92.6%)	2137 (94.1%)	933 (95.4%)
Minority	533 (8.2%)	506 (5.9%)	86 (7.4%)	133 (5.9%)	45 (4.6%)
Charlson CI	3 (1, 4)	2 (1, 4)	3 (2, 6)	2 (0, 3)	3 (1, 5)
Elixhauser CI	3 (2, 5)	3 (2, 4)	3 (2, 4)	3 (2, 4)	3 (2, 4)
CMSHCC Score	1.65 (1.01, 2.6)	1.34 (0.74, 2.27)	1.71 (1.04, 3.04)	1.04 (0.65, 1.73)	1.74 (1.03, 3.07)
Procedure					
THA	3271 (50.1%)	4128 (48.1%)	523 (45.2%)	1029 (45.3%)	502 (51.3%)
TKA	3256 (49.9%)	4461 (51.9%)	635 (54.8%)	1241 (54.7%)	476 (48.7%)

AA: African American; CI: Comorbidity Index; CMSHCC: Centers for Medicare and Medicaid Service's Hierarchical Condition Category; THA: total hip arthroplasty; TKA: total knee arthroplasty.
All p-values < 0.001.

preoperative health expenditure of patients in Cluster 3 was roughly $3500 more than the median preoperative health expenditure of patients in Cluster 2 ($41,730, IQR: $32,980-$55,560). Interestingly, Cluster 3 patients had the lowest number of preoperative inpatient encounters (median: 0; IQR: 0–2), while having the highest number of outpatient encounters (median: 22; IQR: 15–33). In addition, super-utilizer patients in Cluster 3 had the highest preoperative healthcare utilization associated with pharmacy charges ($88,610, IQR: $65,310–124,430) with relatively low charges for the majority of other care categories. Of note, Cluster 4 patients (n = 2,270, 11.6%) had a median preoperative health expenditure similar to Cluster 3 patients (Cluster 4: $41,670, IQR: $32,520-$55,260 vs. Cluster 3: $41,730, IQR: $32,980-$55,560). However, unlike Cluster 3 patients, patients in Cluster 4 had the highest medical/surgical supplies charges ($100,780, IQR: $82,210–132,430) and surgical charges ($32,630, IQR 16,690–55,010) prior to the THA/TKA index admission compared with all other clusters.

Cluster 5 was the smallest cohort (n = 978, 5.0%) and had the lowest proportion of ethnic/racial minority (4.6%; n = 45) patients, as well as the lowest proportion of individuals who underwent TKA (48.7%; 476). Of note, compared with other clusters, individuals in Cluster 5 had the highest comorbidity burden according to the Centers for Medicare and Medicaid Service's Hierarchical Condition Category (OR 1.74, IQR 1.03–3.07). Compared with other super-utilizers, patients in Cluster 5 had by far the highest annual preoperative expenditure (median: $67,292; IQR: $46,470-$100,740). Of note, patients in Cluster 5 had high utilization for several categorizes of care including high pharmacy charges ($133,360, IQR: $20,260–252,240), lab services ($13,330, IQR$ 5670–38,610), and diagnostic radiology services ($13,010, IQR: $4920–31,450).

Overall the classification tree algorithm identified charge categories with high accuracy (83.4% correct estimation) (Fig. 1). Specific clusters included costs for supplies, pharmacy, and intensive care unit utilization (Table 4). For example, patients in Cluster 1 were generally characterized by low costs for supplies (<$20,000), pharmacy (<$54,000), and intensive care unit utilization (<$23,000). In contrast, Cluster 5 patients were characterized as having low supply costs (<$20,000), yet very high pharmacy costs (≥$180,000).

Clinical characteristics and outcomes among the various super-utilization clusters

The most common comorbidity present in the study was chronic pulmonary disease (n = 8,509, 43.6%) followed by congestive heart failure (n = 6,956, 35.6%) (Supplemental Table 2). The rate of comorbidities varied by cluster; particularly malignancy

Table 2
Clinical outcomes by utilization group are presented as N (%) and median (IQR) for categorical and continuous variables, respectively.

Variable	Cluster 1 N = 6527	Cluster 2 N = 8589	Cluster 3 N = 1158	Cluster 4 N = 2270	Cluster 5 N = 978
	Pre-operative Outcomes				
Expenditure: pre-surgery (kUSD)	34.66 (30.51, 42.3)	39.03 (32.01, 51.58)	41.73 (32.98, 55.56)	41.67 (33.52, 55.26)	67.29 (46.47, 100.74)
Inpatient Encounters	2 (1, 3)	2 (1, 3)	0 (0, 2)	1 (1, 2)	1 (0, 3)
Outpatient Encounters	12 (6, 20)	11 (6, 19)	22 (15, 33)	10 (5, 16)	17 (9, 28)
	Operative Outcomes				
Expenditure: surgical (kUSD)	13.88 (12.33, 16.80)	13.22 (11.91, 15.16)	13.27 (11.96, 15.64)	12.76 (11.60, 14.46)	12.99 (11.64, 15.07)
LOS (days)	3 (3, 4)	3 (3, 4)	3 (3, 4)	3 (2, 3)	3 (3, 4)
Any Complication at Index	1079 (16.5%)	889 (10.4%)	93 (8.0%)	156 (6.9%)	117 (12.0%)
	Post-operative Outcomes				
Expenditure: post-surgical (kUSD)	14.73 (2.67, 32.8)	10.14 (1.76, 23.04)	29.24 (11.68, 47.46)	6.71 (1.38, 18.86)	21.14 (4.42, 52.34)
Inpatient Encounters	1 (0, 2)	1 (0, 1)	0 (0, 1)	0 (0, 1)	1 (0, 2)
Outpatient Encounters	9 (4, 17)	8 (4, 15)	17 (10, 26)	7 (3, 13)	12 (5, 21)
Readmission 90day	1584 (24.3%)	1804 (21.0%)	214 (18.5%)	379 (16.7%)	224 (22.9%)
Readmission 30day	835 (12.8%)	942 (11.0%)	109 (9.4%)	205 (9.0%)	118 (12.1%)
Mortality 90day	170 (2.6%)	155 (1.8%)	30 (2.6%)	23 (1.0%)	31 (3.2%)
Mortality 30day	64 (1.0%)	51 (0.6%)	N/A	N/A	N/A

LOS: Length of stay; kUSD: U.S. Dollars in Thousands.
N/A: Cell counts less than 10.
All p-values < 0.001.

Table 3
Preoperative utilization (in thousand US dollars) by utilization group are presented as median (IQR).

Variable	Cluster 1 N = 6527	Cluster 2 N = 8589	Cluster 3 N = 1158	Cluster 4 N = 2270	Cluster 5 N = 978
Medical/Surgical Supplies	6.41 (1.91, 13.61)	31.52 (18.07, 47.62)	1.62 (0.07, 6.82)	100.78 (82.21, 132.43)	26.03 (1.13, 179.47)
Surgery	8.10 (0.49, 15.43)	29.31 (16.11, 45.38)	2.31 (0, 13.09)	32.63 (16.69, 55.01)	21.73 (0, 83.77)
Other	11.07 (4.89, 20.44)	15.39 (7.29, 28.89)	4.87 (1.41, 13.00)	11.67 (4.75, 24.60)	15.26 (4.34, 46.93)
Pharmacy	6.10 (2.94, 12.53)	7.90 (3.79, 16.36)	88.61 (65.31, 124.43)	5.98 (2.79, 12.30)	133.36 (20.26, 252.24)
Lab	6.92 (3.54, 12.08)	9.25 (4.25, 18.13)	5.73 (2.55, 13.04)	5.68 (2.55, 12.30)	13.33 (5.67, 38.61)
Diagnostic Radiology	6.65 (2.78, 13.00)	8.42 (3.81, 16.63)	7.27 (2.11, 17.31)	6.51 (2.71, 12.50)	13.01 (4.92, 31.45)
Emergency	2.16 (0.58, 4.50)	1.86 (0.00, 4.69)	0.61 (0, 2.77)	0.68 (0, 2.82)	1.40 (0, 5.44)
Routine	0 (0, 4.49)	0 (0, 6.5)	0 (0, 0)	0 (0, 3.10)	0 (0, 7.14)
ICU	0 (0, 1.82)	0 (0, 7.53)	0 (0, 0)	0 (0, 4.45)	0 (0, 19.07)
Respiratory Therapy	0 (0, 0.58)	0 (0, 1.60)	0 (0, 0)	0 (0, 0.43)	0 (0, 3.1)
Clinic	0 (0, 0.44)	0 (0, 0.41)	0.37 (0, 1.50)	0 (0, 0.30)	0.15 (0, 1.11)
IV Therapy	0 (0, 0.37)	0 (0, 0.29)	0.95 (0, 3.59)	0 (0, 0)	0.08 (0, 3.56)
Pathology	0 (0, 0.28)	0 (0, 0.51)	0 (0, 0.82)	0 (0, 0.25)	0 (0, 1.01)
Nuclear Medicine	0 (0, 0.11)	0 (0, 0.35)	0 (0, 1.24)	0 (0, 0)	0 (0, 1.48)
Chemotherapy	0 (0, 0)	0 (0, 0)	1.74 (0, 5.76)	0 (0, 0)	0 (0, 2.83)
Radiation Therapy	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)

Dx: Diagnostic; ICU: Intensive Care Unit; IV: Intravenous.

status and dementia. Malignancies were prevalent at a rate 4.3 times higher for Cluster 3 compared with Cluster 4 (52.1% vs. 12.2%). Further, dementia occurred 3.6 times more often in Cluster 1 compared with Cluster 5 (6.1% vs. 1.7%) (both $p < 0.001$).

Although Cluster 1 had the lowest annual preoperative utilization (median expenditure: $34,660), roughly 1 in 6 (n = 1,079, 16.5%) individuals in this cluster experienced a complication during the index hospitalization; Cluster 1 patients also had a relatively high incidence of 30- (n = 835, 12.8%) and 90-day (n = 1584, 24.3%) readmission (Table 2). Interestingly, despite having the highest annual preoperative utilization (median expenditure: $67,292), patients in Cluster 5 had a similar incidence of complications (n = 117, 12.0%) and 30-day readmission (n = 118, 12.1%) as Cluster 1 patients. In addition, patients in Cluster 1 (n = 170, 2.6%) and Cluster 5 (n = 31, 3.2%) had a comparable incidence of 30-day mortality.

Discussion

Healthcare spending is increasingly being scrutinized, especially among high healthcare utilizers. In particular, there has been a growing focus on so-called "super-utilizers" – that subset of patients who represent the most extreme group of patients with the highest healthcare expenses.[2,8,9,12,30,31] The overwhelming majority of previous reports have focused on factors associated with high-cost patients centered around the perioperative period.[7,32,33] Specifically, comorbidity burden, presence of congestive heart failure and male sex have been associated with higher than average Medicare spending.[32,33] To date, however, no study has investigated specific patterns of preoperative utilization among super-utilizers or characterized differences in patterns of preoperative utilization relative to clinical outcomes. Therefore, the current study was important as it examined the broader scope of healthcare utilization surrounding a surgical episode specifically among super-utilizers undergoing TKA or THA. Of note, there was a large degree of heterogeneity within the preoperative super-utilizer population with specific differences in the type of utilization (e.g. pharmacy, lab services, intensive care unit, and diagnostic radiology services). Variations in utilization patterns among super-utilizers suggested that utilization patterns among all super-utilizers was not homogenous. Rather, the data suggested that different targeted interventions may be needed to identify specific areas of cost control among different clusters of super-utilizers.

Previous studies have suggested that various patterns of utilization prior to surgery may be associated with different clinical outcomes. For example, Leeds and colleagues noted that individuals who met with a non-surgeon provider prior to their surgery, despite having increased preoperative total costs, had reduced odds of a perioperative complication and postoperative costs.[6] In a separate study, Barakat et al., reported that participation in fee-based exercise programs at a physiotherapy gym prior to abdominal aortic aneurysm repair was associated with reduced rates of perioperative complications, as well as reduced length of stay.[3] In the current study, we noted that increased spending, as well as variations in the type of spending, among super-utilizers was variably associated with improved perioperative outcomes. Specifically, super-utilizers in Cluster 1 and Cluster 5 had different total expenditures, as well as variations in how costs were associated with different care categories. Despite this, clinical outcomes such as readmission and mortality were comparable among patients in Cluster 1 and Cluster 5. In contrast, patients in Cluster 4 who had highest median preoperative healthcare expenditures characterized by high medical/surgical supplies were noted to have among the lowest complication rates (Table 2). Compared with studies that examined patients across a wide spectrum of health care costs (e.g. low-, medium-, and super-utilizers),[3,6] the current study focused exclusively on super-utilizers. As such, the data suggest that among patients who were already in the top 5% of utilization and costs, that variations in the amount and type of utilization may be associated with outcomes. As such, rather than simply using "hot spotting" techniques to identify super-utilizers, it is important to identify specific patterns of healthcare spending within this super-utilizer cohort of patients itself.[32,33]

Another interesting finding was that the effect of surgery (i.e. THA, TKA) on post-surgical healthcare expenditures varied by cluster category. For example, patients in Cluster 4 had an 84%, roughly $35,000, reduction in healthcare spending following surgery. Interestingly, patients in Cluster 3 had preoperative healthcare utilization driven by medical/surgical supplies and prior surgical expenses. As such, the change in pre-versus post-operative surgical spending was not as dramatic compared to patients in other clusters (i.e. Cluster 3 Δ$12,500).

The varied effect of surgery on post-operative expenditures the year following surgery across the clusters may have been due to the variation in how patients were utilizing the healthcare system prior to surgery, as well as the different reasons associated with preoperative high costs (e.g. pharmacy vs. supplies vs. in-patient hospital-based costs). Previous work from our group note that,

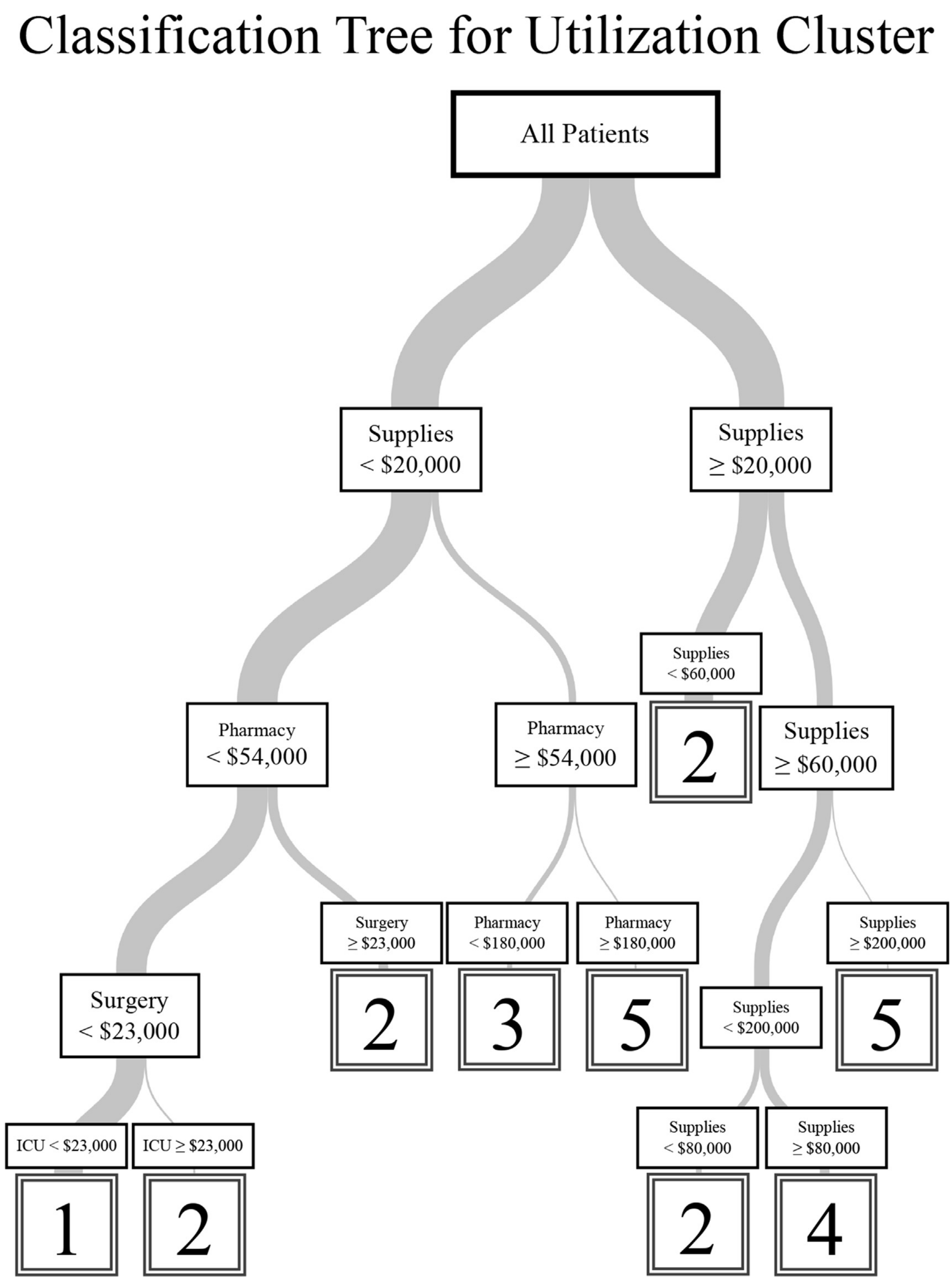

Fig. 1. Classification tree used to assign patients into a utilization pattern cluster.

while median spending per Medicare beneficiary in the year after surgery was higher for super-utilizers compared with low-utilizers [$1837 (IQR: $341-$11,390) vs $18,223 (IQR: $3466-$43,356)], super-utilizers accounted for 13.5% of total postoperative spending.[10] The reduction in adjusted average annual Medicare expenditure ranged from >$15,000 per year for patients undergoing coronary artery bypass grafting to approximately $30,000 per year for patients undergoing a hip replacement. Interesting, a similar variation in healthcare spending reduction had been reported among patients undergoing bariatric surgery.[34] Specifically, Weiner et al. noted marked variations in cost-savings/reduction in health care expenses within the first few years following bariatric surgery.[34] Of note, after a prolonged period of time (i.e. 6 years) the relative benefit/cost-savings associated with bariatric surgery stabilized among paitents.[34] Collectively, the data suggest that among a subset of super-utilizers, surgical intervention was associated

Table 4
Results of discriminant analysis and classification tree algorithm are provided by frequency (relative row frequency) where shaded cells indicate correct cluster assignment from the classification tree algorithm.

Predicted Utilization Cluster	Actual Utilization Cluster					
	1	2	3	4	5	Total
1	5,273 (89.8%)	469 (8.0%)	115 (2.0%)	0 (0.0%)	18 (0.3%)	5,875
2	1,236 (12.6%)	7,748 (79.2%)	115 (1.2%)	488 (6.1%)	193 (2.0%)	9,780
3	18 (1.6%)	136 (12.3%)	920 (83.0%)	0 (0.0%)	34 (3.1%)	1,108
4	0 (0.0%)	236 (11.3%)	0 (0.0%)	1,729 (82.8%)	123 (5.9%)	2,088
5	0 (0.0%)	0 (0.0%)	8 (1.2%)	53 (7.9%)	610 (90.9%)	671
Total	6,527	8,589	1,158	2,270	978	19,522

with a reduction in annual Medicare expenditure in the year after surgery.

The current study utilized a relatively novel approach to examine heterogeneity in healthcare spending within super-utilizers. Previous studies have utilized cluster analysis and discriminant analysis in tandem to cluster observations, as well as identify and discriminate factors.[29,35–38] For example, Soler et. utilized cluster analysis followed by discriminant analysis to identify an algorithm for classifying patients to predict variations clinical outcomes.[29] In a different study, Mariampillai et al. utilized hierarchical clustering analysis to aggregate patients into different idiopathic inflammatory myopathy schemas using phenotypic, biological, and immunologic criteria.[38] Using this methodological approach, subgroups based on clinical-serologic data were classified into four unique subgroups.[38] To the best of our knowledge, the present study was the first study to employ this methodology to examine healthcare utilization patterns among super-utilizers in order to identify discriminant healthcare spending categories. The current study identified five distinct cohorts characterized by unique utilization patterns of spending in supplies, pharmacy, and, among others, intensive care unit. The resultant classification tree was overall able to classify individuals into clusters with 83.4% correct estimation. In turn, the identification of these healthcare utilization patterns may help the Centers for Medicare and Medicaid Services detect individuals with utilization patterns that may be modifiable. In particular, the concept of utilization efficiency (i.e. specific types/patterns of preoperative utilization that can lead to more positive clinical outcomes) may be applied to certain sub-populations of super-utilizers.

Several limitations should be considered when interpreting the currents results. Primarily, the SAFs did not contain detailed clinical information and therefore nuanced differences between super-utilizer subpopulations may have been undetected. In addition, the variables used to cluster patients were determined from the revenue SAF that provides an itemized list of billed services to Medicare and not actual amount reimbursed for specific procedures and treatments by Medicare. As such, we could not determine a detailed estimate of how much each charge was reimbursed. The cohort of patients also consisted exclusively of Medicare patients aged 65 and older; as such, the results may not be generalizable to a younger population or to individuals with health insurance other than Medicare. The study did, however, have multiple strengths. In particular, super-utilizers were identified using a statistically rigorous approach rather than implementing arbitrary cutoff of quartiles, quintiles, or deciles. Rather, a k-medians cluster analysis was used to identify super-utilizers. In addition, a form of unsupervised machine learning, cluster analysis, was utilized to group super-utilizers by patterns of preoperative utilization in order to minimize within cluster variation while maximizing between cluster variability. Using discriminant analysis and classification tree analysis, two methods of machine learning, in tandem, allowed for a statistically sound and rigorous technique to develop an easily implemented decision tree for categorizing super-utilizers into specific clusters.

In conclusion, while super-utilizers consist of only a small proportion of the entire surgical patient population, these patients represent the most extreme of high healthcare utilizers with a heavy burden on Medicare. Data from the current study demonstrated that super-utilizers were not a homogenous group of patients. Rather, distinct sub-populations of super-utilizers were identified with distinct patterns of preoperative healthcare utilization. Patients had variable postoperative clinical outcomes depending on patterns of preoperative utilization; of note, not all high utilization was universally associated with poor outcomes suggesting some measure of utilization efficiency. Further studies should investigate the role of utilization efficiency to determine how to target various patterns of utilization among super-utilizers and possibly among the general surgical Medicare population to control costs while improving quality of patient care.

Declaration of competing interest

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.amjsurg.2020.01.043.

References

1. Martin AB, Hartman M, Washington B, Catlin A. National health care spending in 2017: growth slows to post-great recession rates; Share of GDP stabilizes. *Health Aff*. 2019 Jan;38(1), 101377hlthaff201805085. PubMed PMID: 30521399. Epub 2018/12/07.
2. Mitchell EM, Machlin SR. *Concentration of Health Expenditures and Selected Characteristics of High Spenders*. U.S. Civilian Noninstitutionalized Population; 2015. Statistical Brief (Medical Expenditure Panel Survey (US)). Rockville (MD) 2001.
3. Barakat HM, Shahin Y, Khan JA, et al. Preoperative supervised exercise improves outcomes after elective abdominal aortic aneurysm repair: a randomized controlled trial. *Ann Surg*. 2016 Jul;264(1):47–53. PubMed PMID: 26756766. Epub 2016/01/13.
4. Gillis C, Li C, Lee L, et al. Prehabilitation versus rehabilitation: a randomized control trial in patients undergoing colorectal resection for cancer. *Anesthesiology*. 2014 Nov;121(5):937–947. PubMed PMID: 25076007. Epub 2014/07/31.

5. Harari D, Hopper A, Dhesi J, et al. Proactive care of older people undergoing surgery ('POPS'): designing, embedding, evaluating and funding a comprehensive geriatric assessment service for older elective surgical patients. *Age Ageing*. 2007 Mar;36(2):190–196. PubMed PMID: 17259638. Epub 2007/01/30.
6. Leeds IL, Canner JK, Gani F, et al. Increased healthcare utilization for medical comorbidities prior to surgery improves postoperative outcomes. *Ann Surg*. 2018 Jun 1;271(1):114–121. PubMed PMID: 29864092. Epub 2018/06/05.
7. Gawande A. The hot spotters: can we lower medical costs by giving the neediest patients better care? *New Yorker*. 2011 Jan:40–51. PubMed PMID: 21717802. Epub 2011/07/02. eng.
8. Office CB. *High-Cost Medicare Beneficiaries*. 2005.
9. Stanton MWRM. *The High Concentration of U.S. Health Care Expenditures*. Agency for Healthcare Research and Quality.; 2005.
10. Hyer JM, Ejaz A, Diaz A, et al. *Characterizing and Assessing the Impact of Surgery on Healthcare Spending Among Medicare Enrolled Preoperative Super-utilizers*. Ann Surg; 2019 Jul 10. PubMed PMID: 31305286. Epub 2019/07/16.
11. Hyer JM, Ejaz A, Tsilimigras D, et al. A novel machine learning approach to identify preoperative risk factors of super utilization following surgery: a study in Medicare expenditure. JAMA Surg. In Press.
12. Rinehart DJ, Oronce C, Durfee MJ, et al. Identifying subgroups of adult super-utilizers in an urban safety-net system using latent class Analysis: implications for clinical practice. *Med Care*. 2018 Jan;56(1):e1–e9. PubMed PMID: 27632768. Pubmed Central PMCID: PMC5406260. Epub 2016/09/16.
13. Johnson TL, Rinehart DJ, Durfee J, et al. For many patients who use large amounts of health care services, the need is intense yet temporary. *Health Aff*. 2015 Aug;34(8):1312–1319. PubMed PMID: 26240244. Epub 2015/08/05. eng.
14. Wolf BJ, Ramos PS, Hyer JM, et al. An analytic approach using candidate gene selection and logic forest to identify gene by environment interactions (G x E) for systemic lupus erythematosus in african Americans. *Genes*. 2018 Oct 15;9(10). PubMed PMID: 30326636. Pubmed Central PMCID: PMC6211136. Epub 2018/10/18. eng.
15. Hyer JM, White S, Cloyd J, et al. Can we improve prediction of adverse surgical outcomes? Development of a surgical complexity score using A novel machine learning technique. *J Am Coll Surg*. 2020;230(1):43–52.
16. Gordon L, Austin P, Rudzicz F, Grantcharov T. MySurgeryRisk and machine learning: a promising start to real-time clinical decision support. *Ann Surg*. 2019 Jan;269(1):e14–e15. PubMed PMID: 30048316. Epub 2018/07/27.
17. Benke KK. Data analytics and machine learning for disease identification in electronic health records. *JAMA Ophthalmol*. 2019 May 1;137(5):497–498. PubMed PMID: 30789647. Epub 2019/02/23.
18. Beam AL, Kohane IS. Big data and machine learning in health care. *J Am Med Assoc*. 2018 Apr 3;319(13):1317–1318. PubMed PMID: 29532063. Epub 2018/03/14.
19. Chiew CJ, Liu N, Wong TH, et al. Utilizing machine learning methods for preoperative prediction of postsurgical mortality and intensive care unit admission. *Ann Surg*. 2019 Apr 8. PubMed PMID: 30973386. Epub 2019/04/12. In press.
20. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol*. 2019 May;20(5):e262–e273. PubMed PMID: 31044724. Epub 2019/05/03.
21. Deo RC. Machine learning in medicine. *Circulation*. 2015 Nov 17;132(20):1920–1930. PubMed PMID: 26572668. Pubmed Central PMCID: PMC5831252. Epub 2015/11/18.
22. Hastie T, Tibshirani R, Friedman JH. *The Elements of Statistical Learning : Data Mining, Inference, and Prediction*. second ed. New York, NY: Springer; 2009:745. xxii.
23. Nathan H, Thumma JR, Ryan AM, Dimick JB. Early impact of Medicare accountable care organizations on inpatient surgical spending. *Ann Surg*. 2019 Feb;269(2):191–196. PubMed PMID: 29771724. Epub 2018/05/18. eng.
24. Tan P-N, Steinbach M, Karpatne A, Kumar V. *Introduction to Data Mining*. second ed. Ny Ny: Pearson; 2019:839. xix.
25. Quan H, Li B, Couris CM, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol*. 2011 Mar 15;173(6):676–682. PubMed PMID: 21330339. Epub 2011/02/19.
26. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care*. 1998 Jan;36(1):8–27. PubMed PMID: 9431328. Epub 1998/02/07.
27. Pope GC, Kautter J, Ellis RP, et al. Risk adjustment of Medicare capitation payments using the CMS-HCC model. *Health Care Financ Rev*. 2004 Summer;25(4):119–141. PubMed PMID: 15493448. Pubmed Central PMCID: PMC4194896. Epub 2004/10/21.
28. Institute HCC. *Health Care Cost and Utilization Report. healthcostinstitute.Org*. 2017:2019.
29. Soler ZM, Hyer JM, Ramakrishnan V, et al. Identification of chronic rhinosinusitis phenotypes using cluster analysis. *Int Forum Allergy Rhinol*. 2015 May;5(5):399–407. PubMed PMID: 25694390. Pubmed Central PMCID: PMC4428937. Epub 2015/02/20.
30. Harris LJ, Graetz I, Podila PS, et al. Characteristics of hospital and emergency care super-utilizers with multiple chronic conditions. *J Emerg Med*. 2016 Apr;50(4):e203–e214. PubMed PMID: 26472609. Epub 2015/10/17. eng.
31. Jiang HJ, Weiss AJ, Barrett ML, Sheng M. *Characteristics of Hospital Stays for Super-utilizers by Payer, 2012 Statistical Brief #190. Healthcare Cost and Utilization Project (HCUP) Statistical Briefs*. 2015.
32. Merath K, Chen Q, Johnson M, et al. Hot spotting surgical patients undergoing hepatopancreatic procedures. *HPB*. 2019 Jun;21(6):765–772. PubMed PMID: 30497897. Epub 2018/12/01. eng.
33. Shubeck SP, Thumma JR, Dimick JB, Nathan H. Hot spotting as a strategy to identify high-cost surgical populations. *Ann Surg*. 2019 Jan;269(3):453–458. PubMed PMID: 29342019. Epub 2018/01/16. eng.
34. Weiner JP, Goodwin SM, Chang HY, et al. Impact of bariatric surgery on health care costs of obese persons: a 6-year follow-up of surgical and comparison cohorts using health plan data. *JAMA Surg*. 2013 Jun;148(6):555–562. PubMed PMID: 23426865. Epub 2013/02/22.
35. Panagopoulos GP, Angelopoulou D, Tzirtzilakis EE, Giannoulopoulos P. The contribution of cluster and discriminant analysis to the classification of complex aquifer systems. *Environ Monit Assess*. 2016 Oct;188(10):591. PubMed PMID: 27670889. Epub 2016/09/28.
36. Petrescu AM, Ilia G. Molecular docking study to evaluate the carcinogenic potential and mammalian toxicity of thiophosphonate pesticides by cluster and discriminant analysis. *Environ Toxicol Pharmacol*. 2016 Oct;47:62–78. PubMed PMID: 27636985. Epub 2016/10/31.
37. Novak M, Palya D, Bodai Z, et al. Combined cluster and discriminant analysis: an efficient chemometric approach in diesel fuel characterization. *Forensic Sci Int*. 2017 Jan;270:61–69. PubMed PMID: 27915188. Epub 2016/12/05.
38. Mariampillai K, Granger B, Amelin D, et al. Development of a new classification system for idiopathic inflammatory myopathies based on clinical manifestations and myositis-specific autoantibodies. *JAMA Neurol*. 2018 Dec 1;75(12):1528–1537. PubMed PMID: 30208379. Pubmed Central PMCID: PMC6583199. Epub 2018/09/13.