# Supplementary Online Content

Hyer JM, Ejaz A, Tsilimigras DI, Paredes AZ, Mehta R, Pawlik TM. Novel machine learning approach to identify preoperative risk factors associated with super-utilization of Medicare expenditure following surgery. *JAMA Surg*. Published online August 14, 2019. 10.1001/jamasurg.2019.2979

**eTable 1.** Charlson and Elixhauser Comorbidities in the Crude Population and the Propensity-Matched Population

**eTable 2.** Patient Demographics and Pre-Operative Characteristics for Low- and Super-Utilizers in the Crude Population and the Propensity-Matched Population

This supplementary material has been provided by the authors to give readers additional information about their work.

**eTable 1.** Charlson and Elixhauser Comorbidities in the Crude Population and the Propensity-Matched Population

| | Crude Population | | | Propensity-Matched Population | | |
|---|---|---|---|---|---|---|
| | Low Utilizer | Super Utilizer | Effect Size | Low Utilizer | Super Utilizer | Effect Size |
| | | Charlson Comorbidities | | | | |
| COPD | 124,591 (22.8%) | 49,722 (62.4%) | 0.294 | 14,337 (40.4%) | 8,447 (48.3%) | 0.075 |
| CHF | 39,537 (7.2%) | 51,037 (64%) | 0.538 | 7,574 (21.3%) | 6,316 (36.1%) | 0.158 |
| Leukemia/Lymphoma/ Localized Solid Tumor | 94,099 (17.2%) | 26,249 (32.9%) | 0.175 | 14,777 (41.6%) | 6,304 (36.0%) | 0.051 |
| CKD | 57,496 (10.5%) | 42,646 (53.5%) | 0.391 | 8,133 (22.9%) | 5,628 (32.2%) | 0.099 |
| Metastatic Solid Tumor | 13,212 (2.4%) | 9,829 (12.3%) | 0.133 | 6,401 (18%) | 2,444 (14.0%) | 0.054 |
| Mild Liver Disease | 25,552 (4.7%) | 12,687 (15.9%) | 0.156 | 4,504 (12.7%) | 2,025 (11.6%) | 0.016 |
| DM | 17,082 (3.1%) | 16,374 (20.5%) | 0.258 | 2,376 (6.7%) | 1,989 (11.4%) | 0.080 |
| CTD | 31,916 (5.8%) | 7,022 (8.8%) | 0.041 | 2,858 (8.1%) | 1,758 (10.1%) | 0.033 |
| Hemi- or Para-plegia | 1,747 (0.3%) | 5,577 (7%) | 0.207 | 368 (1%) | 994 (5.7%) | 0.138 |
| Dementia | 4,765 (0.9%) | 4,014 (5%) | 0.118 | 897 (2.5%) | 682 (3.9%) | 0.038 |
| Severe Liver Disease | 1,131 (0.2%) | 1,906 (2.4%) | 0.105 | 425 (1.2%) | 154 (0.9%) | 0.014 |
| AIDS | 142 (0.0%) | 111 (0.1%) | 0.019 | 28 (0.1%) | 18 (0.1%) | 0.004 |
| | | Elixhauser | | | | |
| Uncomplicated HTN | 354,665 (64.8%) | 38,916 (48.8%) | 0.111 | 22,010 (62.0%) | 10,750 (61.5%) | 0.005 |

| | | | | | |
|---|---|---|---|---|---|
| Cardiac Arrhythmia | 87,303 (16%) | 42,189 (52.9%) | 0.304 | 9,663 (27.2%) | 5,403 (30.9%) | 0.038 |
| Fluid & Electrolyte Disorders | 54,038 (9.9%) | 36,973 (46.4%) | 0.345 | 6,509 (18.3%) | 4,187 (23.9%) | 0.066 |
| Hypothyroidism | 105,419 (19.3%) | 13,249 (16.6%) | 0.023 | 6,707 (18.9%) | 3,306 (18.9%) | 0.000 |
| Obesity | 84,140 (15.4%) | 14,077 (17.7%) | 0.021 | 5,871 (16.5%) | 2,898 (16.6%) | 0.000 |
| Renal Failure | 35,447 (6.5%) | 24,600 (30.8%) | 0.276 | 5,016 (14.1%) | 2,635 (15.1%) | 0.013 |
| Complicated HTN | 35,120 (6.4%) | 23,596 (29.6%) | 0.265 | 4,904 (13.8%) | 2,559 (14.6%) | 0.011 |
| Peripheral Vascular Disorders | 23,753 (4.3%) | 19,999 (25.1%) | 0.271 | 3,900 (11.0%) | 2,319 (13.3%) | 0.033 |
| Valvular Disease | 27,918 (5.1%) | 20,642 (25.9%) | 0.259 | 2,751 (7.7%) | 1,953 (11.2%) | 0.056 |
| Metastatic Cancer | 7,185 (1.3%) | 4,915 (6.2%) | 0.117 | 4,095 (11.5%) | 1,240 (7.1%) | 0.069 |
| Weight Loss | 3,872 (0.7%) | 13,954 (17.5%) | 0.337 | 1,081 (3.0%) | 1,081 (6.2%) | 0.075 |
| Rheumatoid Arthritis | 22,758 (4.2%) | 4,240 (5.3%) | 0.019 | 1,975 (5.6%) | 1,078 (6.2%) | 0.012 |
| Coagulopathy | 11,980 (2.2%) | 14,897 (18.7%) | 0.271 | 1,586 (4.5%) | 1,029 (5.9%) | 0.031 |
| Other Neurologic Disorders | 9,969 (1.8%) | 9,401 (11.8%) | 0.192 | 986 (2.8%) | 936 (5.4%) | 0.065 |
| Pulmonary Circulation Disorders | 5,899 (1.1%) | 8,074 (10.1%) | 0.204 | 1,092 (3.1%) | 606 (3.5%) | 0.010 |
| Drug Abuse | 8,178 (1.5%) | 2,837 (3.6%) | 0.036 | 1,083 (3.1%) | 525 (3.0%) | 0.003 |
| Lymphoma | 2,392 (0.4%) | 1,471 (1.8%) | 0.060 | 364 (1.0%) | 407 (2.3%) | 0.051 |
| Blood Loss Anemia | 6,552 (1.2%) | 1,904 (2.4%) | 0.034 | 803 (2.3%) | 372 (2.1%) | 0.004 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Liver Disease | 5,401 (1.0%) | 3,690 (4.6%) | 0.101 | 979 (2.8%) | 311 (1.8%) | 0.030 |
| Alcohol Abuse | 4,594 (0.8%) | 1,914 (2.4%) | 0.051 | 567 (1.6%) | 284 (1.6%) | 0.001 |
| Drug Abuse | 3,235 (0.6%) | 1,200 (1.5%) | 0.052 | 401 (1.1%) | 211 (1.2%) | 0.001 |
| Psychoses | 2,852 (0.5%) | 1,039 (1.3%) | 0.033 | 355 (1.0%) | 207 (1.2%) | 0.008 |
| Peptic Ulcer Disease (excl. bleeding) | 2,608 (0.5%) | 801 (1%) | 0.024 | 241 (0.7%) | 151 (0.9%) | 0.010 |
| Paralysis | 375 (0.1%) | 1,753 (2.2%) | 0.122 | 66 (0.2%) | 104 (0.6%) | 0.034 |
| | | **Additional** | | | | |
| Fibromyalgia | 15,736 (2.9%) | 1,363 (2.7%) | 0.002 | 1,017 (2.9%) | 630 (3.6%) | 0.020 |

COPD: Chronic Obstructive Pulmonary Disease; CHF: Congestive Heart Failure; CKD: Chronic Kidney Disease; DM: Diabetes Mellitus; AIDS: Autoimmune Deficiency Syndrome; CTD: Connective Tissue Disease; HTN: hypertension.

**eTable 2.** Patient Demographics and Pre-Operative Characteristics for Low- and Super-Utilizers in the Crude Population and the Propensity-Matched Population

| Variable | Training Cohort | | | | Validation Cohort | | | |
|---|---|---|---|---|---|---|---|---|
| | Total<br>N = 26,493 | Low Utilizer<br>N = 17,794 | Super Utilizer<br>N = 8,699 | Effect Size | Total<br>N = 26,493 | Low Utilizer<br>N = 17,703 | Super Utilizer<br>N = 8,790 | Effect Size |
| Age | 74 (69, 79) | 74 (69, 79) | 74 (69, 79) | -0.004 | 74 (69, 79) | 74 (69, 79) | 74 (69, 79) | -0.001 |
| Age Category | | | | 0.027 | | | | 0.029 |
| 65 - 69 | 6,329 (23.9%) | 4,388 (24.7%) | 1,941 (22.3%) | | 6,407 (24.2%) | 4,412 (24.9%) | 1,995 (22.7%) | |
| 70 - 74 | 7,218 (27.2%) | 4,799 (27.0%) | 2,419 (27.8%) | | 7,140 (27.0%) | 4,694 (26.5%) | 2,446 (27.8%) | |
| 75 - 79 | 6,101 (23.0%) | 4,026 (22.6%) | 2,075 (23.9%) | | 6,091 (23.0%) | 3,988 (22.5%) | 2,103 (23.9%) | |
| 80 - 84 | 4,167 (15.7%) | 2,802 (15.7%) | 1,365 (15.7%) | | 4,164 (15.7%) | 2,776 (15.7%) | 1,388 (15.8%) | |
| > 84 | 2,678 (10.1%) | 1,779 (10.0%) | 899 (10.3%) | | 2,691 (10.2%) | 1,833 (10.4%) | 858 (9.8%) | |
| Male | 12,616 (47.6%) | 8,415 (47.3%) | 4,201 (48.3%) | 0.009 | 12,587 (47.5%) | 8,343 (47.1%) | 4,244 (48.3%) | 0.011 |
| Race | | | | 0.007 | | | | 0.003 |
| White | 24,348 (91.9%) | 16,369 (92.0%) | 7,979 (91.7%) | | 24,379 (92%) | 16,294 (92.0%) | 8,085 (92.0%) | |
| AA | 1,460 (5.5%) | 969 (5.4%) | 491 (5.6%) | | 1,466 (5.5%) | 975 (5.5%) | 491 (5.6%) | |
| Hispanic | 87 (0.3%) | 61 (0.3%) | 26 (0.3%) | | 76 (0.3%) | 49 (0.3%) | 27 (0.3%) | |
| Other/Unknown | 598 (2.3%) | 395 (2.2%) | 203 (2.3%) | | 572 (2.2%) | 385 (2.2%) | 187 (2.1%) | |
| Procedure | | | | 0.048 | | | | 0.053 |
| Colectomy | 6486 (24.5%) | 4,189 (23.5%) | 2,297 (26.4%) | | 6,417 (24.2%) | 4,186 (23.6%) | 2,231 (25.4%) | |
| AAA | 1,350 (5.1%) | 930 (5.2%) | 420 (4.8%) | | 1,383 (5.2%) | 920 (5.2%) | 463 (5.3%) | |
| CABG | 1,073 (4.1%) | 645 (3.6%) | 428 (4.9%) | | 1,002 (3.8%) | 572 (3.2%) | 430 (4.9%) | |
| Hip | 5,808 (21.9%) | 3,953 (22.2%) | 1,855 (21.3%) | | 5,919 (22.3%) | 3,921 (22.1%) | 1,998 (22.7%) | |
| Knee | 8,713 (32.9%) | 6,018 (33.8%) | 2,695 (31.0%) | | 8,704 (32.9%) | 5,961 (33.7%) | 2,743 (31.2%) | |

| | Training Cohort | | | | Validation Cohort | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | Total<br>N = 26,493 | Low Utilizer<br>N = 17,794 | Super Utilizer<br>N = 8,699 | Effect Size | Total<br>N = 26,493 | Low Utilizer<br>N = 17,703 | Super Utilizer<br>N = 8,790 | Effect Size |
| Lung Resection | 3,063 (11.6%) | 2,059 (11.6%) | 1,004 (11.5%) | | 3,068 (11.6%) | 2,143 (12.1%) | 925 (10.5%) | |
| Comorbidity Indices | | | | | | | | |
| Charlson | 3 (1, 5) | 2 (1, 5) | 3 (2, 5) | -0.010 | 3 (1, 5) | 3 (1, 5) | 3 (1, 5) | -0.006 |
| Elixhauser | 3 (2, 5) | 3 (2, 4) | 3 (2, 5) | -0.007 | 3 (2, 5) | 3 (2, 5) | 3 (2, 5) | -0.006 |
| CMS-HCC | 1.31 (0.75, 2.34) | 1.24 (0.72, 2.25) | 1.44 (0.86, 2.5) | -0.008 | 1.32 (0.74, 2.37) | 1.26 (0.71, 2.31) | 1.43 (0.82, 2.45) | -0.006 |

*: Variable was used in the propensity-matching analysis