**Description of all code files**

These notes briefly highlight the purposes of the code files that are in the tables in each section. All code files listed below are shareable per our Data Use Agreement (DUA) with the Centers for Medicare and Medicaid Services (CMS) and will be posted publicly on our GitLab page.

**Software**

We used SAS 9.4, Stata/MP 15.0, and R Version 3.5.1 for this analysis.

**C1. Identifying Low-Value Services**

All programs were made to be run on the 2006-2014 Medicare data, formatted like the data housed at the National Bureau of Economic Research (NBER). These programs search for low-value services occurring between 2007 and 2014. Note that searching for low-value services in a given year (the “index” year) also requires using data in a “lookback” year, the prior year. We first created Medicare file extracts containing variables needed for low-value service screening, then searched for low-value services, created beneficiary-level covariates for analysis, and combined outputs to create a beneficiary-year-level dataset with beneficiaries’ low-value services and covariates. For each measure and each year, the program produced two sets of low-value services. One set has low-value services identified using a “specific” detection criteria and the other has low-value services identified using a “sensitive” detection criteria.

**Table S11. Files related with identifying low-value services**

|  |  |  |
| --- | --- | --- |
| Program name | Input files (File source) | Output files |
| Mcarextracts.sas  flags.sas  covars.sas  flags2.sas | Beneficiary summary files for the index year and prior year  Carrier (line and claim), outpatient (line and claim), and MedPAR claims files for the index year and prior year  BETOS to HCPCS crosswalk for the index year and prior year  car`year’\_20.sas7bdat  otpt`year’\_20.sas7bdat  medpar`year’\_20.sas7bdat  ourbenes`year’\_20.sas7bdat  car`year’\_20.sas7bdat  otpt`year’\_20.sas7bdat  medpar`year’\_20.sas7bdat  fl\_‘measure\_number’\_`year’\_`sensitivity\_level’.sas7bdat  ourbenescovars.sas7bdat | ourbenes`year’\_20.sas7bdat  car`year’\_20.sas7bdat  otpt`year’\_20.sas7bdat  medpar`year’\_20.sas7bdat  fl\_‘measure\_number’\_`year’\_`sensitivity\_level’.sas7bdat  ourbenescovars.sas7bdat  yranalysis\_20.sas7bdat |

Note: `Year’ ranges from 2007-2014. ‘Measure\_number’ ranges from 1-31. `Sensitivity\_level’ takes on values of “sensitive” and “specific”.

**C2. Identify Primary Care Providers**

We identified for each beneficiary his/her primary care provider as the provider (NPI) with whom the beneficiary had the most allowed charges on primary care claims within each year. To be conservative with our analysis, we constructed our sample based on low-value services identified using the “specific” detection criteria. We limited our sample by only including NPIs with at least 11 patients in years 2007-2014, as per our DUA with CMS.

**Table S12. Files related with identifying primary care providers**

|  |  |  |
| --- | --- | --- |
| Program name | Input files (File source) | Output files |
| AssignNPI.sas  MergeNPI.sas | Carrier (line and claim) files  claims\_npi\_20.sas7bdat  yranalysis\_20.sas7bdat | claims\_npi\_20.sas7bdat  bene\_pcp\_final.sas7bdat |

**C3. Apply Further Denominator Exclusions**

We applied finer exclusion criteria to the low-value service denominators following the denominator definitions in Table 1 of the manuscript. From the carrier, outpatient, and MedPAR claims we identified relevant claims for patients who should be excluded from each denominator for head imaging, back imaging, PAP test, PTH test, and spinal injection. We also used chronic condition flags to exclude patients with history of prostate cancer from PSA test, stroke or transient ischemic attack (TIA) from carotid artery screening.

**Table S13. Files related with applying further denominator exclusions**

|  |  |  |
| --- | --- | --- |
| Program name | Input files (File source) | Output files |
| Denomexclusion.sas  Denomexclusionflag.sas  Denommerge.sas | car`year’\_20.sas7bdat  otpt`year’\_20.sas7bdat  medpar`year’\_20.sas7bdat  bsfcc`year’.sas7bdat  car`year’\_20.sas7bdat  otpt`year’\_20.sas7bdat  bene\_pcp\_final.sas7bdat  cerv\_ex.sas7bdat  hica.sas7bdat  stroke.sas7bdat  backscan\_ex.sas7bdat  head\_ex.sas7bdat  spinj\_ex.sas7bdat | excervcar\_`year’.sas7bdat  excncrpthcar\_`year’.sas7bdat  expthcar\_`year’.sas7bdat  exrhinoctcar\_`year’.sas7bdat  exbackscancar\_`year’.sas7bdat  excervotpt\_`year’.sas7bdat  excncrpthotpt\_`year’.sas7bdat  expthotpt\_`year’.sas7bdat  exrhinoctotpt\_`year’.sas7bdat  exbackscanotpt\_`year’.sas7bdat  exrhcathmedpar\_`year’.sas7bdat  cerv\_ex.sas7bdat  hica.sas7bdat  stroke.sas7bdat  backscan\_ex.sas7bdat  head\_ex.sas7bdat  spinj\_ex.sas7bdat  bene\_pcp\_final\_ex.sas7bdat |

Note: `Year’ ranges from 2007-2014.

**C4. Create Sample Summary Statistics and Add Additional Covariates**

We calculated rates of low-value service utilization as a percentage of qualifying patients who received the service and identified the top eight low-value services with the highest utilization rates. We merged our sample of low-value services with CAHPS survey data by NPI and dropped observations for NPIs who have no CAHPS reviews. Then we created NPI deciles based on per NPI count of beneficiaries and generated 101 random samples stratifying on NPI deciles. Finally, we merged in HCC scores, chronic condition indicators, and zip-level socioeconomic variables and created eight output datasets for the top eight low-value services with the highest utilization rates.

**Table S14. Files related with creating sample summary statistics and adding additional covariates**

|  |  |  |
| --- | --- | --- |
| Program name | Input files (File source) | Output files |
| Exhibit1.sas  Sample.sas  Addvar.sas | bene\_pcp\_final\_ex.sas7bdat  bene\_pcp\_final\_ex.sas7bdat  lvs\_npisamp.sas7bdat | lvsuse\_meancount\_bin.csv  lvs\_npisamp.sas7bdat  lvscle\_backscan.sas7bdat  lvscle\_cerv.sas7bdat lvscle\_ctdasym.sas7bdat lvscle\_head.sas7bdat lvscle\_psa.sas7bdat  lvscle\_pth.sas7bdat  lvscle\_spinj.sas7bdat  lvscle\_t3.sas7bdat |

**C5. Modeling and Analysis**

To make it computationally viable, we ran a model with all fixed effects only to predict the outcome variable of whether patients received a low value service. In R, we used glm() to generate fixed effects predictions separately for each of the eight low value services on the linear (log-odds) scale. The fixed effects included patient characteristics (age, sex, race, chronic conditions indicators, and dual status), geographical characteristics, including household income, percent poverty, percent population with college/less than high school education, and percent population who live alone, year, and HRR (hospital referral regions). Predictions for each low value service at the bene-year level were then used in an interaction with low value service type in a mixed model with NPI random effects (with beneficiaries nested). We use melogit in Stata to run the three-level mixed-effects logistic regressions. The NPI-level random effects were saved as the low-value service exposure composites.

Using these low-value service exposure composites, we ranked NPIs into quintiles and deciles, for different purposes. For Figure 1, we calculated rates of receipts of individual low-value services for PCP patient populations by quintiles of the low-value service exposure composites. We centered and rescaled CAHPS measures on a 0 to 10 scale and created CAHPS score composites for interactions with the personal doctor. We regressed CAHPS measures on low value service exposure composites adjusting for patient characteristics (age, dual status, education, health status) using PROC MIXED procedure in SAS.

**Table S15. Files related with modeling and analysis**

|  |  |  |
| --- | --- | --- |
| Program name | Input files (File source) | Output files |
| Fepredict.R  addfixedpred.sas  Randomeffects.do  Exhibit3.sas  Exhibit5.sas  Regression.sas | lvscle\_backscan.sas7bdat  lvscle\_cerv.sas7bdat  lvscle\_ctdasym.sas7bdat  lvscle\_head.sas7bdat  lvscle\_psa.sas7bdat  lvscle\_pth.sas7bdat  lvscle\_spinj.sas7bdat  lvscle\_t3.sas7bdat  lvscle\_backscan.sas7bdat  lvscle\_cerv.sas7bdat  lvscle\_ctdasym.sas7bdat  lvscle\_head.sas7bdat  lvscle\_psa.sas7bdat  lvscle\_pth.sas7bdat  lvscle\_spinj.sas7bdat  lvscle\_t3.sas7bdat  lvscle\_backscan\_fe.sas7bdat  lvscle\_cerv\_fe.sas7bdat  lvscle\_ctdasym\_fe.sas7bdat  lvscle\_head\_fe.sas7bdat  lvscle\_psa\_fe.sas7bdat  lvscle\_pth\_fe.sas7bdat  lvscle\_spinj\_fe.sas7bdat  lvscle\_t3\_fe.sas7bdat  lvscle`sample’.dta  lvscle\_re`sample'.csv  lvscle\_8measures.sas7bdat  lvscle\_8measures.sas7bdat  lvscle\_re`sample'.csv  lvscle\_re`sample'.csv | lvscle\_backscan\_fe.sas7bdat  lvscle\_cerv\_fe.sas7bdat  lvscle\_ctdasym\_fe.sas7bdat  lvscle\_head\_fe.sas7bdat  lvscle\_psa\_fe.sas7bdat  lvscle\_pth\_fe.sas7bdat  lvscle\_spinj\_fe.sas7bdat  lvscle\_t3\_fe.sas7bdat  lvscle\_8measures.sas7bdat  lvscle`sample’.dta  lvscle\_re`sample'.csv  exhibit3.csv  exhibit5.csv  exhibit4.csv |

Note: `Sample’ ranges from 1-101.