**Supplemental Content D: Geographic and Hospital Variable Selection**

*Geographic Variables*

A number of zip-code-level variables were considered from the American Community Survey from the US Census bureau to capture socioeconomic status, acculturation, and education. These included median income in the patient’s census tract, bilingual status of census tract (operationalized as whether greater than 50% of the adults in the census tract speak a language other than English at home), percentage of the population living below the 200 percent poverty threshold in the patient’s census tract, and percentage with a bachelor’s degree or higher. There was significant multicollinearity between the variable of median income and the other variables including <200 percent poverty (correlation coefficient, *r=* -0.79) and percentage with bachelor’s degree or higher (correlation coefficient, *r*=0.76). There was no significant collinearity between bilingual status and median income (correlation coefficient, *r*= -0.11), thus these two variables (median income and bilingual status) were included in the final model.

*Hospital Variables*

Variables were selected from the hospital financial annual report data to capture hospital size, ED volume, surgery volume, hospital ownership and presence of residents. Hospital control (ownership) was *a priori* deemed to be significant enough to include in the analysis. The following variables were examined:

|  |  |
| --- | --- |
| Variable Category | Variable |
| Hospital Size | Licensed Hospital Beds |
| Available Hospital Beds |
| Staffed Hospital Beds |
| Surgery Volume | Total Operating Room (OR) Minutes |
| Total Surgeries |
| Number of Operating Rooms |
| ED Volume | ED Service Visits |
| Teaching Hospital Status | Total Residents (y/n) |

Collinearity was first assessed within each category for both hospital size and surgery volume that had multiple variables. See matrices below.

Hospital Size:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Licensed Beds | Available Beds | Staffed Beds |
| Licensed Hospital Beds |  |  |  |
| Available Hospital Beds | 0.98 |  |  |
| Staffed Hospital Beds | 0.91 | 0.93 |  |

Surgery Volume:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Total OR Minutes | Total Surgeries | OR Rooms |
| Total OR Minutes |  |  |  |
| Total Surgeries | 0.93 |  |  |
| Number of ORs | 0.87 | 0.86 |  |

Based on these results, one variable from each category was chosen based on clinical judgement. Collinearity of these final variables was then assessed (see table below). Ultimately the final variables selected included: number of licensed beds, number of ED visits and presence of any resident.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Licensed Hospital Beds | Total OR Minutes | ER Visits | Total Residents (y/n) |
| Licensed Hospital Beds |  |  |  |  |
| Total OR Minutes | 0.74 |  |  |  |
| ER Visits | 0.68 | 0.63 |  |  |
| Total Residents (y/n) | 0.46 | 0.50 | 0.42 |  |

*Final Variable List*

Final geographic and hospital characteristics include:

* Geographic characteristics
  + Zip-code median income (continuous)
  + Bilingual zip-code (>50% adults in the ZCTA speak a language other than English, yes/no)
  + County PCP to population rate above or below the state median (yes/no)
* Hospital characteristics
  + Hospital-control
  + Size (licensed bed number, continuous)
  + Safety-net status (proportion of Medicaid/Uninsured patients served by each hospital and defining the top quartile of hospitals as safety-net, yes/no)
  + Teaching hospital status (presence of any residents, yes/no)
  + ED volume (total annual visits, continuous).