**Supplementary Digital Content 1: Data Details**

**Paper: Comparing Driving Miles for** VA-Delivered versus VA-Purchased Cataract Surgery

Contents:

* Preparing data for GIS drive distance analysis
* Data sources

**Preparing data for GIS drive distance analysis**

Here, we list the several steps involved in preparing data for use in ESRI’s ArcGIS Network Analyst extension. This was completed using the instance of ArcGIS within the Veterans Informatics and Computing Infrastructure (VINCI). (VINCI is only accessible within the VA intranet.) That instance of ArcGIS operates at the Advanced level and includes access to the Network Analyst extension.

1. We extracted **cataract surgery events** from VA medical databases using Current Procedural Terminology (CPT) codes 66982 (complex cataract surgeries) and 66984 (routine cataract surgeries). For CC cataract surgeries, we used the Fee and the Fee Basis Claims System (FBCS) along with the Program Integrity Tool (PIT).

We organized the data to represent one surgery per eye per Veteran; in the case of multiple surgeries per eye, we selected only the first index surgery in the FY for each eye (one right and one left). We further added VA Veterans’ sociodemographic data (e.g., age, gender, race), and Veterans’ home latitude and longitude during the quarter in FY15 when they received their surgeries. For procedures that occurred at VA facilities, we added a unique VA facility identifier. For procedures that took place in CC facilities, we identified the rendering provider’s National Provider Identifier (NPI), and this, in turn, allowed us to geocode each provider’s business practice location listed in the National Plan & Provider Enumeration System (NPPES).

1. We identified a **road network dataset** for route analysis. We selected *Streetmap North America* for our *network dataset*. This network dataset was provided as part of the Esri *Data & Maps 10* (Environmental Systems Research Institute, Redlands, California, USA) data collection and is optimized for use with Esri’s *Network Analyst* extension. Network Analyst allows users to specify origins and destinations on a road network—essentially, a weighted network—and to discover and measure both the shortest and fastest routes from starting (origin) to ending (destination) points.
2. We created a GIS point file representing **Veterans’ home locations** as the origin of travel. We used the Veterans’ home longitude and latitude measurements that VA’s Planning Services & Support Group (PSSG) created using an address locator based on Esri’s *Street Maps Premium*. We used the address of record for Veterans during the fiscal quarters in which they had their surgeries. Of note, the address for a Veteran could change if surgeries occurred in different fiscal quarters. We excluded 270 surgeries (217 VA, 53 CC) that could not be geocoded.
3. For VA destinations, we located and loaded a GIS point file representing **VA facility locations**. We used the publicly available VA facility GIS point file provided by the National Center for Veterans Analysis and Statistics (see Supplemental Digital Content 1). Each VA facility in this point file has a unique facility identifier (*Sta6a*) that was linked to the unique VA facility identifier in our cataract surgery data file. We allowed any kind of VA facility (ranging from larger VA medical centers to smaller, local VA clinics) to be a destination location, as long as there was evidence a cataract surgery had taken place there during the study period. In FY15, 123 VA facilities in 48 states provided cataract surgery. We were able to identify almost all (99%) of the driving routes to the closest VA facility, 98% of driving routes to the closest CC facility, 99% of driving routes to the VA facility where the surgery occurred, and 94% of driving routes to the CC facility. With the exception of Puerto Rico, we did not have road network data for the major US territories.
4. We created an **algorithm to select the most likely CC facility address**. For CC facility destinations, we used the NPI associated with each cataract surgery to link to the provider (either rendering provider or attending provider) practice location addresses contained in the October 2014 (start of FY15) NPPES. These addresses were translated into latitude and longitude for GIS using the composite U.S. address locator (geocoded first to a street address, and failing to find that, geocoded to a ZIP Code) provided with the Esri software suite.
5. The last data prep step was to create GIS point files for the **CC facilities**. We geocoded 98.7% of 2,668 CC identified cataract providers (2,383 [89%] by street address; 281 [11%] by ZIP Code), located in all 50 states and Puerto Rico. Multiple addresses were obtained for 3% of CC surgeries. We prioritized provider NPIs for individuals over institution NPIs (corporate or LLC). If one rendering provider was listed, we used that provider’s business practice location address. If two rendering providers were listed, we used the business practice location address for the provider who was listed as having a practice specialty (taxonomy) that indicated the kind of medical expertise (e.g., ophthalmology) associated with conducting cataract surgeries. Data source details and links are provided below.

**Data Sources**

|  |
| --- |
|  |
| **Data** | **Source** |
| * Veteran demographic identifiers
* Veteran home locations (longitude & latitude)
* CPT codes indicating cataract surgery
* surgery dates
* Identifiers for surgery locations
	+ For VA, the facility station identifier (sta6a: a six-character alphanumeric facility identification code
	+ For community care providers the 10-digit National Provider Identifier (NPI))
 | VA Corporate Data Warehouse (CDW) records maintained in the VA Informatics and Computing Infrastructure (VINCI) virtual computing environment. Records include entries into a combined Fee and FBCS dataset as well as select identifiers from PIT.VINCI is available to approved VA researchers and must be accessed from within the VA intranet: <https://vaww.vinci.med.va.gov>  |
| VA facility locations (FY 2016 Q4 GIS pointfile) | National Center for Veterans Analysis and Statistics quarterly GIS pointfile. This file is under the expandable *Miscellaneous* banner (usually 2-3 objects down) at this site: <https://www.va.gov/vetdata/maps.asp> It will usually be a zipped file titled something like *ESRI Shape File of VA Facilities*. The pointfile can be directly imported into an ArcGIS geodatabase. |
| Community Care facility addresses | Provider practice location address as contained in the October 1, 2014 download of the National Plan & Provider Enumeration System (NPPES), and linked to provider using NPI. [www.nber.org/data/npi.html](http://www.nber.org/data/npi.html) |
| *StreetMap North America* (road network) & *Composite\_US* Address locator | *Esri Data & Maps 10* data product. We accessed the version available on VINCI.<https://www.esri.com/library/whitepapers/pdfs/esri-data-and-maps.pdf>  |
| Health Care Provider Taxonomy Code Set CSV | National Uniform Claim Committee archive of HCP taxonomies: <http://www.nucc.org/index.php/code-sets-mainmenu-41/provider-taxonomy-mainmenu-40/csv-mainmenu-57>  |