Hello,

I am interested in looking at the availability of access to OBGYN subspecialists over time (2008 to 2022) in the United States. The topic is physician workforce planning, where is the greatest need for more physicians to serve women 18 years and older. Patients will be driving by car.

I'm hiring because I have too much work to do it all by myself.

What I am providing:

Deliverables:

1) Source code in R

**Sample Abstract**

**Geographic Disparities in Potential Accessibility to Gynecologic Oncologists in the United States from 2008 to 2022.**

Objective: To use a spatial modeling approach to capture potential disparities of gynecologic oncologist accessibility in the United States at the block group level between 2008 and 2022.

Methods: Physician registries identified the 2008 to 2022 gynecologic oncology workforce and were aggregated to each county. The at-risk cohort (women aged 18 years or older) was stratified by race and rurality demographics. We computed the distance from at-risk women to physicians. We set drive time to 30, 60, 180 and 240 minutes.

Results: Between 2008 and 2022, the gynecologic oncology workforce increased. By 2022, there were x active physicians and x% practiced in urban block groups. Geographic disparities were identified, with x physicians per 100,000 women in urban areas compared with 0.1 physicians per 100,000 women in rural areas. In total, x block groups (x million at-risk women) lacked a gynecologic oncologist. Additionally, there was no increase in rural physicians, with only x% practicing in rural areas in 2008-2015 relative to ??% in 2016-2022 (p=?). Women in racial minority populations exhibited the lowest level of access to physicians across all time periods. For example, xx% of American Indian or Alaska Native women did not have access to a physician with a 3-hour drive 2008-2015, which did not improve over time. Black women experience an increase in relative accessibility, with a ??% increase by 2016-2022. However, Asian or Pacific Islander women exhibited significantly better access than ??? women across all time periods.

Conclusion: Although the US gynecologic oncologist workforce increased steadily over 20 years, this has not translated into evidence of improved access for many women from rural and underrepresented areas. However, health care utilization and cancer outcomes may not be influenced only by distance and availability.

Methods:

There are subspecialists in obstetrics and gynecology across the United States in four categories: Gynecologic Oncology, Female Pelvic Medicine and Reconstructive Surgery, Maternal-Fetal Medicine, and Reproductive Endocrinology and Infertility. We gathered data from the National Plan and Provider Enumeration System (NPPES) to extract each obstetrician gynecologist and their address of practice (available at <https://www.nber.org/research/data/national-plan-and-provider-enumeration-system-nppes>). The NPPES file contains a list of every physician practicing in the United States and is the closest thing to a physician data clearinghouse. This data also includes their graduation and license year. Therefore, we identified the approximate year that each subspecialist entered the workforce as a physician. We had data for NPPES files from 2013 to 2023.

Utilization data for each physician were unavailable, so our analysis focused on at-risk women aged 18 years or older in the United States at the block group level (Appendix 0). Analysis was also done at the American College of Obstetricians and Gynecologists District level (Appendix 1). We collected American Community Survey 5-year estimates and decennial census data. Along with the total adult female population in each cross-section, we also collected population on race to identify potential disparities among women in minority populations (Appendix 2). The race categories included White, Black, Asian or Pacific Islander, and American Indian or Alaska Native.

To measure the distance between the patient location and the point of care (hospital or emergency department), the shortest or "straight-line" distance (i.e., the geodetic or great circle distance) is commonly used because it can be readily calculated (e.g., through statistical software programs such as SAS®). An alternative distance metric that has been used is the driving distance or driving times that can be obtained from various mapping software such as HERE, Google Maps, MapQuest, OpenStreetMaps, and ArcGIS Network Analyst. Multiple organizations prefer drive time for measuring access such as the Veteran’s Administration, and the Department of Transportation. We use the HERE API to calculate optimal routes and directions for driving with traffic on the third Friday in October at 0900 (Appendix 3).

Additionally, our analysis examined potential access to obstetrician gynecologist subspecialists across the United States. We used drive time isochrones.

**Appendix 0:**

Choosing block groups over counties for data analysis could be driven by several factors, depending on the nature of the analysis you are conducting:

* Resolution of Data: Block groups provide a finer resolution of data compared to counties.
* Local Trends and Patterns: Block groups are small enough to reveal local trends and patterns that might be obscured when data is aggregated at the county level.
* Socioeconomic Analysis: For studies that involve socioeconomic factors, block groups can provide more precise information about demographic and economic conditions since they represent smaller communities.
* Policy Impact Assessment: When assessing the impact of local policies or interventions, block groups may be more appropriate because policies might vary significantly within a county.
* Spatial Analysis: For spatial analyses that require precision, such as identifying hotspots or conducting proximity analysis, the smaller geographic units of block groups are more useful.

We utilized the tigris package to get the block group geometries for different years. This required downloading each state’s data and joining them into one national file for years 2018 and earlier.

**Appendix 1: The American College of Obstetricians and Gynecologists (ACOG) is a professional organization representing obstetricians and gynecologists in the United States. ACOG divides its membership into various geographical regions known as "ACOG Districts."**

**State ACOG\_District State\_Abbreviations**

Alabama District VII AL

Alaska District VIII AK

Arizona District VIII AZ

Arkansas District VII AR

California District IX CA

Colorado District VIII CO

Connecticut District I CT

District of Columbia District IV DC

Florida District XII FL

Georgia District IV GA

Hawaii District VIII HI

Illinois District VI IL

Indiana District V IN

Iowa District VI IA

Kansas District VII KS

Kentucky District V KY

Louisiana District VII LA

Maryland District IV MD

Massachusetts District I MA

Michigan District V MI

Minnesota District VI MN

Mississippi District VII MS

Missouri District VII MO

Nebraska District VI NE

Nevada District VIII NV

New Hampshire District I NH

New Jersey District III NJ

New Mexico District VIII NM

New York District II NY

North Carolina District IV NC

North Dakota District VI ND

Ohio District V OH

Oklahoma District VII OK

Oregon District VIII OR

Pennsylvania District III PA

Puerto Rico District IV PR

Rhode Island District I RI

South Carolina District IV SC

South Dakota District VI SD

Tennessee District VII TN

Texas District XI TX

Utah District VIII UT

Vermont District I VT

Virginia District IV VA

Washington District VIII WA

West Virginia District IV WV

Wisconsin District VI WI

**Appendix 2: US Census Bureau Codes Dicennial Census and the American Community Survey.**

*Dicennial Census – Demographic and Housing Characteristics File (API variables:* [*https://api.census.gov/data/2020/dec/dhc/variables.html*](https://api.census.gov/data/2020/dec/dhc/variables.html)*)*

*(*[*https://www2.census.gov/programs-surveys/decennial/2020/technical-documentation/complete-tech-docs/demographic-and-housing-characteristics-file-and-demographic-profile/2020census-demographic-and-housing-characteristics-file-and-demographic-profile-techdoc.pdf*](https://www2.census.gov/programs-surveys/decennial/2020/technical-documentation/complete-tech-docs/demographic-and-housing-characteristics-file-and-demographic-profile/2020census-demographic-and-housing-characteristics-file-and-demographic-profile-techdoc.pdf) *)*

*American Community Survey*

*(API variables:* [*https://api.census.gov/data/2022/acs/acs1/variables.html*](https://api.census.gov/data/2022/acs/acs1/variables.html) *)*

B01001\_026E Estimate \_Total \_Female

female\_10\_to\_14 = "B01001\_029",

female\_15\_to\_17 = "B01001\_030",

female\_18\_to\_19 = "B01001\_031",

female\_20years = "B01001\_032",

female\_21years = "B01001\_033",

female\_22\_to\_24 = "B01001\_034",

female\_25\_to\_29 = "B01001\_035",

female\_30\_to\_34 = "B01001\_036",

female\_35\_to\_39 = "B01001\_037",

female\_40\_to\_44 = "B01001\_038",

female\_45\_to\_49 = "B01001\_039",

female\_50\_to\_54 = "B01001\_040",

female\_55\_to\_59 = "B01001\_041",

female\_60\_to\_61 = "B01001\_042",

female\_62\_to\_64 = "B01001\_043",

female\_65\_to\_66 = "B01001\_044",

female\_67\_to\_69 = "B01001\_045",

female\_70\_to\_74 = "B01001\_046",

female\_75\_to\_79 = "B01001\_047",

female\_80\_to\_84 = "B01001\_048",

female\_85\_over = "B01001\_049"

**Appendix 3: Characteristics of Isochrones**

The bespoke R code generates individual maps for each drive time, providing a visual representation of the accessible areas on a map. The function shapefiles, which are geospatial data files used for storing geographic information, including the boundaries of the reachable areas. The HERE API (here.com) was utilized because traffic and time could be standardized for each year. Each year the isochrones are built on the third Fridays in October at 0900 defined as “posix\_time”. We imagined that patients would see their primary care provider at this time of year for an influenza vaccination or other issue. The hereR package (<https://github.com/munterfi/hereR/> ) works as a wrapper around the R code that calls the HERE REST API for isoline routing (platform.here.com) and returns it as an sf object. There is a cost of $5.50 for every 1,000 isolines created (<https://www.here.com/get-started/pricing#here---platform---pricing---page-title> ).

* October 18, 2013
* October 17, 2014
* October 16, 2015
* October 21, 2016
* October 20, 2017
* October 19, 2018
* October 18, 2019
* October 16, 2020
* October 15, 2021
* October 21, 2022

R code utilizing the hereR package with the isoline library. The range of isochrones was 30 minutes, 60 minutes, 120 minutes, and 180 minutes.

hereR::isoline(

poi = row\_data,

range = c(1),

datetime = posix\_time,

routing\_mode = "fast",

range\_type = "time",

transport\_mode = "car",

url\_only = FALSE,

optimize = "balanced",

traffic = TRUE,

aggregate = FALSE)