Healthcare Access in Quantitative Research: Operationalizing Geographic Access to Care

By

Kaelee Mae Belletto

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF ARTS

in

Sociology

August 13, 2021

Nashville, Tennessee

Approved:

Bianca N. Manago, Ph.D.

Evelyn J. Patterson, Ph.D.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure | Page

INTRODUCTION

In extant literature on health, geographic access to a healthcare facility has been used as a predictor of both health-seeking behaviors and health outcomes. Access, while a broad term, is often represented by two proximal factors: time to a healthcare facility and distance to a healthcare facility. Though they often proxy as equivalent measures of the same "healthcare access" concept, these two measures may be capturing different aspects of a person's journey to healthcare. There are varying results on how increased distance or time to a care facility impacts a person's health, with some research finding a positive and other research finding a negative relationship between proximity and health. This raises two questions: first, are time and distance equivalent measures of health outcomes? Second, if not, what are their differences?

To examine these questions, this study uses data from the 2017 Health Reform Monitoring Survey to assess (what I argue are) two distinct measures of a person's journey to healthcare. In this study, I examine the relative utility of both self-reported time and distance to a person's most frequently sought health facility to explain a self-rated health measure. This approach has two main contributions to the literature.

First, the relationship (or lack thereof) between two methodologically similar time and distance measures is assessed for their independent predictive power on health. Existing research has yet to distinguish what differences may exist between these two measures and often relies on assumptions about a person's most frequently sought practitioner or means of transportation to medical care (Ricketts and Goldsmith 2005; Zahnd and McLafferty 2017). I suggest that "time to a practitioner generally" yields a more accurate prediction of healthcare access—and thus health—because of its innate ability to capture externalities which a strict distance measure ignores (e.g., mode of transportation, reliance on others for transit, etc.).

Second, I use self-report measures of respondents' journeys to healthcare facilities. Existing literature focuses on software-generated estimates (e.g., from Geographic Information Systems (GIS) and their potential relationship to health outcomes. While providing a potentially more objective measure, it may be incapable of capturing other factors that affect individuals' journey to healthcare. Therefore, this paper investigates the utility of a unique measurement method, i.e., self-report, to reflect respondents' experiences of access to healthcare.

LITERATURE REVIEW

In health literature, there is an existing connection between healthcare access, receiving care, and health outcomes (NASEM 2018); however, the definition and measurement of healthcare access varies widely. One of the most common definitions of healthcare access is, "the timely use of personal health services to achieve the best health outcomes" (IOM 1993). Implicit in this definition is the direct link between attending some sort of healthcare facility and improving personal health as a result. Health services research often focuses on a multi-faceted definition of access that relates the social, behavioral, and physical factors—for instance, "availability, accessibility, accommodation, affordability and acceptability,"of care—to a specific patient's propensity to both seek and employ healthcare services (Higgs 2004:121).

Two of the most common factors that are used to represent physical access to a healthcare facility are time and distance to healthcare provider (Kelley et al. 2016). In terms of distance to care, earlier studies relating geographic factors to health, like McGuirk and Porrell (1984), investigate spatial access to hospitals by time and distance metrics estimated by Geographic Information Systems (GIS). Largely, these studies examine the "distance decay theory"—which assumes worsening health outcomes as distance travelled to a provider increases

(Billi et al. 2014). Across this field of literature, localized studies that focus on effects of distance produce conflicting results on the relationship between this metric and the likelihood of care utilization or health outcomes both because of methodological inconsistency in the operationalization of the measure and in its effect on health (Caniglia et al. 2019). For instance, two studies examining breast cancer treatment and progression in the Southeastern U.S. find significantly different effects of distance traveled to providers and receiving care. In their study of Florida breast carcinoma treatment, Voti et al. (2006) found that distance was negatively associated with receiving treatment. In contrast, Wheeler et al. (2014) found that rural North Carolina patients living more than 10 miles from a care facility were *more likely* to receive treatment than individuals living within a 10-mile radius. Thus, while some research finds that the closer one is to a care facility the better their health outcomes, other research finds that the closer one is to a care facility, the worse their health outcomes. This suggests that distance may not be a precise measure of healthcare access because of inconsistent methodological assumptions regarding this measure, as well as wide variability in its operationalizations (Caniglia et al. 2019).

One reason why distance may be an inaccurate measure of healthcare access is due to how it is measured. Measures of both time and distance to care are currently and most frequently generated with software estimation methods, e.g., GIS (Higgs 2004; Kelly et al. 2016). Even when estimated through GIS, time and distance capture substantially different stories of a person's travels to a healthcare facility. Distance measures are largely characterized as a straight-line (Euclidian) estimates usually from one's home to the nearest healthcare facilities or GIS-generated methods of measurement that imputes roadway data to estimate driving distance to a facility; this reduces a person's commute to a path to either one that is unrealistic because of built environment or one which is only possible through one form of transit, with significant differences in distances generated by each method for the same start and end points (Jones et al. 2010). Tabling the assumptions about where that person is travelling from or what provider they are heading to, distance measures are often a proxy for access when they actually represent whether or not there are practitioners nearby a person's home.

Similar to measures of distance, GIS estimates of time are also typically derived from travel routes that employ roadways. In this way, time estimates through GIS offer some amount of advantage to some of their strictly-distance counterparts because of their ability to include a potential realistic travel path; however, these estimations still do not account for the actual travel of people seeking care. For example, Banke-Thomas et al. (2021) compare multiple time estimation methods to actual patient travel times and find substantial underestimation of times from almost all time-based estimation methods.

Even if two studies are both using GIS estimates of distance or time, their methodological approaches to measuring distance may differ. This can bias research findings about the effect of distance on health outcomes. For example, the means of generating distance via GIS (i.e. straight-line distance from a person's home to a practitioner versus an estimate including public infrastructure like roads) can produce aggregation error in sample geographic data; this generates unreliable and unreproducible findings (Apparicio et al. 2008; Apparicio et al. 2017). Compounding the issue of estimation error, medical data often utilizes "geo-scrambling" to protect the confidentiality of the individuals whose access is being assessed. Elkies et al. (2005) explain that the intentional perturbations of household data for many widely-cited surveys generates non-negligible bias that can lead to consistent error in various forms of statistical analysis that use household address data or census data—something which is true for both time and distance estimates in GIS.

Compared to the ubiquity of GIS measures for estimating the effect of distance on health outcomes, there is relatively little health research uses or examines self-report measures of distance/time to care for predicting health outcomes. Notably, although subjective, self-report measures have been found to be robust measures of important sociological phenomena. For example, self-reported health is a robust indicator of health and highly predictive of mortality (Idler and Benyamini 1997). In contrast to relatively objective, physician-collected measures of health (e.g., blood pressure), self-rated health may capture important complexities that are otherwise missed. For instance, self-rating offers a participant to refer to their health relative to individuals occupying similar social positions. In addition, it may capture some of the complexities or existing illness and symptom management that alternate measures are not equipped to encapsulate.

The strengths of self-reports may extend to subjective measures of healthcare access. The direct impact of time or distance to a healthcare facility may vary by geographic context, methodological approach, and health outcomes. For medical sociology, time or distance to healthcare are critical measures of the socio-spatial context that a person is accessing care from. Much like the robustness of the self-rated health measure, participants' self-reports of time/distance to a frequently-sought practitioner allows data to reflect personal journeys to healthcare. By allowing an individual to respond based on their own point of departure and the facility they attend, self-reports may resolve some of the errors in assumption that GIS estimates cannot capture. Key criteria that GIS estimates usually generalize include the practice or physician a person may see, how they get to the physician, and where they are travelling from (as aforementioned, typically one's place of residence. Because deaths attributable to chronic or persistent health issues substantially outnumber death attributable to acute accidents (Kochanek, Xu, and Arias 2019), proximity to any possible healthcare provider in a given area does not provide the same substantive empirical information as a person's most frequently sought place of care. Self-reported data allow respondents to base their response to survey criteria considering all these factors and reflecting the actual details of their travel to care.

Therefore, in this project, I suggest that self-reported measures may be useful for assessing healthcare access and predictive of health outcomes. Using participant self-estimated measures of time and distance to healthcare allows for a direct comparison of the potential latent influences that would affect the report of each of these measures. While these measures do share an intrinsic relationship, i.e., the further one lives (distance) the more time it probably takes to get to a physician, I argue that the predictive ability of these two measures are distinct. Specifically, I suggest that time to a healthcare provider is more likely to encompass latent factors that a strict distance measure does not capture. For example, temporal measures of access encompass how mode of transportation may impact a person's journey to healthcare. While distance-based estimates proxy as a measure of access in much of health access literature's existing work, time-based access measures may provide information which more accurately reflects a person's travel to healthcare providers or facilities.

Using data from the 2017 Health Reform Monitoring Survey (HRMS), I examine the variation in self-reported health as a product of both self-reported distance and self-reported time to individuals' healthcare facility. Utilizing regression analyses, I examine whether and how the effect of time and distance on health is significantly different and whether either of these measures produces better model fit in these estimated models.

METHODS

Data for this project come from the 2017 Health Reform Monitoring Survey, Quarter 3 (HRMS)—the most recent wave available at the start of this project. This data set was collected using random sampling by U.S. postal address, with approximately 7,500 respondents in each wave of the semi-annual survey. The variables of interest from this survey include self-rated health, self-reported distance to medical provider, self-reported time to medical care, means of transit to medical care, and respondents' report of access to public transit from the place they live. HRMS examines a limited age range (the 18-64 year old population) which constrains this study to examining only non-elderly and non-youth populations. The effects of excluding this age cohort limits the generalizability to older populations, but still allows for a complete analysis of a range of age cohorts who may be seeking care for themselves or their dependents. The final analytic sample includes 5,392 respondents (see Missing Data, below).

Demographics

Table 1 presents demographic data for the sample. Though missing cases were removed, the distributions of demographic variables remained comparable after listwise deletion was completed. The sample was majority white (69.6%). Hispanic-identified participants comprised 15.1% of the sample, followed by those reporting Black (8.8%) and "other" (3.8%) or two or more races (2.7%). Approximately 49.8% of the sample identified as male, with the remaining 50.2% identifying as female. Average age was 45, which is slightly younger than the mean age of the U.S. population (this is an artifact of the survey's age limitations). The income of this sample was captured ordinally; the distribution of this variable was somewhat bimodal, with 19.8% of the sample reporting a household income below $25,000 annually and 35.1% reporting a household income of over $100,000 a year. This group of respondents had relatively high educational attainment, with 64.6% having at least some college education, a Bachelor's degree, or higher. The majority of respondents lived in urban or suburban areas, with 86.9% living in a census-designated Metropolitan Statistical Area. Pertinent for this study, Table 1 shows that 94.5% of the sample reported being insured at the time the survey was given.

Missing Data

The primary variables of interest were self-rated health, self-reported distance to medical provider, self-reported time to medical care, means of transit to medical care, and respondents' report of access to public transit from the place they live. Only individuals who responded that they had a practitioner or number of practitioners who they routinely sought for care were asked about the time and distance to these practitioners. If the respondent answered "no," they did not have a practitioner from whom they routinely sought care, they were not prompted to answer the time and distance questions. Approximately 23% of respondents (2,140 respondents) reported not having a place they usually went to get healthcare. Accordingly, the majority of missing data was due to this skip logic, specifically, no respondents were missing data on demographic measures,

To assess the potential effects of this skip-logic and the resulting missing data on my findings, I estimated binary logistic regression models to determine what groups were most likely to be missing data on the main independent (time and distance) and dependent (self-rated health) variables. I find that persons identifying as LGBT, who were uninsured, and who had lower income all were significantly less likely to indicate having a place to routinely seek care. This suggests that the data was not missing at random and therefore, I am unable to impute these values without creating additional bias in the dataset (Allison 2002). For this reason, cases without responses on the time, distance, transit, and health questions were removed from the data set using listwise deletion to minimize bias. After missing cases were removed, the sample used for analysis contained 5,392 observations.

Dependent Variable

The primary dependent variable is self-reported health. This variable asks respondents "In general, would you say your health is:" with options ranging from "poor" health (coded as 1) to "excellent" health (coded as 5). Self-reported health is a commonly used and powerful predictor of potential morbidity or mortality; use of self-rated health as the primary dependent variable provides an outcome consistent with the sociological health literature (Idler and Benyamini 1997). One additional strength of using a self-reported health measure is that individuals who may not have access to medical facilities to receive formal diagnosis are still able to report ill health (Idler and Benyamini 1997). There were very few respondents who reported having poor (2.65%) or fair (11.07%) health, therefore, these categories were combined for statistical power (see Table 1).

Independent Variables

Unlike other nationally-representative data sets, this dataset contains self-report measures of the primary independent variables of interest (distance and time to most frequently attended medical facility). Unfortunately, this dataset does not contain geo-coded data that would permit a GIS analysis; therefore, I am unable to compare the predictive ability of GIS measures to self-report measures. I am, however, able to examine and compare the predictive ability of time and distance in relation to health.

For time, respondents were asked "About how long does it take you to get to the place you usually go when you are sick or need advice about your health?" Answer choices ranged from <15 minutes to more than one hour, in five evenly distributed, 15-minute intervals. Very few respondents lived 45 minutes to 1 hour away (4.66%) or more than one hour away (2.91%), so these categories were collapsed to 45 or more minutes away, resulting in 4 total categories: <15 min, 15-30, 30-45, and 45 or more minutes (see Table 1).

For distance, respondents were asked "About how far do you travel to get to the place you usually go when you are sick or need advice about your health?" Respondents could choose one of six responses for how far a person travelled to their usual healthcare facility: less than 1/2 a mile, 1/2 a mile to under one mile, one mile to under two miles to under five miles, five miles to under 10 miles, and 10 or more miles. The lowest two categories (less than ½ a mile and ½ a mile to under one mile) were collapsed into one category titled "under one mile" to both generate more even distribution across analytic categories and ultimately provide more comparable model complexity across time and distance models (see Table 1). For both time and distance, odds were calculated relative to the lowest category (i.e., closest distance or smallest amount of time to care).

Control Variables

To more precisely measure the effects of time and distance on self-reported health, I included theoretically relevant control variables in the models. These variables include: gender, race, age, and educational attainment (Read and Gorman 2010; Read and Emerson 2005; Ross and Wu 1996). Due to their association with improved chances of accessing a practitioner, I included the following control variables: household income, educational attainment, Metropolitan Statistical Area (MSA) residence, and health insurance status (Link and Phelan 1995). Variables were kept in their original form (e.g., continuous or ordinal), unless otherwise specified. Odds were computed relative to white respondents (race), lowest educational attainment category (less than high school), lowest income category (less than $25,000 a year), and respondents who were not insured or not living in MSAs.

Finally, I included a control for mode of transportation. Because a large portion of the sample reported driving themselves to care (83.2%) versus using other modes of transit (16.8%), I dichotomized the variable for those who drive themselves compared to others (e.g. ride-sharing, ambulance or medical vehicle, walking, being driven by someone else, or public transit).

ANALYTIC STRATEGY

To test the hypotheses, I estimated three ordinal logistic regression models and use post-estimation to generate predicted probabilities (Long and Freese 2014). To determine if ordinal logistic regression models were appropriate, I estimated Brant tests of the parallel regression assumption. As is common with Brant tests (Long and Freese 2014), the ordinal logit model violates the parallel regression assumption ($p<0.001$). Specifically, mode of transportation, age, education, income, and insurance status all have relationships with self-rated health that may not be strictly ordered in nature ($p<.05$). However, neither of the two main variables of interest (i.e. time and distance) violate the parallel regression assumption, which suggests that ordinal logit may be appropriate to analyze these effects. When looking at model fit statistics, the AIC (which does not penalize for model complexity) prefers the multinomial logit model ($\Delta = 71.6$), but the BIC more strongly prefers the ordinal logit model (Δ=-231.6). Finally, to further examine the modeling strategy, I compared the findings of the multinomial and ordinal logit models (Long and Freese 2014), and find no meaningful differences in effect size, direction, nor significance. Based on these factors, I use ordinal logit for my models.

Model 1 contained only distance, Model 2 contained only time, and Model 3 contained both time and distance. I use Model 3 to compare the effect sizes between time and distance. Effect size is examined by comparing the average marginal effects (AMEs) of the range of distance and time (closest to furthest proximity) on the probability of reporting each of the five health outcomes (i.e., difference of differences tests) (Long and Freese 2014). I also examine the predictive ability of the models using goodness of fit statistics. The effect size and fit statistics are jointly assessed to determine whether time or distance is preferred as a predictor for health, as well as to assess the effect of the transportation covariates within this set of models.

RESULTS

Overall Covariance

Since time and distance typically proxy for the same latent "geographic access" concept, I first examine their correlation with one another. Distance traveled and time spent travelling to one's most frequently sought healthcare provider are moderately correlated (Spearman's correlation coefficient=0.5433, $p<.001$). Although this level of correlation is relatively high for social science research, they are not perfectly correlated. This offers some support that these variables may offer somewhat different and distinct measurements of access, i.e., they may not be measuring the same latent concept. This discrepancy supports further investigation of the underlying access factor for each of these variables.

Effect of Time and Distance on Self-Rated Health

I first examine the effects of time and distance using the combined model. Overall, and consistent with most existing literature, both distance and time have a negative relationship with self-rated health. As either measure increased, the likelihood of positive health outcomes decreased. As shown in Figures 1 and 2, at the smallest distance and time intervals (i.e. less than one mile to care and less than 15 minutes from care), respondents are most likely to report having ‘good’ and ‘very good’ health. As distance and time to practitioner increases, the likelihood of ‘very good’ and ‘excellent’ health decreases. The probabilities of ‘good’ and ‘poor/fair’ health increased with greater travel time and distance to care.

First Differences

First, I examine the change in predicted probability of each category of self-rated health over the range of time and distance (see Table 3). The probability of ‘poor/fair’ health increased as a function of both time (.031, $p<0.05$) and distance (.028, $p<0.05$). Similarly, the probability of good health increased as a function of both time (.028, $p<0.05$) and distance (.029, $p<0.05$). In contrast, but keeping with overall trends, the probability of very good and excellent health decreased with both time ($\Delta_{\text{very good}} = -0.033$, $\Delta_{\text{excellent}} = -0.026$, $p<0.05$ both) and distance ($\Delta_{\text{very good}} = -0.030$, $\Delta_{\text{excellent}} = -0.027$, $p<0.05$ both).

Difference of Differences

Next, I compare the size of the effects for time and distance on each category of self-rated health over the range of time and distance (see Table 4). Across all outcomes of self-rated health, the effects of time and distance were not significantly different from one another ($p>0.80$, all comparisons). This lack of difference suggests that the effect of time and distance on perceived health are comparable.

Model Fit

Finally, I compare the overall model fit to determine if time and distance provide distinct enough information to warrant the inclusion of both in a model, or if one of these measures is sufficient (see Table 2). When looking at both the AIC and BIC, Model 2 (distance only) is preferred over all other models (Raftery 1995). In contrast to my hypothesis that time was a superior measure of access than distance, this finding suggests that the distance measure may be a strong and sufficient predictor of self-rated health and that when added to a model time does not improve predictive ability. This conclusion is further supported by the results of a likelihood ratio test which shows that time does not significantly improve model fit (X^2=6.22, df=3, p=0.102).

Overall, although there are some discrepancies in the measures, time and distance appear to be comparable measures of perceived health. Additionally, and contrary to my hypothesis, the distance measure appears to provide a better fit to and reflection of the data for this sample than time. Furthermore, fit statistics suggest that distance provides sufficient information for predicting health and that time did not improve the model fit for predicting self-rated health.

CONCLUSION

The utility of geospatial access measures in health research provides important insight into how people realize medical services. While time and distance estimates to healthcare practitioners has often represented the same access concept, my results suggest that time and distance measures capture different information reflecting respondents' journeys to medical care. Both relative fit and effect size varied between time versus distance models based on these self-report data. While the most preferred model based on fit to data alone was that which included solely a distance measure, the effect of time on health provided comparable estimations of each self-rated health category. This finding suggests that time and distance may be comparable measures of healthcare access and similarly able to predict self-reported health. While distance has been an oft-preferred measure for healthcare access (Kelly et al. 2016), time may be able to capture latent factors of access that can illustrate substantively individuals' journeys to healthcare.

Although I did not make initial predictions about the role of modes of transportation, a transit covariate substantially improved model fit across models. Individuals who did not drive themselves to a physician consistently reported lower self-rated health. Notably, this relationship may be attributable to other access factors (disability or chronic illness, severity of illness, etc.) which could influence a person's ability to transport themself. This relationship suggests that mode of transportation and transit access are important considerations when including geographic access factors in quantitative work. This finding is an important consideration for future literature assessing geographic access factors. While not consistently included in health access literature, this study's findings support further exploration into the utility of transportation covariates in health access and outcomes literature.

While recognizing the contributions of this research, there were also several limitations imposed by the data. First, the data is limited to individuals 18 to 64. These age constraints limit the generalizability of these findings to elderly populations seeking healthcare. Second, the lack of data on individuals who did not report having routine healthcare added additional constraints to who the data within the sample represented. Income, insurance status, and LGBT identity were all associated with higher odds of not reporting having a routine place to go for care. Future research should address the specific healthcare access needs of these communities and, consequently, how each of the geographic and transit measures model health for these groups. In addition, the single method of data collection via self-report does not allow these results to be directly compared to spatial analytic data like GIS estimates for this same group. Continued research on healthcare access should consider direct comparison of multi-method geographic access measures. In addition, future research should consider a multi-variable approach for assessing health; variation in self-rated health, physician-appraised health, and mental health over time and distance to practitioner may exist, which the present study is not able to explicate.

Consistent access to healthcare can dramatically impact a person's health; accordingly, the ways that contemporary literature operationalizes this broad access concept has a direct effect on how well social science researchers are able to capture the inequality caused by differential access. This paper predicted that for a variety of reasons, time to healthcare may be a superior measure of healthcare access than distance to care. The findings suggest that time and distance to care are comparable—but not equivalent—operationalizations of healthcare access. It is clear that these two measures are both valuable to consider in the healthcare access puzzle. As this study demonstrates, though, the way social science research chooses to operationalize access differentially impacts estimates of a person's journey to care. Time, distance, and transportation all provide valuable insights about health as a result of access to care. The results of this paper offer that—while time and distance host different benefits to modeling access—the relationship between proximal factors and health is complex. As access variables are integrated into medical sociology research, researchers' continued assessment of the measures of healthcare access is critical to the validity and accuracy of quantitative health literature.

Bibliography

Allison, Paul D. 2002. *Missing Data*. Thousand Oaks, CA: SAGE Publications.

Andersen, Ronald M. 1995. "Revisiting the Behavioral Model and Access to Medical Care: Does It Matter?" *Journal of Health and Social Behavior* 36(1):1–10.

Apparicio, Philippe, Mohamed Abdelmajid, Mylène Riva, and Richard Shearmur. 2008. "Comparing Alternative Approaches to Measuring the Geographical Accessibility of Urban Health Services: Distance Types and Aggregation-Error Issues." *International Journal of Health Geographics* 7(1):7.

Apparicio, Philippe, Jérémy Gelb, Anne-Sophie Dubé, Simon Kingham, Lise Gauvin, and Éric Robitaille. 2017. "The Approaches to Measuring the Potential Spatial Access to Urban Health Services Revisited: Distance Types and Aggregation-Error Issues." *International Journal of Health Geographics* 16(1):32. doi: 10.1186/s12942-017-0105-9.

Banke-Thomas, Aduragbemi, Kerry L. M. Wong, Francis Ifeanyi Ayomoh, Rokibat Olabisi Giwa-Ayedun, and Lenka Benova. 2021. "'In Cities, It's Not Far, but It Takes Long': Comparing Estimated and Replicated Travel Times to Reach Life-Saving Obstetric Care in Lagos, Nigeria." *BMJ Global Health* 6(1):e004318. doi: 10.1136/bmjgh-2020-004318.

Billi, John E., Chih-Wen Pai, and David A. Spahlinger. 2007. "The Effect of Distance to Primary Care Physician on Health Care Utilization and Disease Burden." *Health Care Management Review* 32(1):22.

Caniglia, Ellen C., Rebecca Zash, Sonja A. Swanson, Kathleen E. Wirth, Modiegi Diseko, Gloria Mayondi, Shahin Lockman, Mompati Mmalane, Joseph Makhema, Scott Dryden-Peterson, Kalé Z. Kponee-Shovein, Oaitse John, Eleanor J. Murray, and Roger L. Shapiro. 2019. "Methodological Challenges When Studying Distance to Care as an

Exposure in Health Research." *American Journal of Epidemiology* 188(9):1674–81. doi: 10.1093/aje/kwz121.

Elkies, Noam, Günther Fink, and Till Bärnighausen. 2015. "'Scrambling' Geo-Referenced Data to Protect Privacy Induces Bias in Distance Estimation." *Population and Environment* 37(1):83–98.

Hawthorne, Timothy L., and Mei-Po Kwan. 2012. "Using GIS and Perceived Distance to Understand the Unequal Geographies of Healthcare in Lower-Income Urban Neighbourhoods." *The Geographical Journal* 178(1):18–30.

Higgs, Gary. 2004. "A Literature Review of the Use of GIS-Based Measures of Access to Health Care Services." *Health Services and Outcomes Research Methodology* 5(2):119–39.

Idler, Ellen L., and Yael Benyamini. 1997. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior* 38(1):21–37.

Institute of Medicine (US) Committee on Monitoring Access to Personal Health Care Services. 1993. *Access to Health Care in America*. edited by M. Millman. Washington (DC): National Academies Press (US).

Jones, Stephen G., Avery J. Ashby, Soyal R. Momin, and Allen Naidoo. 2010. "Spatial Implications Associated with Using Euclidean Distance Measurements and Geographic Centroid Imputation in Health Care Research." *Health Services Research* 45(1):316–27. doi: 10.1111/j.1475-6773.2009.01044.x.

Kelly, Charlotte, Claire Hulme, Tracey Farragher, and Graham Clarke. 2016. "Are Differences in Travel Time or Distance to Healthcare for Adults in Global North Countries Associated with an Impact on Health Outcomes? A Systematic Review." *BMJ Open* 6(11):e013059.

Kochanek, Kenneth, Jiaquan Xu, and Elizabeth Arias. 2019. “Mortality in the United States, 2019.” *Centers for Disease Control National Center for Health Statistics*.

Link, Bruce G., and Jo Phelan. 1995. “Social Conditions As Fundamental Causes of Disease.” *Journal of Health and Social Behavior* 35:80–94.

Long, J. Scott & Jeremy Freese. 2014. Regression Models for Categorical Dependent Variables Using Stata, 3rd Edition. College Station, TX: *Stata Press*.

Mao, Liang, and Dawn Nekorchuk. 2013. “Measuring Spatial Accessibility to Healthcare for Populations with Multiple Transportation Modes.” *Health & Place* 24C:115–22.

McGuirk, Marjorie A., and Frank W. Porell. 1984. “Spatial Patterns of Hospital Utilization: The Impact of Distance and Time.” *Inquiry* 21(1):84–95.

National Academies of Science, Engineering, and Medicine Committee on Health Care Utilization and Adults with Disabilities. 2018. *Factors That Affect Health-Care Utilization*. Washington, D.C.: National Academies Press (US).

Raftery, Adrian E. 1995. “Bayesian Model Selection in Social Research.” *Sociological Methodology* 25:111–63.

Read, Jen’nan Ghazal, and Michael O. Emerson. 2005. “Racial Context, Black Immigration and the U.S. Black/White Health Disparity.” *Social Forces* 84(1):181–99.

Read, Jen’nan Ghazal, and Bridget K. Gorman. 2010. “Gender and Health Inequality.” *Annual Review of Sociology* 36(1):371–86. doi: 10.1146/annurev.soc.012809.102535.

Ricketts, Thomas C., and Laurie J. Goldsmith. 2005. “Access in Health Services Research: The Battle of the Frameworks.” *Nursing Outlook* 53(6):274–80. doi: 10.1016/j.outlook.2005.06.007.

Ross, Catherine E., and Chia-Ling Wu. 1996. “Education, Age, and the Cumulative Advantage in Health.” *Journal of Health and Social Behavior* 37(1):104–20. doi: 10.2307/2137234.

Semenza, Daniel C., Deena A. Isom Scott, Jessica M. Grosholz, and Dylan B. Jackson. 2020. “Disentangling the Health-Crime Relationship among Adults: The Role of Healthcare Access and Health Behaviors.” *Social Science & Medicine* 247:112800.

Talen, Emily. 2003. “Neighborhoods as Service Providers: A Methodology for Evaluating Pedestrian Access.” *Environment and Planning B: Planning and Design* 30(2):181–200.

Turner, Rachel A., Lucy Szaboova, and Gwynedd Williams. 2018. “Constraints to Healthcare Access among Commercial Fishers.” *Social Science & Medicine; Oxford* 216:10.

Voti, Lydia, Lisa C. Richardson, Isildinha M. Reis, Lora E. Fleming, Jill MacKinnon, and Jan Willem W. Coebergh. 2006. “Treatment of Local Breast Carcinoma in Florida.” *Cancer* 106(1):201–7.

Ward, Carol J., Michael R. Cope, and Jorden Jackson. 2020. “Healthcare Access among Older Rural Women Veterans in Utah*.” *Rural Sociology* 85(4):966–90.

Watanabe, Ryo, and Hideki Hashimoto. 2012. “Horizontal Inequity in Healthcare Access under the Universal Coverage in Japan; 1986-2007.” *Social Science & Medicine* 75(8):1372–78.

Wheeler, Stephanie B., Tzy-Mey Kuo, Danielle Durham, Brian Frizzelle, Katherine Reeder-Hayes, and Anne-Marie Meyer. 2014. “Effects of Distance to Care and Rural or Urban Residence on Receipt of Radiation Therapy Among North Carolina Medicare Enrollees With Breast Cancer.” *North Carolina Medical Journal* 75(4):239–46.

Zhand, Whitney E., and Sara L. McLafferty. 2017. “Contextual Effects and Cancer Outcomes in the United States: A Systematic Review of Characteristics in Multilevel Analyses - ClinicalKey.” *Annals of Epidemiology* 27(11):739–48.

Table 1: Descriptive Statistics[1] (N = 5392)

	Mean/Prop.	SD	Min.	Max.	Median
Race/Ethnicity					
White	69.6%				
Black	8.8%				
Hispanic	15.1%				
2+ Races	2.7%				
Other	3.8%				
Age	45.26	12.97	18.00	64.00	47.00
Gender					
Female	50.2%				
Male	49.8%				
Education, less categories					
Less than HS	8.0%				
HS	27.4%				
Some college	32.0%				
Bachelors or higher	32.6%				
Household Income					
<$25k	19.8%				
$25k-<$50k	17.2%				
$50k-<$75k	15.6%				
$75k-<$100k	12.4%				
$100k+	35.1%				
MSA residence[2]	86.9%				
Insured at time of survey	94.5%				
Drives self to care	83.2%				
Distance to Med. Care					
<1mi	12.0%				
1mi-<2mi	12.0%				
2mi-<5mi	26.0%				
5mi-<10mi	28.0%				
10mi+	22.0%				
Time to Med. Care					
<15 min	38.0%				
15-30 min	41.0%				
30-45 min	13.0%				
45+	8.0%				
Self-Rated Health					
Poor/Fair	14.0%				
Good	35.0%				
Very Good	39.0%				
Excellent	12.0%				

[1] Data from 2017 Health Reform Monitoring Survey, QIII

[2] MSA indicates census-designated residence in a Metropolitan Statistical Area

Table 2. Ordinal Logistic Regression, Self-Rated Health - Odds Ratios and (SEs)[1] – N=5392

	(1)		(2)		(3)	
	Model 1		Model 2		Model 3	
Female[2]	1.002	(0.05)	1.006	(0.11)	1.004	(0.07)
Race[3]						
Black	0.925	(-0.84)	0.960	(-0.44)	0.972	(-0.31)
Hispanic	1.090	(1.10)	1.083	(1.00)	1.093	(1.12)
2+ Races	0.880	(-0.80)	0.905	(-0.63)	0.911	(-0.58)
Other	0.911	(-0.70)	0.921	(-0.62)	0.937	(-0.49)
Age	0.976^{***}	(-11.66)	0.974^{***}	(-12.24)	0.975^{***}	(-12.14)
Education[4]						
HS	1.728^{***}	(4.97)	1.691^{***}	(4.75)	1.692^{***}	(4.76)
Some college	2.011^{***}	(6.22)	1.935^{***}	(5.84)	1.941^{***}	(5.86)
Bachelors+	3.695^{***}	(10.76)	3.539^{***}	(10.36)	3.560^{***}	(10.40)
MSA	0.975	(-0.33)	0.978	(-0.28)	0.988	(-0.15)
Income[5]						
$25k-<$50k	1.699^{***}	(6.15)	1.600^{***}	(5.40)	1.580^{***}	(5.24)
$50k-<$75k	2.104^{***}	(8.27)	1.958^{***}	(7.38)	1.925^{***}	(7.17)
$75k-<$100k	2.479^{***}	(9.33)	2.245^{***}	(8.15)	2.209^{***}	(7.97)
$100k+	3.322^{***}	(13.50)	3.074^{***}	(12.43)	3.022^{***}	(12.20)
Insured	0.857	(-1.35)	0.852	(-1.39)	0.855	(-1.36)
Time to care[6]						
15-<30 min			0.856^{**}	(-2.71)	0.937	(-1.02)
30-<45 min			0.734^{***}	(-3.81)	0.841	(-1.83)
45+ min			0.620^{***}	(-4.64)	0.764^{*}	(-2.26)
Distance to care[7]						
1mi-<2mi			0.909	(-0.92)	0.932	(-0.67)
2mi-<5mi			0.811^{*}	(-2.30)	0.837	(-1.92)
5mi-<10mi			0.720^{***}	(-3.63)	0.765^{**}	(-2.75)
10mi+			0.669^{***}	(-4.28)	0.767^{*}	(-2.38)
Doesn't drive self to care			0.524^{***}	(-8.51)	0.547^{***}	(-7.76)
aic		12711		12763		12711
bic		12863		12902		12882

Exponentiated coefficients; *t* statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
[1] Data from 2017 Health Reform Monitoring Survey, QIII
[2] Omitted group: 'male'
[3] Omitted group: 'white'
[4] Omitted group: 'less than high school'
[5] Omitted group: 'less than $25k"
[6] Omitted group: 'less than 15 min.'
[7] Omitted group: 'less than 1 mi.'

Table 3: First Differences of Self-Rated Health Across Range of IVs [1] (N = 5392)

	Pred. Prob.	p-value	ll	ul
Poor/Fair				
Distance	0.028	0.016	0.005	0.051
Time	0.031	0.031	0.003	0.058
Good				
Distance	0.029	0.018	0.005	0.053
Time	0.028	0.018	0.005	0.051
Very Good				
Distance	-0.03	0.015	-0.054	-0.006
Time	-0.033	0.031	-0.063	-0.003
Excellent				
Distance	-0.027	0.020	-0.05	-0.004
Time	-0.026	0.017	-0.046	-0.005

[1]Data from 2017 Health Reform Monitoring Survey, QIII; estimates generated from Model 3

Table 4: Second Differences of Time and Distance MEs Across Health Outcomes[1] (N = 5392)

	Pred. Prob.	p-value	ll	ul
Poor/Fair	-0.003	0.907	-0.045	0.04
Good	0.001	0.954	-0.039	0.041
Very Good	0.003	0.89	-0.043	0.049
Excellent	-0.002	0.922	-0.039	0.035

[1]Data from 2017 Health Reform Monitoring Survey, QIII; estimates generated from Model 3

FIGURE 1: PREDICTED PROBABILITY OF SELF-RATED HEALTH BY DISTANCE

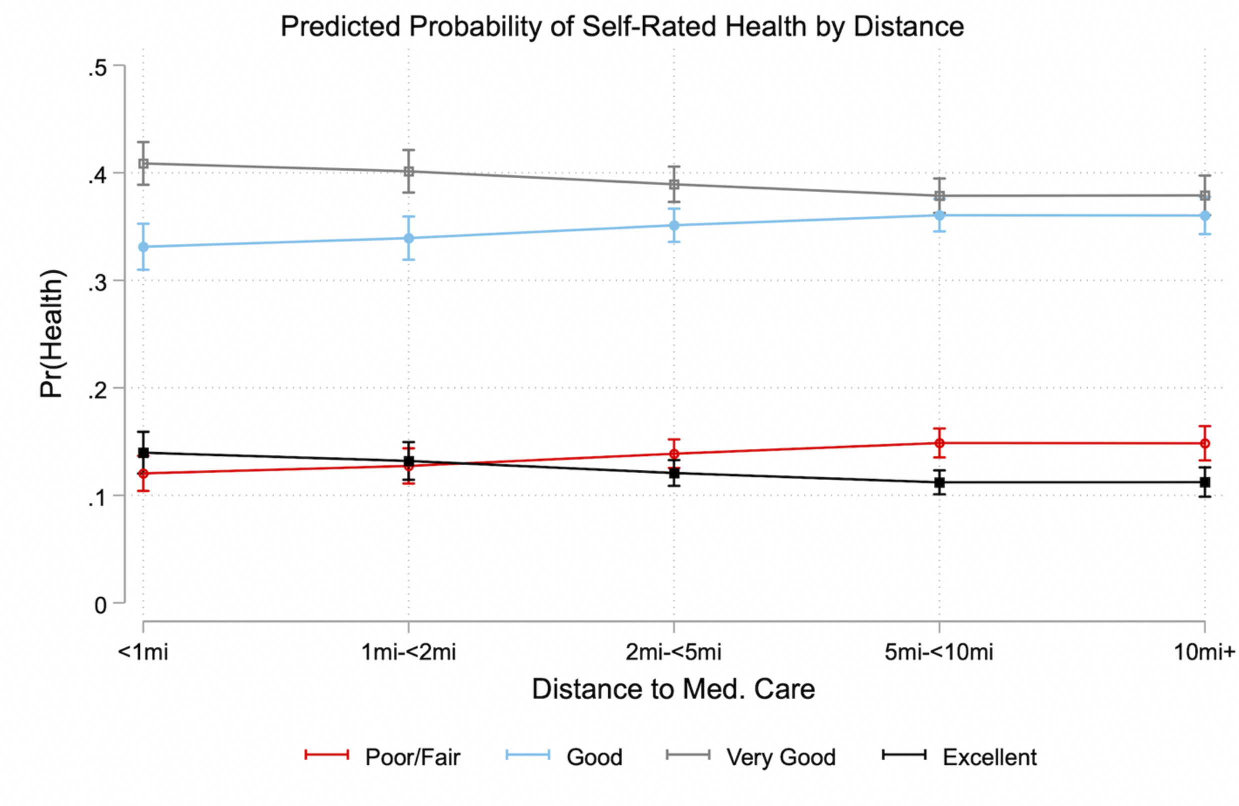

FIGURE 2: PREDICTED PROBABILITY OF SELF-RATED HEALTH BY TIME