2015

Planning Towards Equal Spatial Accessibility of NCI Cancer Centers Across Geographic Areas and Demographic Groups in the U.S.

Cong Fu
Louisiana State University and Agricultural and Mechanical College, fucong87@gmail.com

PLANNING TOWARDS EQUAL SPATIAL ACCESSIBILITY OF NCI CANCER
CENTERS ACROSS GEOGRAPHIC AREAS AND DEMOGRAPHIC GROUPS IN
THE U.S.

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Geography and Anthropology

by
Cong Fu
B.E., China University of Geosciences, 2008
M.S., Wuhan University, 2010
August 2015

Table of Contents

List of Tables

List of Figures

Abstract

The Cancer Centers designated by the National Cancer Institute (NCI) form the "backbone" of the cancer care system in the United States. Awarded via a peer-review process and being re-evaluated every 3 to 5 years, an NCI Cancer Center receives substantial financial support from NCI grants. When the quality standard is not compromised, we argue that an additional criterion for improving and promoting equal accessibility should be factored into the designation and planning process of NCI Cancer Centers. With the help of regression and dummy variables, this research evaluates geographic disparities in spatial accessibility of the NCI Cancer Centers across geographic area, divisions and urbanicity. It also evaluates demographic disparities across ethnic and poverty groups. Then this research examines two planning objectives to minimize the inequalities in accessibility. One is to minimize the geographic inequality while the other is to minimize the racial disparities. Two types of optimization scenarios are considered in this exploratory research for the objective of minimizing inequality of spatial accessibility. One is to allocate additional resources to existing NCI Cancer Centers, and the other is to designate new centers from the most likely candidates (e.g., existing academic medical centers or AMCs). Quadratic Programming (QP) and Particle Swarm Optimization (PSO) are used to solve different optimization problems. Several scenarios are used to illustrate the impact of optimization on reducing geographic and demographic disparities. Results from the study may inform the public policy decision making process in planning of the NCI Cancer Centers towards equal accessibility.

Chapter 1 Introduction

Equality in health care is widely considered as an important goal of public policy. Among a diverse set of principles of equity, equal access to health care (for those in equal need) is considered the most appropriate principle for health care policy makers to pursue. Inequality in health care access comes at a personal and societal price, evidenced in disparities in various health outcomes. Outcomes include differential rates in infant mortality and birth weight, in vaccination, in complications from preventive and common diseases, in late-stage cancer diagnosis, in quality patient care and survival, among others.

The United States has one of the highest age-standardized rates for all cancers (excluding non-melanoma skin cancer) for men and women in the world (World Health Organization, 2014). Over 340 men and 290 women for every 100,000 people suffer from cancer. This makes cancer a leading cause of death in the United States, second only to heart disease (Jamal et al., 2010). The spatial access to cancer care heavily influences patients' usage of the cancer care services and the outcomes (Onega, 2008). The spatial distribution of cancer care services is not uniform, and calls for effective planning and allocation to match demands. Policies and strategies are needed to reduce inequality in cancer care accessibility so that the gaps in accessibility can be reduced.

The National Cancer Institute (NCI) in the U.S. has designated dozens of Cancer Centers, hereafter referred to as the NCI Cancer Centers (or simply NCI-CCs). These cancer centers are considered the backbone of the U. S. cancer care system because of their "scientific excellence and the capability to integrate a diversity of research approaches to focus on the problem of cancer" (National Cancer Institute, 2013). The NCI has built the NCI Thesaurus, which provides an up-to-date and comprehensive science-based cancer

terminology (Angela, 2010) and reflects current best practice (Coronado, 2004; Jennifer, 2006). The study of NCI-CCs is selected to demonstrate how analytical spatial analysis of accessibility can improve our understanding of disparity in health care service allocation and possible strategies to mitigate it.

Accessibility may be related to spatial and nonspatial factors. *Spatial access* emphasizes the importance of spatial separation between supply and demand as a barrier or a facilitator, whereas the *aspatial access* is related to many demographic and socioeconomic variables that characterize various demographic groups. Given our primary interest in geographic issues, this study will emphasize spatial accessibility. However, the study will also analyze how some asptial factors interact with spatial accessibility. For example, residents somewhere may suffer from poor access because (1) they are far from any NCI Cancer Centers, or (2) they are composed of disproportionally higher ratios of disadvantaged population groups (e.g., low-income and minority residents) that lack economic or transportation means to gain the access. The former leads to *spatial (geographic) disparity*, and the latter leads to *aspatial disparity*.

Awarded via a grant using a peer-review process every three to five years, the current designation criterion for a NCI Cancer Center focuses on quality of research and care in cancer prevention, diagnosis, and treatment. When the quality standard is not compromised, we argue that an additional criterion for improving and promoting equal accessibility should be factored into the designation and planning process of NCI Cancer Centers. The goal of this research is to build the methodological foundation for identifying scientific formulation of feasible policy scenarios that reduce disparities in accessibility of NCI Cancer Centers. Both spatial and aspatial disparities will be examined, and

corresponding planning problems for minimizing the disparities will be formulated and solved.

Specifically, this dissertation has four research tasks:

(1)   It quantifies how the spatial accessibility of NCI Cancer Centers varies across geographic areas (i.e., counties and census tracts) in the U.S.

(2)   It examines which demographic groups (e.g., racial-ethnic groups) suffer from poorer access and whether such a (racial) disparity is statistically significant.

(3)   It formulates and solves the problem of minimizing geographic disparity in spatial accessibility of NCI Caner Centers under various plausible policy scenarios (e.g., allocation a fixed amount of financial resource or designation a given number of new Centers).

(4)   It formulates and solves the problem of minimizing aspatial disparity given similar policy scenarios.

Tasks 1 and 2 measure the geographic and racial disparities, respectively; and Tasks 3 and 4 seek optimization that reduces the corresponding disparities. Each task forms a major chapter (Chapters 4-7) after the literature review (Chapter 2) and discussion of the data preparation (Chapter 3). The dissertation concludes with a summary and an outlook for future work (Chapter 9).

The dissertation belongs to a large body of literature in location-allocation analysis. As detailed in Chapter 2, one major difference of our work from the existing literature is the perspective of formulating equality in terms of *equal accessibility*. It builds upon the advancement of GIS-based spatial accessibility measures (e.g., Luo and Wang, 2003) and

another more recent advancement in location-allocation analysis by proposing a new optimization objective of *minimal disparity in accessibility* (e.g., Wang and Tang, 2013).

With its focus on methodological issues, the contributions of this study can be outlined in three aspects:

(1)   It is the first attempt to formulate the optimization problem of minimal racial disparity (or in general, minimal disparity across demographic groups) as the existing work is limited to the problem of minimal geographic disparity.

(2)   It formulates an alternative measure of minimal disparity (i.e., minimum of maximum absolute error or MINIMAX) in addition to the existing work that minimizes the sum of the absolute deviations (MAD).

(3)   It experiments with various algorithms in solving the optimization problems such as Particle Swarm Optimization (PSO) in addition to what has been used in the existing work (i.e., quadratic programming and integer programming), and recommends the best practice based on their performances and other criteria.

Chapter 2 Literature Review

Given the methodological focus of this study, the literature on methods of measuring accessibility and optimization techniques related to location-allocation analysis is then reviewed in great details in two separate sections. Between the two major sections, the middle section provides a brief survey of research on the substantive issue, i.e., accessibility of NCI Cancer Centers.

2.1 Measures of Spatial Accessibility

Accessibility refers to the relative ease by which an activity, here health care, can be reached from a given location (Kwan, 1998, 1999). *Spatial accessibility* emphasizes how easily residents (demand) at a location can overcome the spatial separation to obtain the service (supply) elsewhere.

As noted by Geurs and van Wee (Geurs, 2004), accessibility is "the extent to which land-use and transport systems enable individuals to reach activities or destinations by means of transport mode." Early measures of spatial accessibility emphasize the proximity to supply locations in terms of distance or time. For instance, one may use minimum travel time to the closest cancer care facility to measure accessibility to cancer care service. More generally, accessibility can be measured as travel cost including both travel time and waiting time (Shavandi, 2006). When accounting for the contribution of other service locations beyond the nearest one to the overall accessibility at a demand location, Hansen (1959) used a simple gravity-based potential model to define accessibility as the ease of reaching desirable destinations. In other words, the accessibility is the sum of all surrounding supply capacities, each of which is discounted by its distance (time) from the demand. Hansen's work showed one of the first efforts by the policy planners to consider

the effects of the capacity of a supply and its proximity as well as the benefit of availability of multiple supplies.

The above definitions do not consider that resources such as health care are scarce, and an adequate measure of accessibility also needs to account for the competition for services by the amount of demand involved. A simple method along this line of work measures spatial accessibility by the supply-demand match ratio in an area (usually an administrative unit such as township or county). For example, the DHHS (2008:11236) uses a minimum population-to-physician ratio of 3000:1 within a "rational service area" as a basic indicator for defining Health Professional Shortage Areas (HPSAs). However, this neither reveals the detailed spatial variations within an area unit (e.g., a county or a sub-county area), nor accounts for interaction between population and physicians across areas.

The gravity-based accessibility index, proposed originally by Weibull (1976), is a significant improvement over the classic Hansen (1959) model and considers the competition for supply by demand. The model is written as:

$$A_i = \sum_{j=1}^{n} \frac{S_j d_{ij}^{-\beta}}{V_j}, \text{ where } V_j = \sum_{k=1}^{m} D_k d_{kj}^{-\beta}. \tag{1}$$

In equation (1), A_i is the accessibility at demand location i, S_j is the capacity of supply at location j, D_k is the demand (e.g., population) at location k, d_{ij} (d_{kj}) is the distance or travel time between i (k) and j, β is the travel friction coefficient, and n and m are the total numbers of supply and demand locations, respectively. The contribution of the supply at each location j is first discounted by the rule of distance decay, further discounted by the crowdedness of all its surrounding population (at locations k) captured by the potential V,

and finally aggregated over all supply locations j. A larger A_i indicates better accessibility to services at the demand location i.

The popular two-step floating catchment area (2SFCA) method was developed by Luo and Wang (2003) to measure spatial accessibility to primary care physicians. In essence, the 2SFCA method measures spatial accessibility as a ratio of supply to demand (e.g., population), implemented by two steps:

(1)   It first assesses "service provider's availability" at the provider's location as the ratio of its supply capacity to its surrounding demand (e.g., within a threshold distance or travel time from the provider).

(2)   It sums up the availability of providers (derived in the first step) around each demand location (i.e., within the same threshold distance or travel time) to yield the accessibility value.

Other models have been proposed to improve the 2SFCA by assuming different distance decay behaviors, and most studies are from health care studies (e.g., Guagliardo 2004; Dai, 2010; Shi et al. 2012). A common fundamental assumption underlines these models: the number of nearby facilities (and their capacities) and distance to a facility can compensate each other. In other words, availability and accessibility are mutually compensating. The revelation of the inherent relationships between the methods under different names leads to the understanding that their differences are technical rather than conceptual.

After all, Wang (2012) reviewed many existing accessibility methods that count for interactions between supply and demand located in different areas, and synthesized them into a generic model such as:

$$A_i = \sum_{j=1}^{n} [S_j f(d_{ij}) / (\sum_{k=1}^{m} D_k f(d_{kj}))] \tag{2}$$

where D_k is the estimated number of patients (demand) at location (census tract) k, S_j is the capacity of supply (e.g., number of hospital beds) at location (cancer center) j, d is the travel time between them, and n and m are the total numbers of hospital locations and population locations, respectively.

The above model is termed *generalized 2SFCA* (G2SFCA) method. As the model shows, the main difference of various accessibility models is the function $f(d_{ij})$ that captures the distance decay effect. For example, in the gravity-based accessibility model in Equation (1), the function $f(d_{ij})$ is a gravity kernel (i.e., a power function of distance).

In addition to spatial accessibility, health care access is also influenced by nonspatial factors that include demographic (e.g., seniors, children, women of child-bearing ages), socioeconomic status (e.g., poverty, female-headed households, home ownership and median income), housing conditions (e.g., crowdedness, basic amenities), linguistic barriers and education, etc. For example, disadvantaged population groups (e.g., low-income and minority residents) often suffer from poor access to certain activities or opportunities because of their lack of economic or transportation means that limit their residential choice. Given the focus of our research on spatial access, our interest in nonspatial factors is on how these variables interact with spatial accessibility. In addition, for lack of available socio-demographic variables at the individual level, our data on nonspatial factors are extracted from the Census and thus ecological in nature. In other words, nonspatial factors such as racial-ethnic groups and poverty status are aggregated data in area units such as county and census tracts, etc.

8

2.2 Spatial Access to NCI Cancer Centers

Cancer is a leading cause of death in the United States, only second to heart disease (Centers for Disease Control and Prevention, 2010). Spatial access to cancer care can be particularly important to patients' utilization of the services (Onega et al., 2008), and thus the outcomes. Research suggests that longer travel time to cancer care services increases risk of advanced cancer (e.g., Gumpertz et al., 2006), reduces utilization of certain therapy (e.g., Celaya et al., 2006), and limits enrollment in clinical trials (e.g., Avis et al., 2006).

The Cancer Centers designated by the National Cancer Institute (NCI) in the U.S. (hereafter referred to as "NCI Cancer Centers" or NCI-CCs) have demonstrated "scientific excellence and the capability to integrate a diversity of research approaches to focus on the problem of cancer" (National Cancer Institute, 2013). The NCI-CCs are not the only cancer care providers in the U.S. Other specialized cancer care facilities also include (1) the NCI-CC satellite facilities, (2) the Community Clinical Oncology Programs (CCOPs), and (3) academic medical centers (AMCs) in the Council of Teaching Hospitals and Health Systems (COTH). Nevertheless, the NCI-CCs represent perhaps the cancer care of the highest quality. There are currently a total of 66 NCI Cancer Centers. Patients cared for by the NCI Cancer Centers have much lower mortality rates (Onega et al., 2009) and higher five-year cancer-free rates in multiple types of cancer (Jemal et al., 2008; Landis et al., 1999).

A recent study by Shi et al. (2012) reports that there is a great deal of variability in spatial accessibility of the NCI Cancer Centers and other academic medical centers (AMCs), and much demand for quality cancer care is left unfulfilled. Uneven distributions of cancer care facilities and population lead to *geographic disparity* in accessibility,

exemplified by presence of ample service in some areas and absence or paucity of service in others. Furthermore, disproportionally higher numbers of racial and ethnic minorities often suffer from poor access to health care including cancer care (National Cancer Institute, 2008), commonly referred to as *racial disparity*. Both geographic and racial disparities contribute to deep gaps in access to care and health outcomes in the U.S.

Given the important role of NCI Cancer Centers as the "backbone" of the cancer care system in the U.S., it is important to further examine the disparities of spatial accessibility of NCI Cancer Centers across geographic areas and socio-demographic groups and to explore possible policy options to mitigate the problems.

2.3 Optimization Methods in Health Care Studies

On the methodological front, there is a rich collection of models in the study of planning for health care facilities, but most follow the line of classic location–allocation problems (Church 1999).

The first classic optimization problem is the classic location-allocation (or location modeling) problems (e.g., Church, 1999). The *p median problem* from Church's paper seeks to locate a given number of facilities among a set of candidate sites so that the total travel distance or time between demands and supply facilities is minimized. The *location set covering problem* (LSCP) minimizes the number of facilities needed to cover all demand within a critical distance or time. See Shavandi and Mahlooji (2008) for an application of LSCP in allocating health care facilities at different levels in Iran. The *maximum covering location problem* (MCLP) maximizes the demand covered within a desired distance or time threshold by locating p facilities. For example, Pacheco and Casado (2005) used it in allocating health care resources in Burgos, Spain.

The above models emphasize various objectives such as minimal travel, minimal resource, maximal coverage or a combination of them (i.e., multi-objective). Some recent work considered spatial accessibility. For example, Perry and Gesler (2000) used a target ratio of health personnel versus population and a maximum travel distance as criteria to adjust health personnel distribution to improve overall access. Gu et al. (2010) used a bi-objective model to identify optimal locations for health care facilities that maximize total coverage of population as well as their total accessibility.

However, none of these studies has equity as an objective. The value of equity is a matter of ethical obligation and needs to be recognized as rights to medical care (Fried, 1975). Equity in health and health care may be defined as equal access to health care, equal utilization of health care service or equal (equitable) health outcomes among others (e.g., Culyer and Wagstaff, 1993). Most agree that equal access is the most appropriate principle of equity from a public health policy perspective (Oliver and Mossialos, 2004: 656).

Most recently, Wang and Tang (2013) formulate the issue of equity in health care delivery as equal accessibility of health care services, and specifically proposes a new objective of minimizing inequality in accessibility of public services. More specifically, the objective is to minimize the variance (i.e., least squares) of accessibility index across all population locations by redistributing the total amount of supply among health care facilities. Tao et al. (2014) used a similar approach in planning residential care facility locations in Beijing, China, in order to achieve the maximum equality of accessibility for senior residents.

11

This research further advances this line of work on solving the planning problem of maximal equality in accessibility and applies the method to planning NCI Cancer Centers.

Chapter 3 Study Area and Data Preparation

The study area is the contiguous 48 states in the U.S. (excluding Hawaii and Alaska). In Hawaii or Alaska, the accessibility issue is confined within the state and thus more straightforward to address. The accessibility measures of the 48 contiguous states need to count for complex interaction between supply and demand across state borders, and thus can benefit more from the advanced methods to be discussed in the study.

Three variables are needed in defining spatial accessibility to cancer care: supply, demand and the geographic relationship between them. In this study, the supply is the NCI Cancer Centers and their corresponding capacities (e.g., numbers of beds), the demand for cancer care is potential cancer patients across geographic areas (e.g., counties or census tracts) in the U.S., and the link is travel time between them.

On the supply side, besides the NCI Cancer Centers, comprehensive hospitals are the second tier of cancer care facilities in the U.S. Most such hospitals are academic medical centers (AMCs), which are either independent or integrated with medical schools, and are members of the Council of Teaching Hospitals and Health Systems (COTH). The AMCs also provide high quality research and cancer care (Onega et al., 2008), and in fact, there is also considerable overlap between the NCI Cancer Centers and the AMCs. The 243 AMCs that are currently not NCI Cancer Centers thus may be the best candidates for the future designation. This study considers the 58 NCI Cancer Centers that currently provide care to patients in the contiguous 48 states as existing NCI Cancer Centers, and the 243 AMCs (currently not NCI Cancer Centers) as possible candidates for future designations. The combined data set of the 301 hospitals, including their geographic locations and existing beds counts, is shown in Figure 1. Data of the NCI Cancer Centers

are available from the NCI website (NCI, 2013). Data of the number of staffed hospital

beds were extracted from the American Hospital Directory

(http://www.ahd.com/state_statistics.html) and websites of individual hospitals.
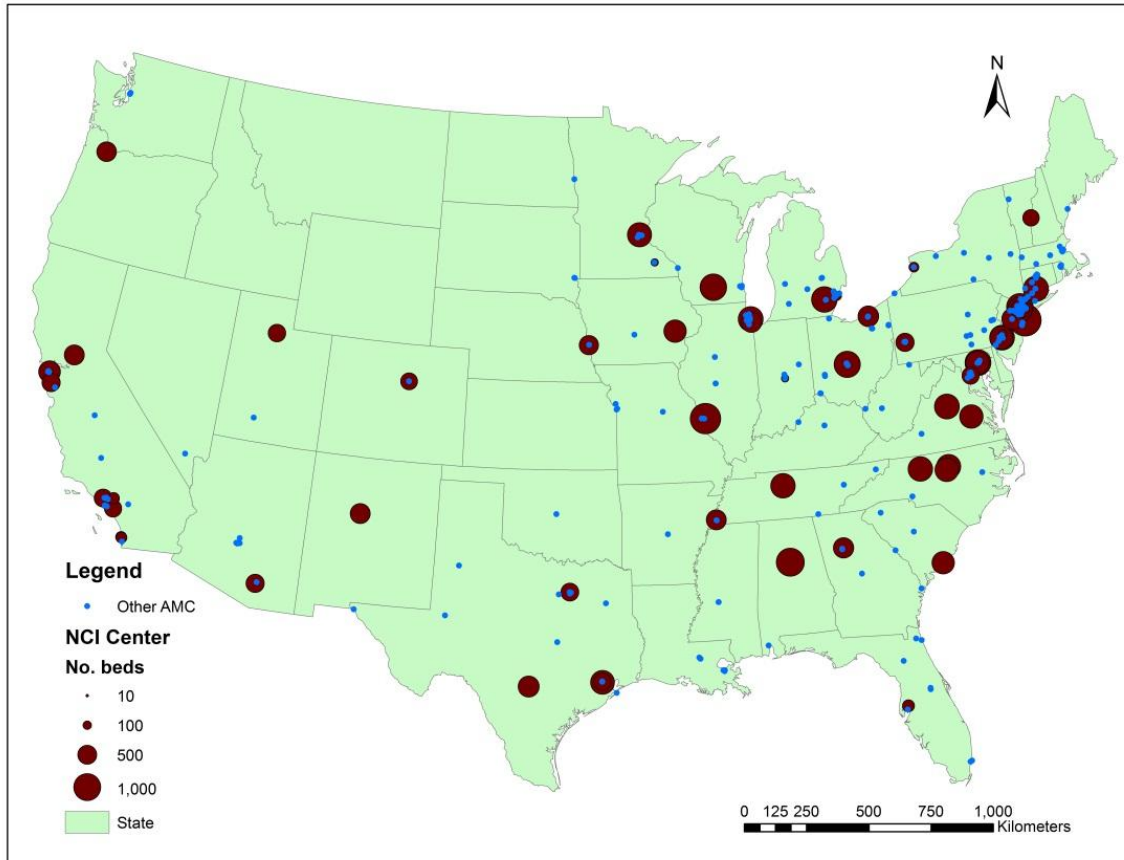
Figure 1 NCI Cancer Centers and other AMCs in the contiguous U.S.

As shown in Figure 1, most of the NCI Cancer Centers (NCI-CCs) are in the east

part of the country, and there is a cluster of NCI Cancer Centers in the New England Area.

A few centers are in west and the Great Plains expect for the clusters around San Francisco

and Los Angeles in California. There are three centers in the three largest cities in Texas

(Houston, Dallas and San Antonio). The sizes of these centers vary a great deal. Among

the 58 NCI Cancer Centers, the smallest one is the USC Norris Comprehensive Cancer

Center in Los Angeles with 26 staffed beds, and the largest one is the Albert Einstein Cancer Center at the Montefiore Medical Center in New York with 1409 staffed beds.

On the demand side, we will use the analysis units of both county and census tract to examine the presence of *possible modifiable areal unit problem (MAUP)*. For the purpose of illustrating the methodology, the number of cancer patients was estimated using the 2010 Census data (http://www.census.gov/) and the nationwide standardized cancer prevalence rates by age, sex, and race and ethnicity. The number of potential patients in each demographic group in a census tract was calculated by multiplying the population in that category by the corresponding cancer rate. The total estimate was the sum of patients from all categories (Shi et al., 2012). To more accurately represent the location of each county, we calculated population-weighted centroid for each county based on the 2010 census tract data. Similarly, we used population-weighted centroid for each census tract based on the 2010 census block data. We are aware of the spatial and temporal variability of cancer rates. More accurate estimate of cancer care demand will require the cancer prevalence data by age, sex, and race and ethnicity at sharper geographic resolutions from the North American Association of Central Cancer Registries (NAACCR) (www.naaccr.org) data and SEER*Stat (http://seer.cancer.gov/seerstat/). Future research will provide refined and more in-depth analysis of the issue.

According to the 2010 Census data, the study area contains 3,138 counties, 22 of which have zero population and thus not included in the analysis. The remaining 3,109 counties represent the demand side at the county level (as shown in Figure 2). There are 72,238 census tracts in the contiguous U.S., 210 of which have no population. The 72,028

census tracts with non-zero population form the demand side at the census tract level (Figure 3).
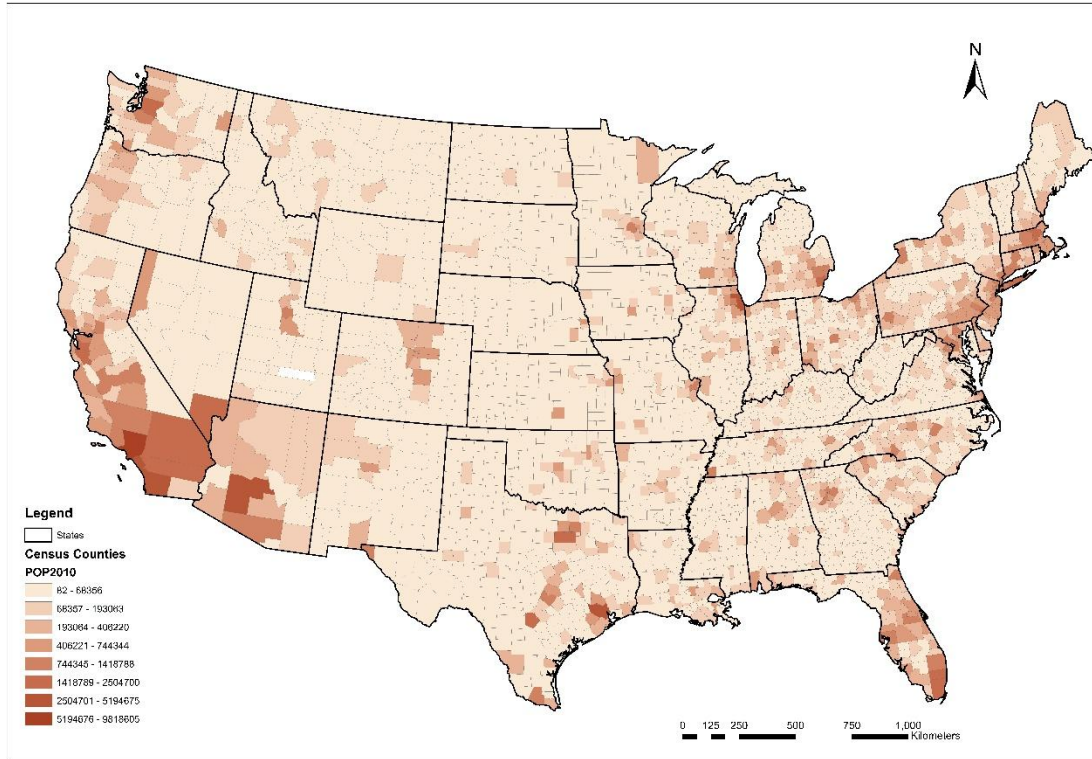
Figure 2 Population at the county level in the contiguous U.S.

The travel time between each county (census tract) centroid and each hospital was computed using the ArcGIS Network Analyst module. The road network data for this task were extracted from the 2012 ESRI StreetMap USA data that came with the ArcGIS 10.1 release. We considered all the streets and major roads, including interstate, the U.S. and state highways, due to the computational limitation of both software and hardware. The estimation of travel time assumes that travelers take the shortest path and follow the speed limit posted on each road section. This approach is adequate for capturing the travel impedance between patients and hospitals at the national scale for planning and public policy analysis. In addition, it assumes that patients seeking the specialized cancer care in

16

NCI-CCs travel by private vehicles. Conceivably, people may also choose other transportation modes such as by air or railway. The former incurs considerably high financial cost and the latter is very limited in the U.S., and neither is considered by this study. In preparing for the computation of travel time, building the road network dataset took over 14 hours on a HP Pavilion dv7 laptop (2.00GHz CPU and 6.00GB memory). After that, the computation time for estimating the travel time matrix between 3,109 county centroids and combined 301 hospitals was negligible. However, the computation for the travel time matrix between 72,043 census tract centroids and 301 hospitals took about 26 hours on the same computer.
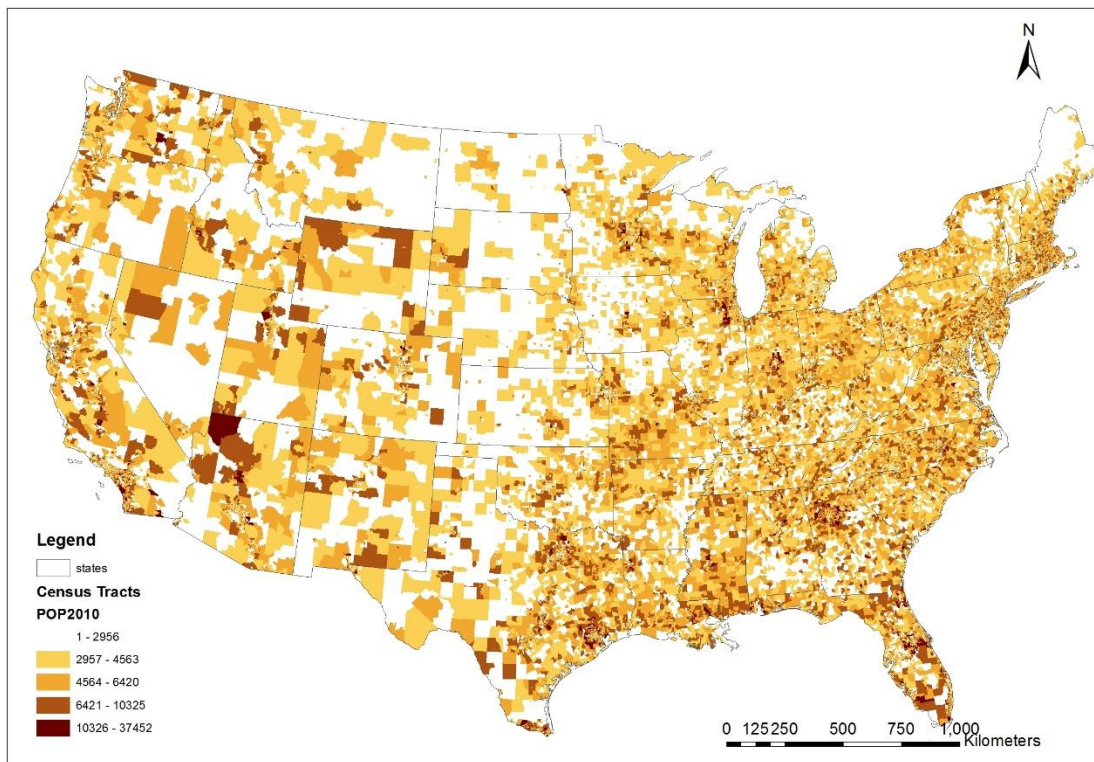
Figure 3 Population at the census tract level in the contiguous U.S.

Chapter 4 Measuring Spatial Accessibility and Its Geographic Disparity

4.1 The Multiple Floating Catchment Areas Method

Based on the literature review in Chapter 2, the most recent advancement in methods that count for complex interactions between supply and demand located in different areas is synthesized into the generalized 2SFCA method. The method is summarized in Equation (2) in Chapter 2, and rewritten here for convenience:

$$A_i = \sum_{j=1}^{n} [S_j f(d_{ij}) / (\sum_{k=1}^{m} D_k f(d_{kj}))]$$

where A_i is the accessibility to NCI Cancer Centers at county or census tract i; S_j is the number of staffed beds representing the capacity of NCI Cancer Centers at a cancer center location (supply) j; D_k is the estimated number of patients at county or census tract (demand) k; d is the travel time between them; and n and m are the total numbers of supply locations and demand locations, respectively.

The derived accessibility value is basically the ratio of supply capacity (i.e., number of beds) to demand (i.e., number of patients), i.e., number of beds per potential patient. Therefore, a higher A_i value indicates better accessibility. In fact, it is proven that the weighted average accessibility score across the whole study area is the ratio of total supply to total demand (Shen, 1998; Wang, 2006).

One challenge in implementing the above method is to define the distance (here measured in travel time) decay function $f(d)$. This research adopts a decay schema with six discrete values based on a Gaussian function. Such a strategy is similar to the 6-ring catchment area method suggested by Shi (2012), which also examined the accessibility of specialized cancer care in the U.S. that included all academic medical centers (AMCs). However, Shi (2012) used a different model in measuring accessibility to fit his purpose.

18

The similarity between our methods lies in the way of conceptualizing the distance decay behavior in visiting specialized cancer care hospitals.

More specifically, the interaction between a patient and a hospital declines with travel time, and the declining pattern is designed according to several discrete values (weights). In this case, we choose six values corresponding to six 30-minute increments (rings) of travel time. The innermost ring defines the area in which travel to the facility takes 0 to 30 minutes, the second ring defines the area between 30 and 60 minutes, … and the outermost ring defines the area taking 150 to 180 minutes. The six discrete values for $f(d_{ij})$ are calculated using a Gaussian function:

$$f(d_{ij}) = \ell^{\ln(0.01)(\frac{x-1}{7})^2} \quad (x = 1, 2 \ldots 6)$$

where $f(d_{ij})$ is the weight for ring x (= 1, 2, ... 6 from inner to outer) and e is the natural base (2.71828 ...). The corresponding weights are summarized in Table 1. The three-hour cap for travel time is almost the limit for a patient to travel to a hospital (one way), obtain some service and return home in a single day. For areas beyond the sixth ring, the weight is arbitrarily set as 0.015. The six rings are considered six service-to-patient catchment areas, and therefore the accessibility measure can be termed the *Multiple Floating Catchment Areas Method*.

The chosen Gaussian function resembles a bell-shaped curve, widely believed to capture the spatial behavior of hospital visits (Shi, 2012). Nevertheless, it is a choice that ideally should be made by analyzing the actual hospital visitation data, as attempted by Delamater et al. (2013). See Páez et al. (2012) for more discussion on the selection of distance decay functions and related parameters. Our future work will resolve this issue.

Table 1 Table for converting travel time to weight using a Gaussian function (Shi, 2012)

Travel time (minutes): d_{ij}	Declining weight: $f(d_{ij})$
0-30	1.00
30-60	0.91
60-90	0.69
90-120	0.43
120-150	0.22
150-180	0.10
> 180	0.015

4.2 Spatial Accessibility across Counties and Census Tracts

Applying the Multiple Floating Catchment Areas Method to estimated patients at the county level yields the spatial accessibility of NCI-CCs at the county level. As stated previously, the accessibility scores may be interpreted as numbers of NCI-CC hospital beds per potential patient, and thus very small numbers. The values are thus multiplied by 100,000 to indicate the numbers of beds per 100,000 patients. The results are shown in Figure 4.

Figure 4 shows that the accessibility generally exhibits a concentric decline from each NCI-CC often in a large city. Much of the country such as the massive Great Plains, the west edge of Midwest, west of Texas and much of the Deep South along the Gulf of Mexico, have the lowest accessibility scores since they are not covered by any NCI-CCs within the 3-hour travel time limit. In areas where the NCI-CCs are close to each other such as the long stretch from Massachusetts to Maryland, patients may reach two or more centers within 3 hours and enjoy better accessibility.
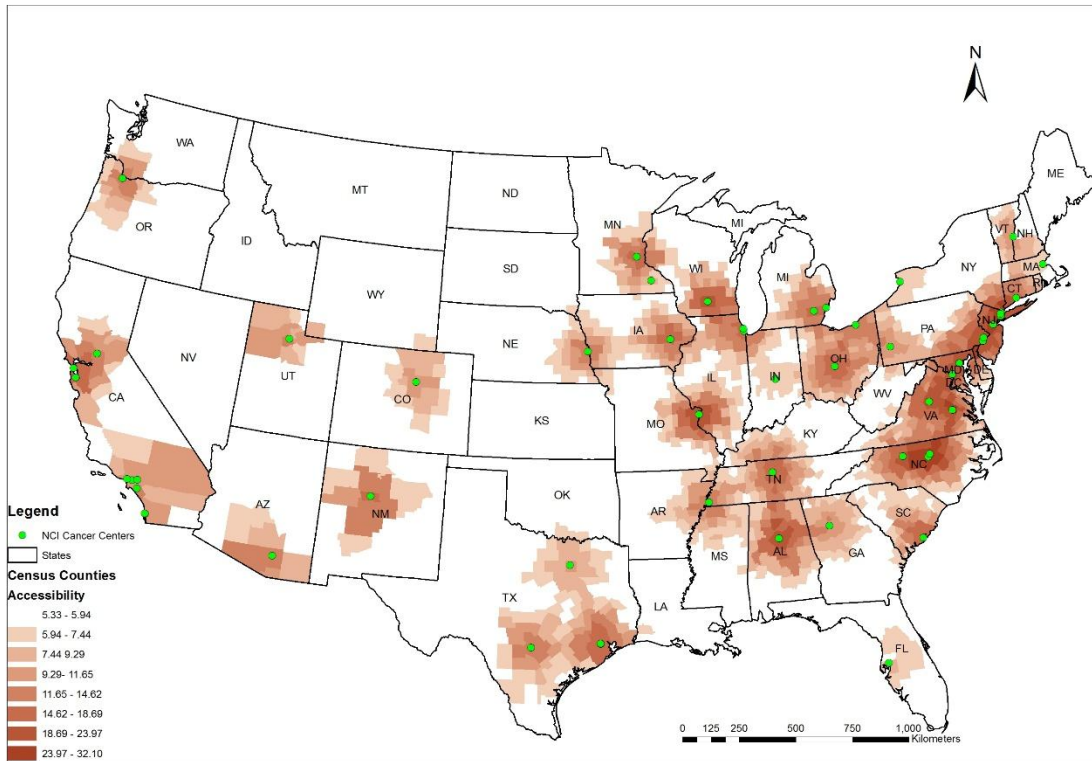
Figure 4 Accessibility of NCI Cancer Centers at the County Level

To further highlight the geographic pattern of spatial accessibility, we aggregated the accessibility scores to the nine census divisions. Denoting the population in a county i as P_i and its accessibility as A_i, the population-weighted accessibility in a census division is defined as:

$$A = \sum_{i=1}^{m}(A_i P_i)/\sum_{i=1}^{m}P_i$$

The result is shown in Figure 5. Census divisions are groupings of states and the District of Columbia, and each is identified by a single-digit census code. Puerto Rico and the Island Areas are not part of any census division. One may also aggregate the result to even larger areas such as the four census regions—Northeast, Midwest, South, and West. The resolution is too coarse at the census region level, and the result is not discussed here.

21
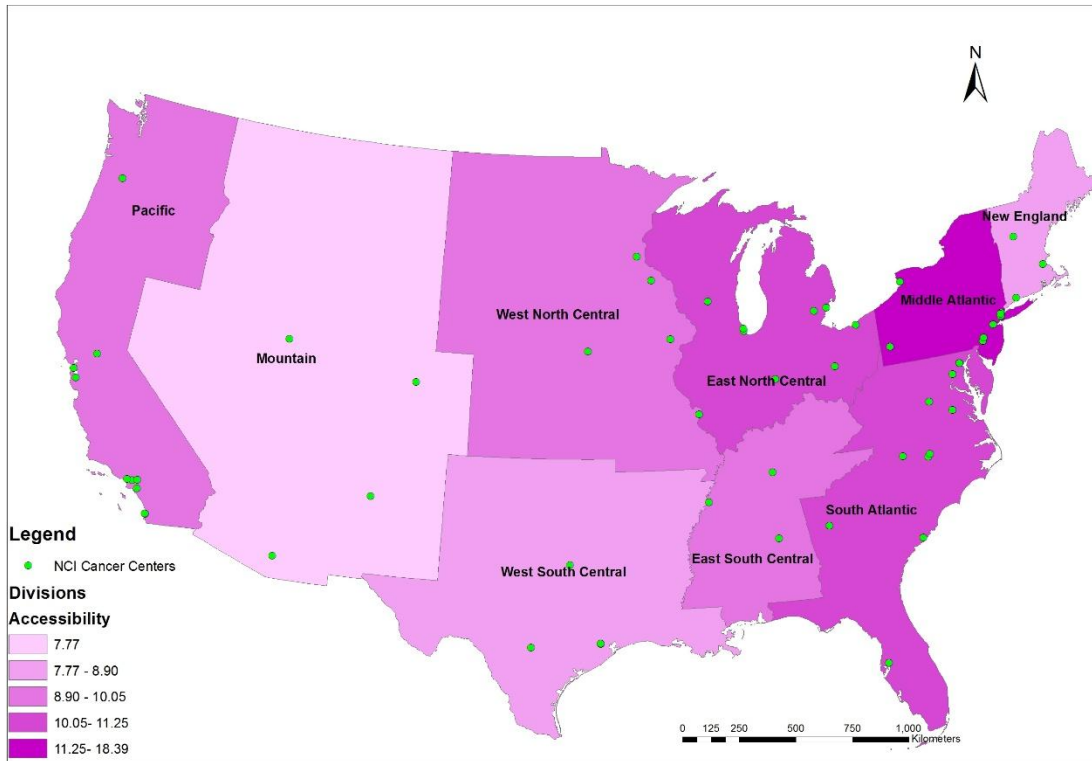
Figure 5 Accessibility of NCI Cancer Centers across Census Divisions

Figure 5 shows that the Middle Atlantic Region enjoys the highest accessibility overall, followed by its western and southern neighbors, i.e., the East North Central Region and South Atlantic Region. These three regions are known for relative higher population densities and more urbanized. On the other end, the Mountain Region with the lowest population density suffers from the poorest accessibility. The New England Region, though well developed, remains less urbanized and has the second lowest accessibility. The West South Central Region, composed of Arkansas, Louisiana, Oklahoma, and Texas, also has relatively low accessibility. The other three regions (Pacific, West North Central and West South Central) have medium accessibility scores. The next section will examine the variability of accessibility by urbanicity levels in depth.

22

Similarly, when the estimated patients at the census tract level are used to define the demand side, the spatial accessibility of NCI-CCs is derived at the census tract level, as shown in Figure 6. The general pattern in Figure 6 is consistent with that in Figure 4, but at a finer resolution.
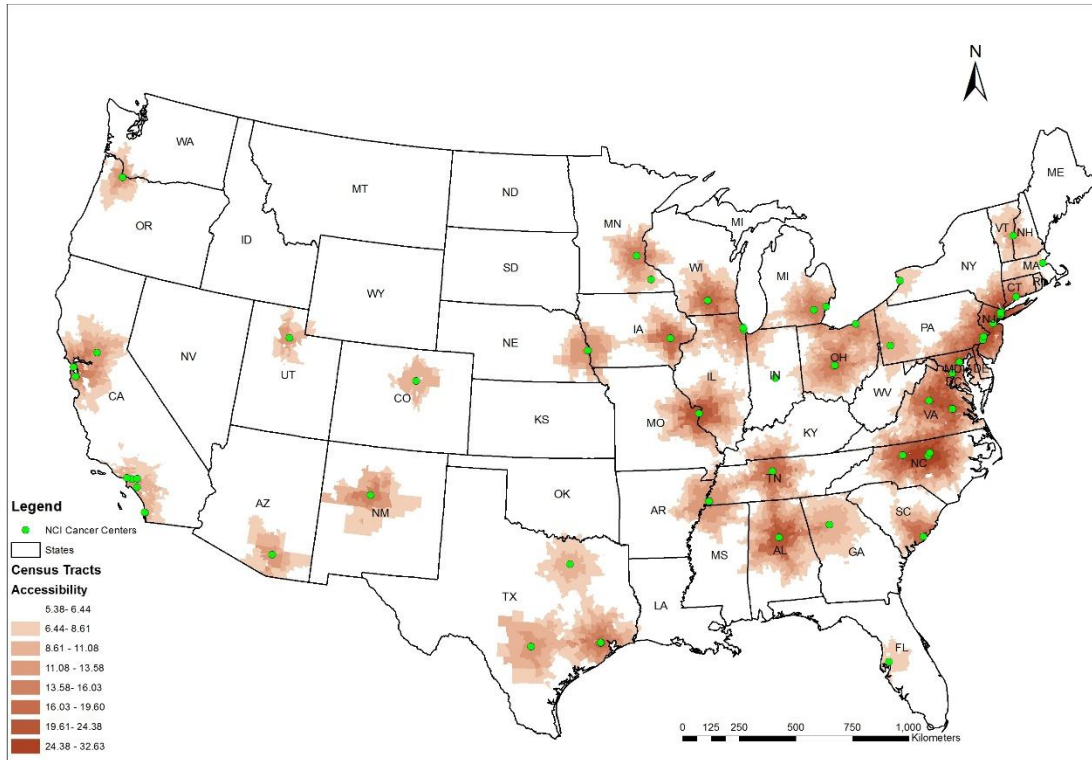
Figure 6. Accessibility of NCI Cancer Centers at the Census Tract Level

4.3 Spatial Accessibility by Levels of Urbanicity

Findings from the previous section suggest that the variability of spatial accessibility of NCI-CCs be related to urbanicity levels. This section explicitly examines this likely association.

For urbanicity, we first use the 2013 NCHS Urban–Rural Classification Scheme for Counties prepared by the National Center for Health Statistics (NCHS, 2006). There are six urban–rural categories such as large central metro, large fringe metro, medium

23

metro, small metro, micropolitan and noncore (Figure 7). This definition of urbanicity is used for the county-level analysis.
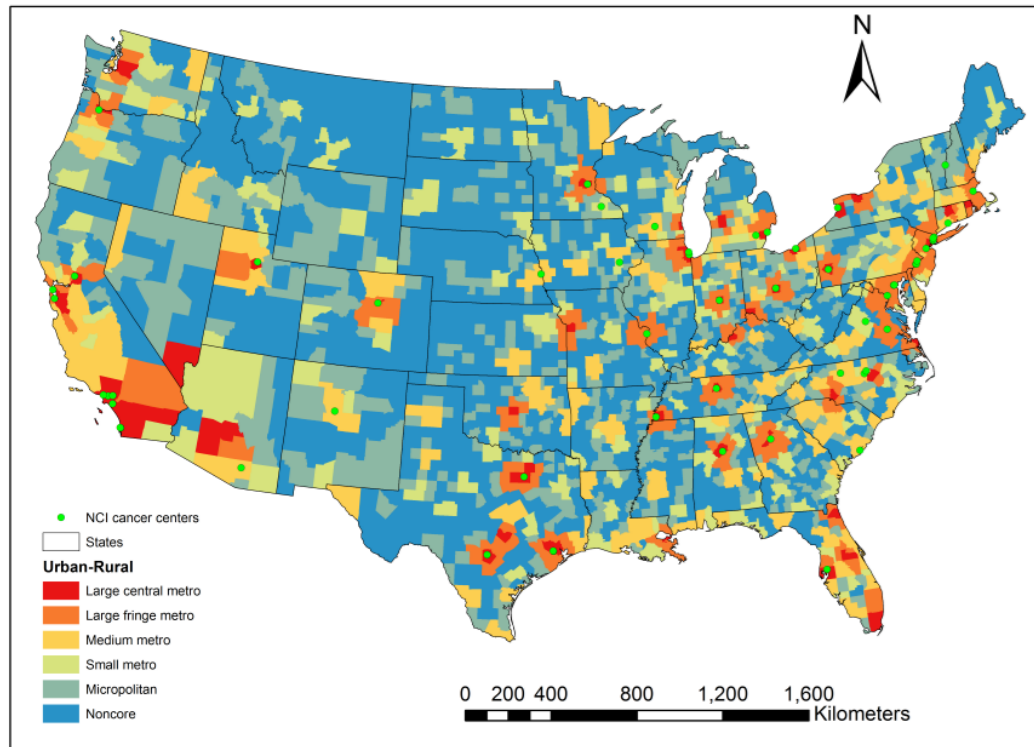
Figure 7 NCHS definition of urbanicity at the county level

Since the NCHS is based on the county unit, we use a different definition to capture the urbanicity at the census tract. For the analysis at the census tract level, this research uses the 2010 Census Urban and Rural Classification (U.S. Census Bureau, 2015). The Census Bureau defines an urban area on census tracts and/or census blocks that meet minimum population density requirements. Urbanized areas (UAs) (50,000 or more people) and urban clusters (UCs) (at least 2,500 and less than 50,000 people) are two types of urban areas (Figure 8). A tract is classified as (1) Urbanized Area, (2) Urban Cluster or (3) rural if its centroid falls within an Urbanized Area, Urban Cluster or rural area, respectively.
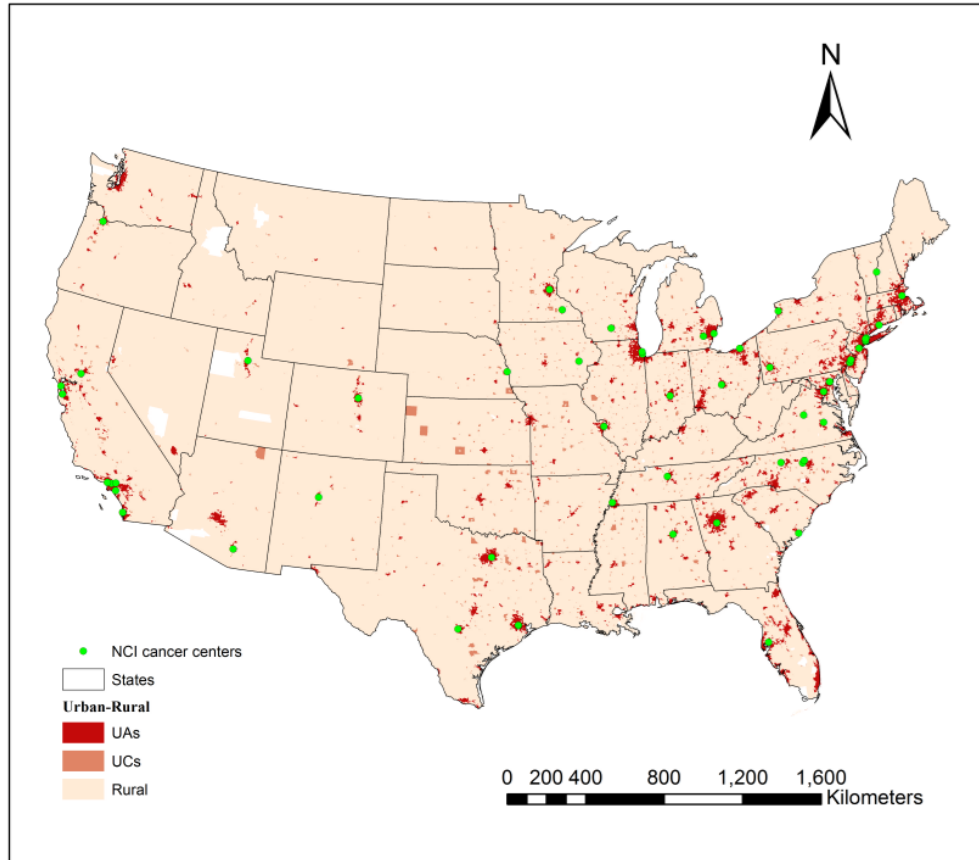
Figure 8 Census Bureau definition of urbanicity

Table 2 presents the average values of accessibility in various urban-rural categories. The average accessibility at the county level declines with urbanicity level from large central metro counties to large fringe metro counties, …and to noncore (rural) counties (also shown in Figure 9a). At the census tract level, clearly the average accessibility values are the highest in Urbanized Areas; however, its values are slightly lower in Urban Clusters (mostly on the urban fringe) than in rural areas, but the difference is minor (shown in Figure 9b). The urban advantage remains in general at the census tract level, and the slightly reversed order between Urban Clusters and rural areas may be attributable to generally higher demands (i.e., population) in Urban Clusters than rural areas, which has been captured by the G2SFCA method.

Table 2 Average Accessibility by Urbanicity Levels

	Accessibility (County)	Accessibility (Census Tract)
Urban-rural classification (County)		
Large central metro	13.92	
Large fringe metro	12.04	
Medium metro	8.40	
Small metro	7.43	
Micropolitan	6.90	
Noncore	6.25	
Urban-rural classification (Census Tract)		
Urbanized areas		12.67
Urban clusters		7.92
Rural		8.33

Are the observed differences in the average values of accessibility statistically significant across urban-rural categories? A simple regression model with dummy variables is formulated to answer this question (Xu et al. 2015). The variable of interest, accessibility value, defines the dependent variable in the regression; and the independent variables are the dummy variables that code the urban-rural categories. For instance, five dummy variables are used to code six urbanicity categories at the county level ("large central metro" as the reference type); and two dummy variables are used to code three urbanicity categories at the census tract level (rural tracts as the reference type).

The results are reported in Table 3. Taking the model for the county level as an example, the intercept from the regression model (13.92) is the average accessibility for the reference category (i.e., large central metro counties), and the coefficient for a category is the difference between the average accessibility of the reference category and this

category. For instance, the coefficient -1.88 for large fringe metro counties indicates that the average accessibility of large fringe metro is 1.88 below that of large central metro. This confirms that the average accessibility for large fringe metro is 13.92-1.88 = 12.04, consistent with that shown in Table 2. What is new from the regression model is the t-statistic value associated with this coefficient (here, -4.03, in parentheses in Table 3), indicating that the difference is highly significant (i.e., significant at 0.001). Similarly, as the corresponding coefficient becomes more and more negative from large fringe metro, to medium metro, small metro, micropolitan, and non-core, the average accessibility keeps declining (consistent with those reported in Table 2). Moreover, the t-values indicate that the differences in average accessibility between the reference category (large central metro counties) and any of the five other categories are all statistically significant.

Similarly, for the central tracts, the result from Table 3 indicates that the difference in average accessibility between Urbanized Areas and Urban Clusters (or rural areas) is statistically significant, and the order of their values is Urbanized Areas > rural > Urban Clusters.
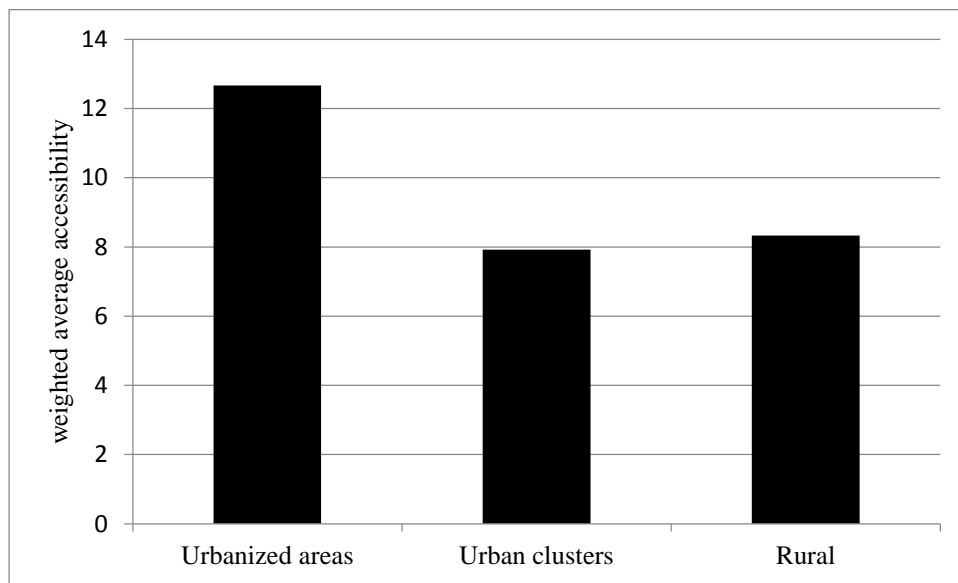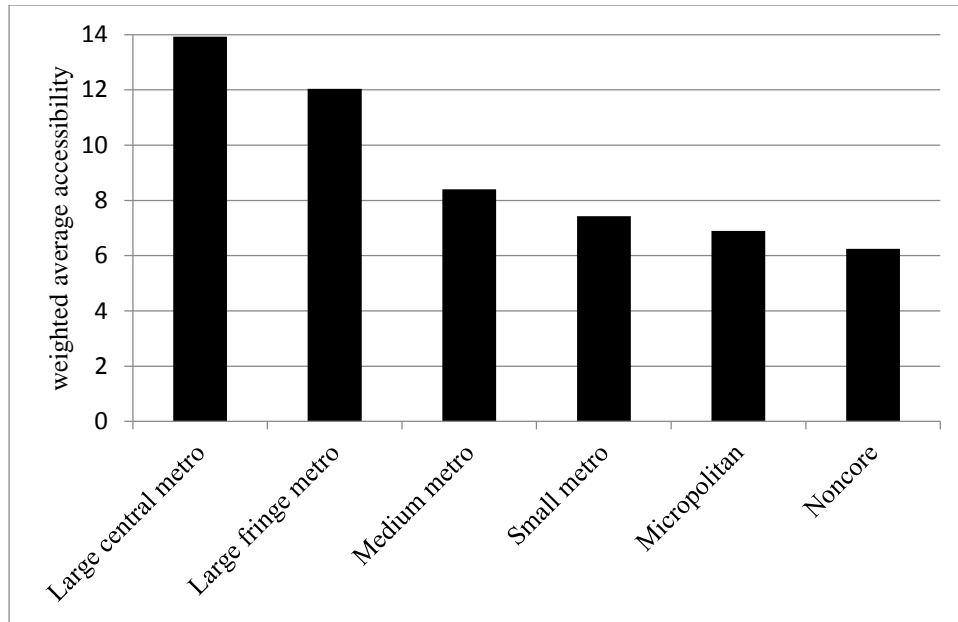
Figure 9 Average accessibility across urbanicity levels: (a) by six categories at the county level, (b) by three categories at the census tract level

Table 3 Regression Results for Testing Statistical Significance in Variability of Accessibility across Urban-Rural classifications

	Accessibility (County)	Accessibility (Census Tract)
Urban-rural classification (County)		
Large central metro	13.92***	
	(32.43)	
Large fringe metro	-1.88***	
	(-4.03)	
Medium metro	-5.52***	
	(-11.83)	
Small metro	-6.49***	
	(-13.84)	
Micropolitan	-7.02***	
	(-15.53)	
Non-core	-7.67***	
	(-17.41)	
Urban-rural classification (Census Tract)		
Urbanized areas		12.67***
		(463.87)
Urban clusters		-4.75***
		(-45.34)
Rural		-4.34***
		(-89.84)

*** $p<=0.001$, t-value in parentheses.

Chapter 5 Disparities of Accessibility by Demographic Groups

As stated in Chapter 2, both geographic and racial disparities in access to health care are of great concern to public policy as both contribute to inequality in health outcomes. Chapter 4 examined geographic disparity in accessibility of NCI Cancer Centers. This chapter analyzes racial disparity in accessibility. The methodology illustrated here can be applied to analysis of disparity across other demographic groups (e.g., people of various income groups, educational attainments, employment status or family structures etc.). Once again, the analysis is limited to the ecological nature of the data. The finding is not consequentially transferrable to individuals. In other words, without access to individual data, we cannot clearly compare the accessibility of one racial-ethnic group to others at the individual level, rather the difference among various groups *on average*.

5.1 Average Accessibility by Demographic Groups

Demographic information including race and poverty level is extracted from the 2010 Census. Poverty rate is the estimated percent of people of all ages in poverty. In this study, three levels of poverty status are defined. A county (census tract) with the poverty rate lower than 10% is defined as low poverty, between 10% and 20% is defined as medium poverty, and higher than 20% is high poverty. Figures 10a-c show the spatial patterns of the percentage of Black, Minority (or non-White), and population under the poverty line at the county level. Figures 11a-c show the patterns of corresponding variables at the census tract level. Black and non-White are used as examples as we focus our analysis on disparity across racial-ethnic groups. The poverty pattern is shown to highlight that patterns of other socioeconomic variables may be consistent with that of race to some extent but may also differ significantly.
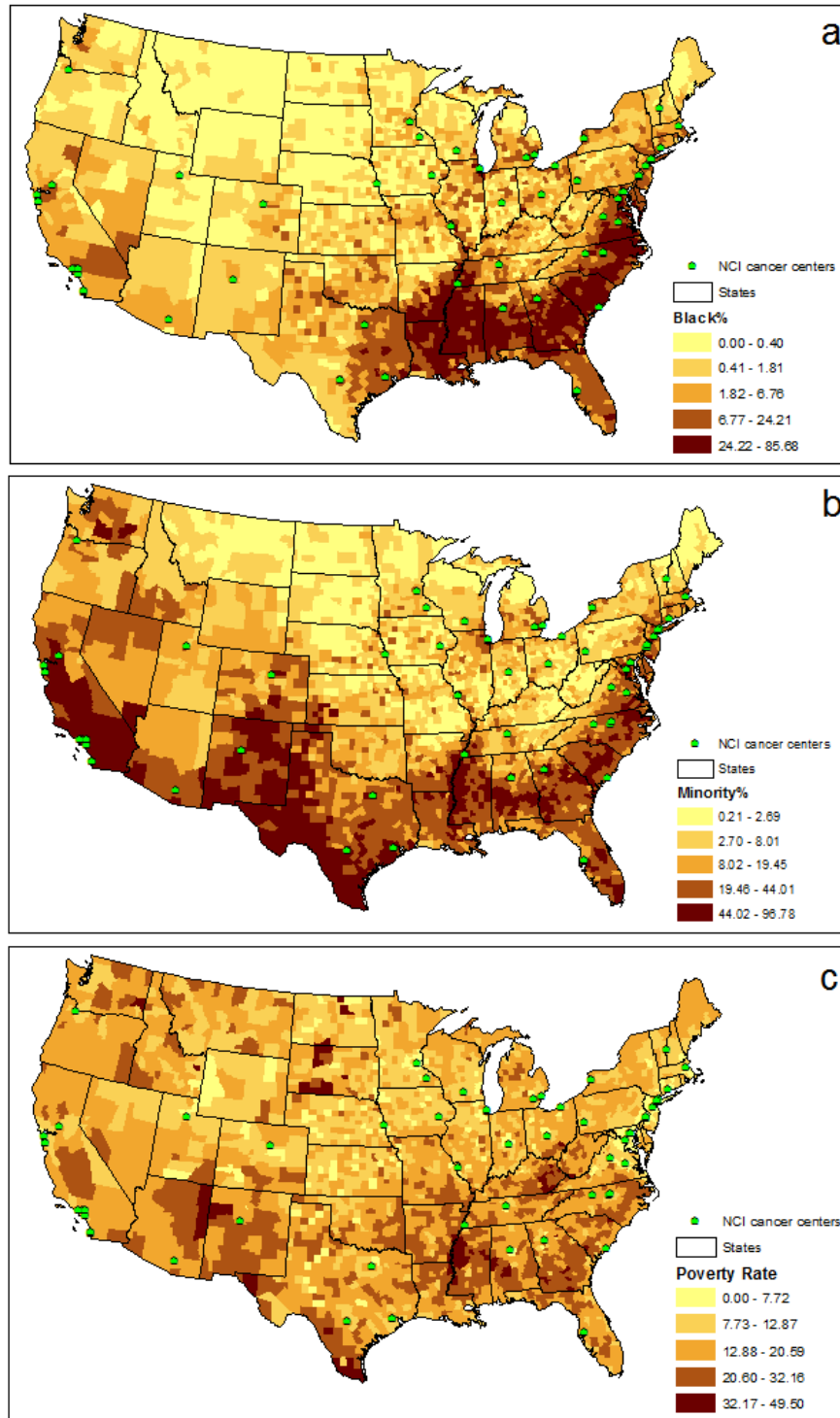
Figure 10 Distribution of demographic groups at the county level: (a) percentage of
Black; (b) percentage of Minority; (c) poverty rate
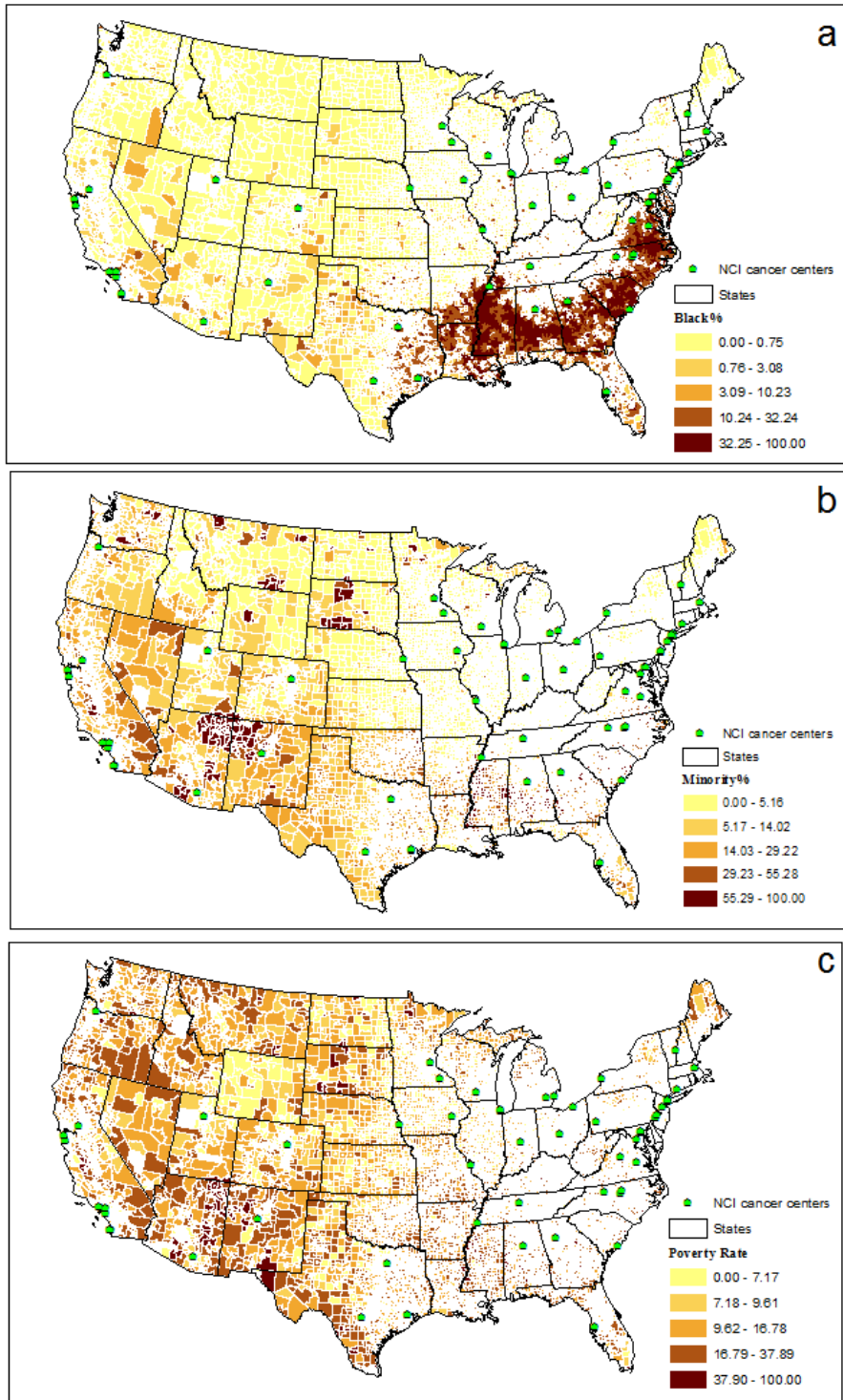
Figure 11 Distribution of demographic groups at the census tract level: (a) percentage of Black; (b) percentage of Minority; (c) poverty rate

Figures 10a and 11a show that high percentages of Blacks are observed in the south and also southeast coastal states including LA, MS, AL, FL, GA, SC, NC, VI, MD and stretching towards northeast. There are over 20 NCI Cancer Centers located in those states. For the non-White minorities, Figures 10b and 11b add areas in the southwest bordering the U.S. and Mexico such as CA, AZ, NM and TX with a significant presence of Hispanic population. In addition to seven NCI Cancer Centers in CA, there are three in TX, one in NM and one in AZ. The poverty pattern is more scattered.

Note that people of various racial-ethnic groups could be present in the same area (a county or a census tract) and have the same spatial accessibility of NCI-CCs. It is the variability of their various concentrations (i.e., percentages) across geographic areas that lead to disparity. This section begins with comparing the average accessibility values for different groups to gain some preliminary understanding of the issue. Table 4 reports the weighted average accessibility for various racial-ethnic groups.

From Table 4, Whites have the lowest accessibility score to the NCI cancer centers among the ethnic groups, followed by Hispanics, Minorities in general, Blacks, and Asians. It may be attributable to that a disproportionally high number of Whites tend to locate in the suburbia or rural areas, and thus suffer from the poorest spatial accessibility. Onega (2008) also found that African Americans have a shorter median travel time from their nearest NCI-CCs (69 minutes) than the Whites (86 minutes), and Asians were found to have the shortest median travel time (28 minutes). The somehow surprising finding implies that when it comes to spatial accessibility, Blacks and Hispanics actually enjoy better accessibility than Whites. We may term this as "*reversed racial (dis)advantage*."

Table 4 also shows that low-poverty areas (either at the county or census tract level) enjoy the best average accessibility. However, high-poverty areas have a slight edge over medium-poverty areas in average accessibility but the difference is minor.

Table 4 Weighted average accessibility for various demographic groups at county and census tract levels

	Accessibility (County)	Accessibility (Census Tract)
All population	11.14	11.14
By Race-Ethnicity		
White	10.59	10.58
Black	13.00	12.99
Hispanic	11.43	11.51
Asian	13.90	13.92
Minority (non-White)	12.36	12.40
By Poverty Level		
Low poverty (<10%)	14.62	12.09
Medium poverty (10-20%)	10.38	10.25
High poverty (>20%)	10.52	10.72

5.2 Testing Statistical Differences in Accessibility by Demographic Groups

Are the aforementioned differences in average accessibility scores among demographic groups statistically significant? Prior to the formulation of statistical tests, note that the variables of interest are constructed differently. For racial-ethnic groups, each group (e.g., White, Black, etc.) is theoretically scattered in every county (tract), and the weighted average accessibility for a group is computed by using its population as the weight across all counties (tracts). For poverty status, the study area is divided into three types of areas according to their poverty percentages, and the weighted average for each is computed within that subset of areas and the weight is each county's (tract's) entire

population. Therefore, the statistical tests are designed differently. Since census tract carries a sharper spatial resolution than county, the following analysis uses the tract as the analysis unit.

For racial disparities, one way to formulate the null hypothesis (H_0) is such as: the ratios of a racial-ethnic group in areas with above-average accessibility values are the same as those areas with below-average accessibility values. If this null hypothesis for the one-tailed test is rejected, the alternative that the ratios of this particular population group are significantly higher (lower) in areas of above-average accessibility is accepted. This may be measured by conducting a pooled t-test to compare the sample mean of the group's percentages in areas with accessibility higher than average with those below the average. For simplicity and easy interpretation, we follow the weighted regression model proposed by Irkam et al. (2015) to implement the test. The weighted ordinary-least-squares (OLS) regression model is formulated as

$Y = a + b * Flag$

where the dependent variable *Y* stands for ratios of ethnic groups in various census tracts, the independent variable *Flag* is a binary dummy variable (= 0 or 1, flagging whether a tract has an accessibility value above or below the average), and *a* and *b* are parameters to be estimated. Population in each tract is used as the weight in the model.

For example, using 72,028 census tracts, the average accessibility value is 11.12. All the tracts are split into two parts: tracts in Part 1 are coded as "*Flag=0*" in which the accessibility values are larger than or equal to 11.12, the rest in Part 2 with value less than 11.12 are coded as "*Flag=1*".

The regression results are presented in Table 5. For example, the model result for Whites is

$Y = 0.2651 + 0.1961 * Flag$

(31.96)

The parentheses underneath the equation is the corresponding t value for slope b. Based on the result, when $Flag=0$, $Y= 0.2651$, which is the intercept, i.e., the sample mean of White ratios in tracts above the average accessibility value. When $Flag = 1$, $Y = 0.2651 + 0.1961 = 0.4612$, which is the sample mean of White ratios in tracts below the average value. The slope $b = 0.1961$ indicates that the difference between the two sample means. Here, a positive b implies higher White ratios in part 2 (below-average-accessibility tracts) than in part 1 (above-average-accessibility tracts). The corresponding t value $= 31.96$ indicates the difference is highly significant ($p<0.001$). By employing a weight term in the regression, the error term is weighted heavier in a census tract/county with more population than one with less population.

Table 5 Percentages of racial-ethnic groups in census tracts

	Above-average	Below-average	Difference	t value
White	26.51	46.12	19.61	31.96***
Black	6.48	6.20	-0.28	-37.78***
Hispanic	6.82	9.58	2.76	5.89***
Asian	2.71	1.89	-0.82	-49.72***
Minority	16.02	17.67	1.63	39.06***

*** $p<=0.001$

From Table 5, at the census tract level, rates of the White, Hispanic and Minorities are higher in the areas with below-average accessibility values. Black and Asian ratios are

higher in the areas with above-average accessibility values. These findings based on rigorous statistical analysis are consistent with the preliminary findings suggested by the simple weighted averages for White, Black and Asian, but differ from the preliminary findings for Hispanic and Minorities. That highlights the value of rigorous statistical analysis that interpretation from simple comparison of average accessibility scores may be misleading.

For analysis of accessibility across areas of various poverty levels, the model is straightforward and similar to the regression model with dummy variables used in Chapter 4. Using the census tracts with low poverty (poverty rate <10%) as the reference category, two dummy variables are used to code three categories of poverty levels in the regression model, and the dependent variable is accessibility value in each tract. The result is presented in Table 6, completely consistent with that reported in Table 3. For example, the intercept 12.09 is the average accessibility score for low-poverty tracts (reference category). The coefficient for medium-poverty tracts (poverty rate between 10% and 20%) is -1.84, indicating that the average accessibility score for medium-status tracts is -1.84 below 12.09 (reference category), i.e., 12.09-1.84=10.25. Similarly, the average accessibility score for high-poverty tracts (poverty rate >=20%) is 12.09-1.37=10.72. What new we can learn from the regression model is the statistical significances associated with the above findings based on the t values corresponding to the dummy variables. Based on Table 6, the differences across the three poverty levels are all statistically significant at the 0.001 level. In other words, tracts with the low poverty level generally enjoy the best accessibility, followed by high-poverty tracts and then medium-poverty tracts.

Table 6 Regression result for variability of accessibility across census tracts of different poverty levels

	Accessibility (Census Tract)
Low poverty	12.09***
	(330.13)
Medium poverty	-1.84***
	(-33.37)
High poverty	-1.37***
	(-24.23)

*** $p <= 0.001$

Chapter 6 Optimization for Minimal Geographic Disparity in Spatial Accessibility

Chapter 4 examines the geographic variability of spatial accessibility of NCI Cancer Centers across counties and census tracts. This chapter proposes several planning problems for minimizing the disparity (inequality) in accessibility across geographic areas by some adjustment in resource allocation. The optimization of equality became an issue after equality planning problems arose (Ostor, 2008). The scenarios are formulated as optimization problems composed of an optimal objective and a set of constraints, and various solution methods are explored and compared.

6.1 Formulating the Planning Problems

A planning (optimization) problem is to find the best available value for the objective function given a set of constraints.

We begin the discussion on what the *objective* is and how to formulate it. As suggested by Schrage (1986), one may maximize equality, or alternatively minimize inequality, across geographic areas by:

1) minimizing the sum of the absolute deviations (MAD), i.e.,

$$\min Varf(x)$$

2) minimizing the maximum absolute error of the range (MINIMAX), i.e.,

$$\min \left| \max f(x) - meanf(x) \right|$$

The second formulation focuses on the gap between a single area with the best accessibility and the overall average accessibility, and seeks to minimize that gap. It is conceivable that the maximum accessibility is observed in an area with very few people, and closing up the gap would have limited value. It is also more technically challenging in solving an

optimization problem with the MINIMAX objective. This chapter considers only the MAD objective.

For the MAD objective, variance is a measure of how far a set of values is spread out. Given an accessibility measure as defined in Equation (2) in Chapter 4, it is known that the weighted mean of the accessibility a is equal to the ratio of the total supply S $(=S_1+S_2+...+S_n)$ to the total demand D $(=D_1+D_2+...+D_m)$ in a study area (Shen, 1998), denoted by a constant a, such that

$$a = \sum_{i=1}^{m}(D_i / D)A_i = (1/D)(D_1A_1 + D_2A_2 + ... + D_mA_m) = S/D$$

Therefore, the MAD objective is to minimize the absolute variance (i.e., least squares) of the accessibility index A_i across all population locations, written as:

$$\min \sum_{i=1}^{m} D_i(A_i - a)^2 \qquad (3)$$

where accessibility gaps ($A_i - a$) are weighted by corresponding demand D_i.

Now, we turn our attention to the formulation of *constraints* and *decision variables* to be solved. Recall that the purpose of this research is to optimize spatial accessibility that is influenced by both supply and demand. On the demand side, people choose where to live for multiple reasons, including birth, work, education, family, political policy, weather, and environment. Access to cancer care may be one of these reasons, but it is usually not the determining factor, which is typically family or work. As a result, we cannot change people's living location to optimize the equality of visiting the NCI Cancer Centers, meaning that we cannot plan for the demand side. Therefore, we have to consider the supply, the allocation of resources related to NCI Cancer Centers in this case, as the

variable in the planning. As a starting point, we may speculate two kinds of decisions regarding the NCI resources: relocating the existing resources, or allocating new resources. Considering the high cost of moving employees and care equipment, we consider it a more realistic scenario of allocating some new available resources. These scenarios are set up in accordance with the most feasible policy options: allocating funds currently available only to existing NCI Cancer Centers and the periodical designation of new centers. These options include (1) expanding existing centers (e.g., measured in bed counts) and (2) designating new centers.

The first option involved in planning is how to allocate additional resources. Resources may be quantified financially, either by personnel or capacity. For illustration, this study examines how to allocate a fixed number of additional beds, denoted by a constant B. The decision variables are the new capacity (number of beds) of the existing NCI Cancer Centers, denoted by X_j. Note that the additional beds increase the total supply to $S+B$, and thus, also change the average accessibility $a = (S + B) / D$. The optimization problem is subject to the following constraints:

$$\sum_{j=1}^{n} X_j = S + B \text{, and}$$

$$X_j - S_j \geq 0 \text{ for all } j = 1, 2, \dots, n.$$

In practice, the number of beds X_j can be considered a continuous non-negative real number instead of an integer. However, for multiple reasons, like land use, transportation, and other urban planning reasons, one hospital or center cannot be given unlimited resources. As a result, we may cap the increase ratio at each hospital. For illustration purposes, we arbitrarily specify this cap to be 25%, so another constraint is added:

$1.25S_j \geq X_j - S_j \geq 0$.

The other planning option is to designate new NCI Cancer Centers from the existing AMCs. Say a given number (n_0) of new centers is to be designated. We denote the capacities (number of beds) of all NCI and AMC hospitals as S_j. However, this time $j = 1$, 2, …, 301. In other words, the number of potential supply sites is expanded from only the existing NCI Cancer Centers in the previous decision option to the total number of combined hospitals. We now introduce the binary decision variable $x_j = 0, 1$, where 1 indicates the corresponding hospital designated as an NCI Cancer Centers and 0 otherwise. For the existing 58 NCI Cancer Centers, $x_j = 1$ is already the predetermined solution; for the remaining 243 AMCs, x_j is the real variable to be solved. The optimization problem thus becomes a 0-1 integer programming problem.

Figure 12 summarizes the structure and various scenarios of optimization for minimizing geographic disparity in spatial accessibility.

More specifically, for decision option 1 by allocating new resource, the objective function in Equation (3) is rewritten as:

$$\min \sum_{i=1}^{m}[D_i(\sum_{j=1}^{n}\frac{(B_j + X_j)f(d_{ij})}{\sum_{i=1}^{m}D_i f(d_{ij})} - (X + S)/D)^2] \qquad (4)$$

For decision option 1, the constraints are:

$$\sum_{j=1}^{n}X_j = X + S \qquad (5)$$

$$1.25S_j \geq X_j - S_j \geq 0 \qquad (6)$$

where the variables X_j is the number of increased staffed beds at each NCI Cancer Center.
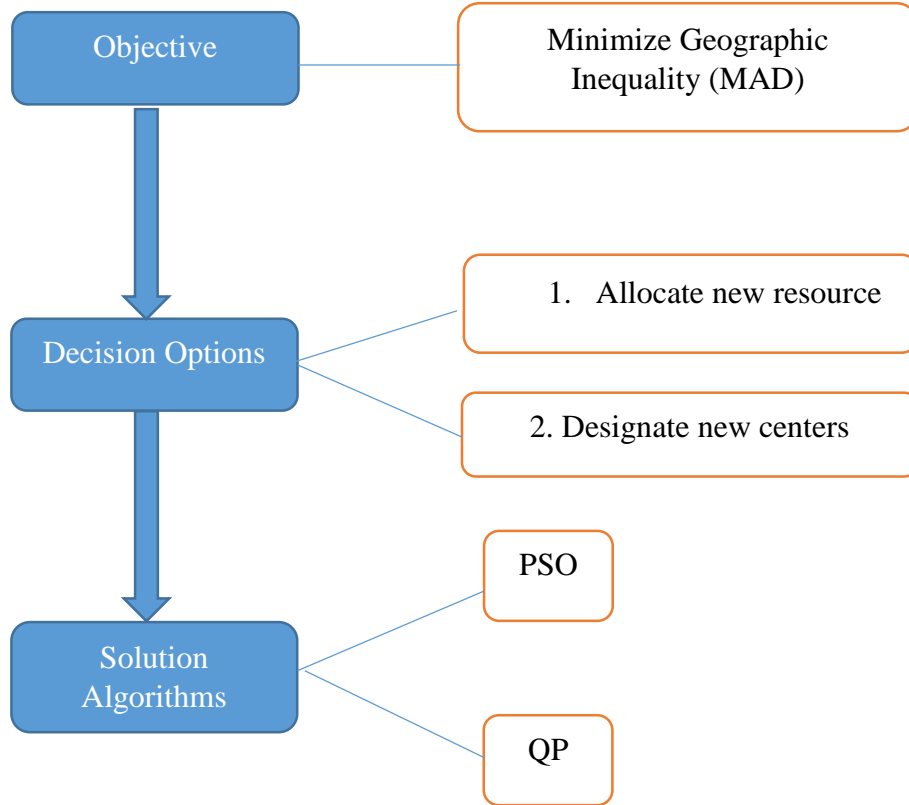
Figure 12 Formulating Planning Problems for Minimizing Geographic Inequality

For decision option 2, new NCI Cancer Centers will raise the total supply

capacity, and the objective function is updated to:

$$\min \sum_{i=1}^{m} (\sum_{j=1}^{n} \frac{B_j X_j f(d_{ij})}{\sum_{i=1}^{m} D_j f(d_{ij})} - (\sum_{j=1}^{n} S_j X_j)/D)^2 \qquad (7)$$

The constraints are:

$$\sum_{j=1}^{n} x_j = m + X , \qquad (8)$$

where *m* (=58) is the number of original NCI-CCs; *n* (=58+243=301) is the number of original NCI-CCs and AMCs; *X* is the number of new NCI-CCs to be designated; $x_j = 1$ (fixed) for the existing NCI-CCs, and $x_j = 1, 0$ (to be solved) for the other AMCs.

Technical details for solving the above optimization problems are presented in Chapter 8.

6.2 Results and Discussion

To set up a baseline for comparison, we begin with Scenario A0 by removing the constraint in Equation (6) that caps the increased bed count of each NCI-CC by 25%. For example, the decision option is to allocate 2500 beds (i.e., about 5 times the average staffed beds of NCI-CCs) to the existing 58 NCI Centers in order to minimize the inequality in accessibility at the *census tract* level. The result is shown in Figure 12.

As shown in Figure 13, most of the additional 2500 beds are allocated to NCI-CCs in southern California, New England, and Indiana. This indicates that the optimal allocation of resources for the purpose of equality is far from a trivial solution that could be obtained by simply reading a map or tabulating cancer care providers. In this case, some of the most needy regions (that is, those receiving additional services) appear to already have multiple facilities on site or nearly. However, they are also major urban areas with a lot more demand. Table 6 shows that the additional beds also improve the overall accessibility (a larger mean accessibility score of 11.19 than the mean value 11.12 in the existing condition) by increasing the total supply capacity, as discussed previously. To normalize the measure of the dispersion of the accessibility distribution across census tracts, the coefficient of variation is computed to evaluate the impact of optimization. The coefficient of variation is reduced from the current 0.5630 to 0.5611.

44
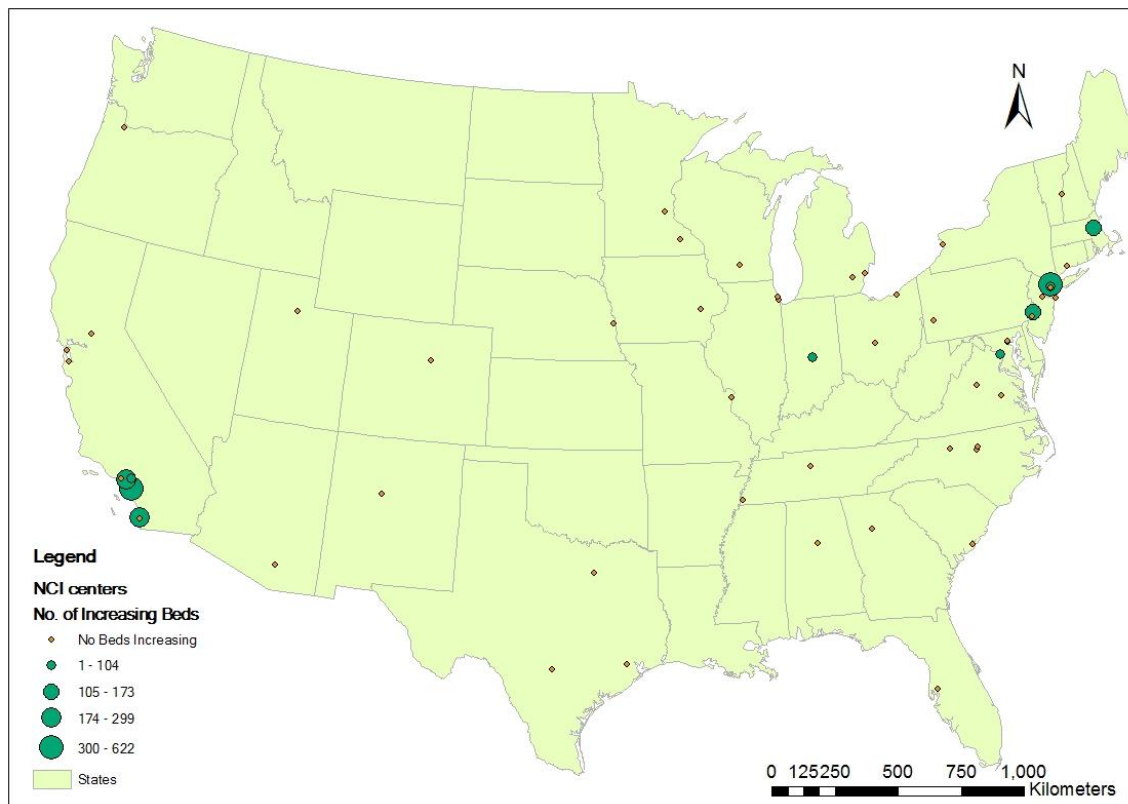
Figure 13 Allocation of 2500 additional beds without caps at the census tract level

From this case, allocating 2500 additional staffed beds, some hospitals received more than 500 beds, doubling their sizes and raising the concern of feasibility in practice. For example, City of Hope Comprehensive Cancer Center, Beckman Research Institute in Los Angeles would receive 622 new staffed beds, while its original number of beds is 185. That is over 3 times from the original capacity. The Scripps Green Hospital in La Jolla, CA, would receive 287 new beds; it has 173 now. That is more than double the size. As we mentioned before, the number of staffed beds is the easiest way to judge the capacity of an NCI Cancer Center. Doubling or tripling the number of beds means that a great increase in supply. It seems more feasible for a cancer center to increase its capacity by a small number, like 10% or 25%.

Table 7 Basic statistics for accessibility before and after optimization of geographic disparity

Scenarios	Mean	St. Dev.	Cof. Var.
Census tract level			
Existing condition	11.12	6.26	0.5630
A0. Adding 2500 beds	11.19	6.28	0.5611
A1. Adding 2500 beds with 25% cap	11.19	6.29	0.5619
A2. Designating 5 new centers	12.01	6.70	0.5576
County level			
Existing condition	7.63	4.08	0.5347
B1. Adding 2500 beds with 25% cap	7.69	4.10	0.5331
B2. Designating 5 new centers	8.24	4.37	0.5299

At the census tract level, Scenario A1 is now designed by imposing a new constraint that the increasing percentage in each center cannot exceed 25% of its existing capacity in bed count (as in Equation (6)). The result is shown in Figure 14, and also Table 7. With the additional constraint of an expansion cap, the additional beds are now allocated to a wider range of facilities across the U.S. (note the absence of hospitals with more than 200 increased beds). From Table 6, the resulting coefficient of variation is inflated slightly from 0.5611 to 0.5619. As an additional constraint is imposed, the objective function's minimal value is raised as expected.

Scenario B1 is the corresponding scenario at the county level by allocating 2500 additional beds with 25% expansion cap. The result is shown in Figure 15 and also Table 7. The result is largely consistent with that at the census tract level. A closer look reveals subtle differences in some regions.

Scenarios A2 and B2 minimize the same objective function by designating five new centers from the existing AMCs at the census tract level and the county level, respectively.

The results are identical as the same AMCs are selected to be designated as new NCI-CCs (shown in one map as in Figure 16 and Table 8). Statistics are reported in Table 7. Among the five hospitals, 1 is put in Texas filling the void in the middle of the heavily urbanized Dallas-San Antonio-Houston triangle, 1 is put in Mississippi, and 1is put in southern Alabama to serve the populous Gulf coastal area. There is also 1 placed in Denver to enhance the service capacity of the heartland, and 1 was in Rhode Island to fix the shortage in the northeast corner. Yet, this scenario exerts the highest impact on equalizing the accessibility by reducing the coefficient of variation by 0.948% to the lowest value 0.5576 at the tract level. In other words, the most cost-efficient policy in reducing the geographic disparities in accessibility of NCI Cancer Centers is to go beyond the allocation of resources on the existing list and establish new centers.
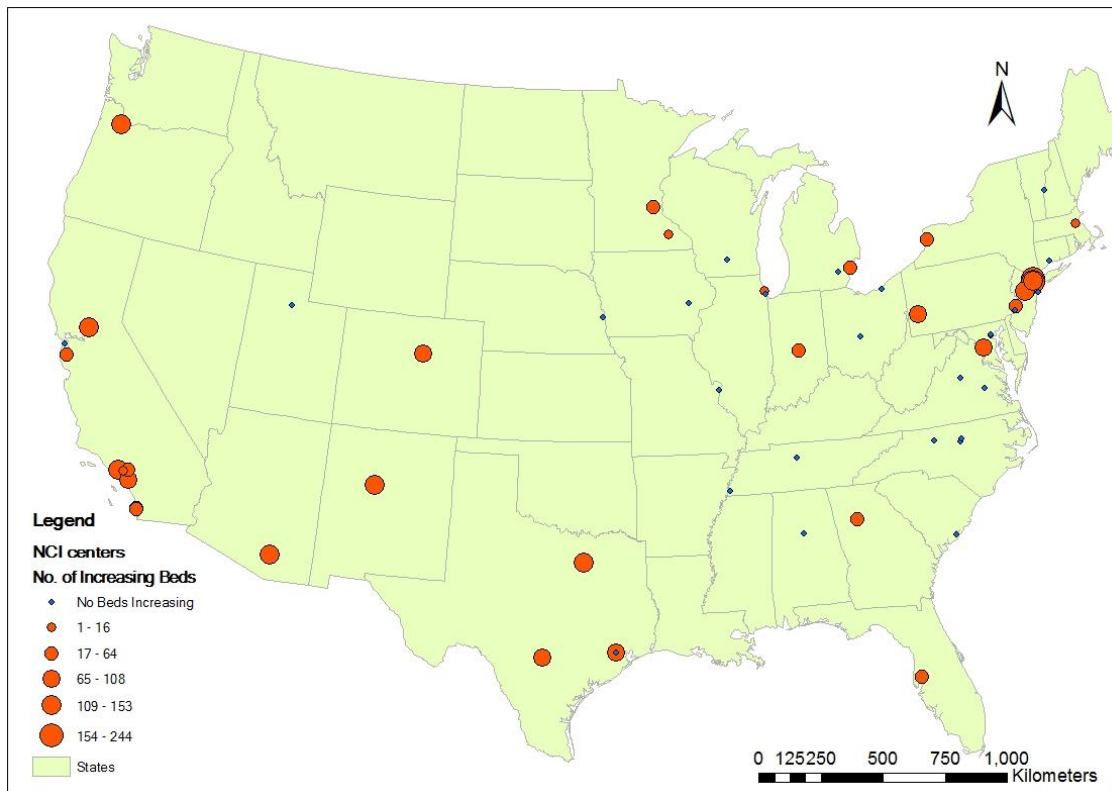
Figure 14 Allocation of 2500 additional beds with 25% expansion cap at the census tract level

47
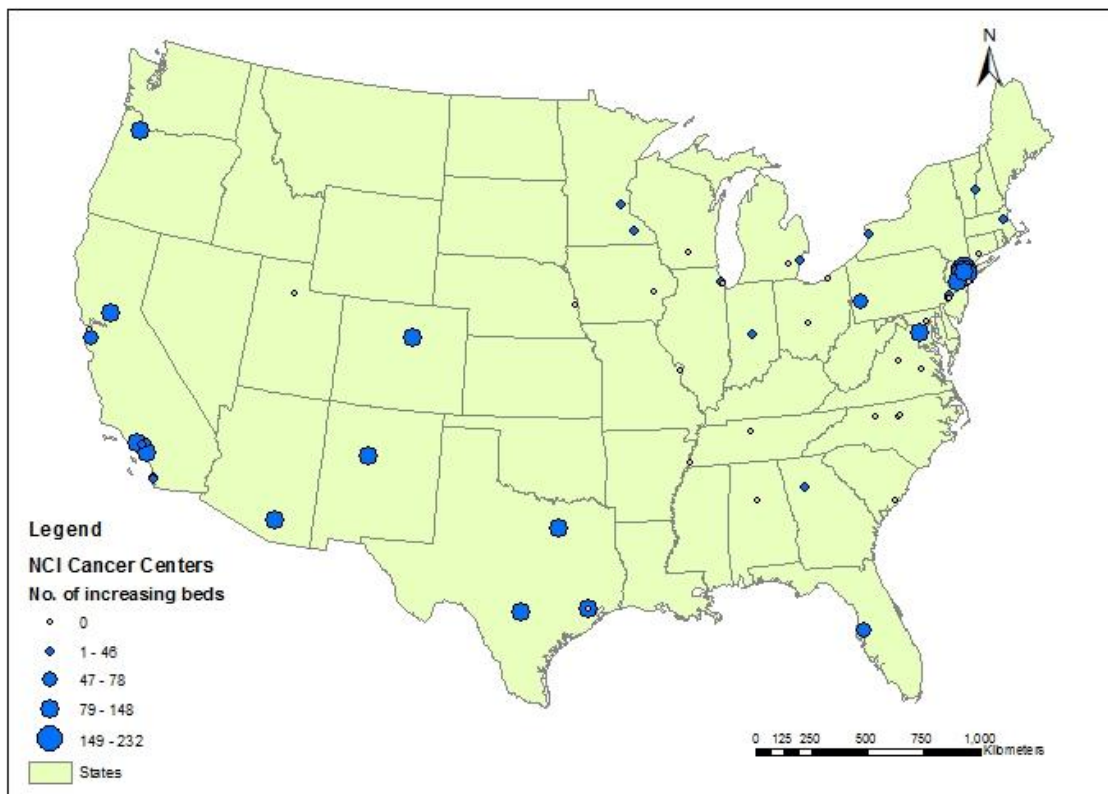
Figure 15 Allocation of 2500 additional beds with 25% expansion cap at the county level

A recent paper by Delmelle et al. (2014) focuses the decision choices between increasing capacity of existing facilities and adding new facilities in a school location. Our study suggests that adding new facilities is a more favorable policy option because it significantly reduces the inequality of accessibility. One likely reason is that the new facilities reduce travel time for patients in resource-deprived areas in addition to the added capacity of the service supply.
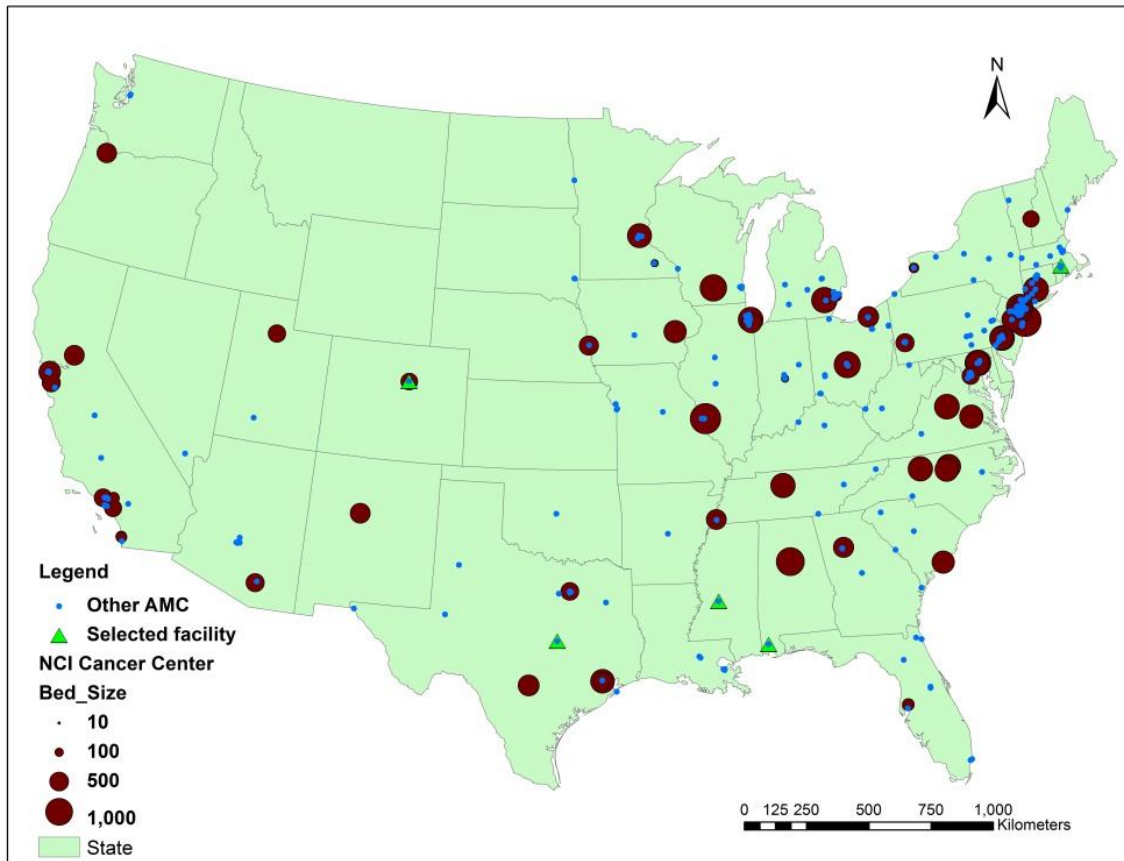
Figure 16 Designating five new NCI Cancer Centers at the county (tract) level

Table 8 Five new designated NCI Cancer Centers for minimal geographic disparity

Hospital	Address	No. of beds
University of South Alabama Medical Center	2451 Fillingim St, Mobile, AL 36617	131
Scott & White Hospital	2401 S 31st St, Temple, TX 76508	116
Methodist Rehabilitation Center	1350 E Woodrow Wilson Blvd, Jackson, MS 39216	124
National Jewish Health	1400 Jackson St, Denver, CO 80826	24
Memorial Hospital of Rhode Island	111 Brewster St, Pawtucket, RI 02860	150
Total		545

49

Chapter 7 Optimization for Minimal Racial Disparity in Spatial Accessibility

Chapter 5 shows the disparity in spatial accessibility to NCI Cancer Centers across different racial-ethnic groups at the county and tract levels. This chapter formulates and solves the planning problem of minimizing such a disparity. According to our knowledge, it is the first attempt to propose an optimization approach to reduction of racial disparity. For that reason, the study is exploratory with an emphasis on the methodological development.

7.1 Formulating the Planning Problems

The formulation of optimization problems for minimal racial disparity shares much similarity with that for minimal geographic disparity. Here only the differences are highlighted in this chapter.

This research proposes a new way to formulate the objective function of minimal racial equality by minimizing the gap between the weighted averages of accessibility between two groups. For example, the objective function is written as

$$\min(\sum_{i=1}^{m} W_i A_i / W - \sum_{i=1}^{m} M_i A_i / M)^2 \tag{9}$$

where W_i, the estimated number of the white patients (demand) at a location (e.g., census tract) i, is constant; M_i, the estimated number of the minority patients at location i is constant; W and M are total white patients and total minority patients in the study area; Ai is the accessibility index calculated at location i; and m is the number of areas (e.g., census tracts).

The constraints and decision options are the same as those in optimization problems of minimizing geographic disparity as discussed in Chapter 6. Decision option 1 allocates

new resources (i.e., additional beds), and decision option 2 designates new centers. Once again, the analysis will be conducted at both the county and census tract levels.

Specifically, the objective function for decision option 1 is updated to:

$$\min(\sum_{i=1}^{m} W_i(\sum_{j=1}^{n}[\frac{(B_j + X_j)f(d_{ij})}{(\sum_{k=1}^{m} D_k f(d_{kj}))}])/W - \sum_{i=1}^{m} M_i(\sum_{j=1}^{n}[\frac{(B_j + X_j)f(d_{ij})}{(\sum_{k=1}^{m} D_k f(d_{kj}))}])/M)^2 \quad (10)$$

With the same constraints as presented in Chapter 6.

For decision option 2 of designating a given number of new NCI Cancer Centers, the objective function is revised to:

$$\min(\sum_{i=1}^{m} W_i(\sum_{j=1}^{n}[\frac{(S_j X_j)f(d_{ij})}{(\sum_{k=1}^{m} D_k f(d_{kj}))}])/W - \sum_{i=1}^{m} M_i(\sum_{j=1}^{n}[\frac{(S_j X_j)f(d_{ij})}{(\sum_{k=1}^{m} D_k f(d_{kj}))}])/M)^2 \quad (11)$$

The constraints remain the same as discussed in Chapter 6.

Technical details for implementing the optimization problems and solving them by various algorithms are discussed in Chapter 8.

7.2 Results and Discussion

We examine the scenarios for minimizing the racial disparity in accessibility between Whites and each minority (Black, Hispanic, Asian), and also between Whites and non-White minorities. The pairings form four different objective functions. Similarly to the planning scenarios proposed in Chapter 6, we use the same two decision options such as (1) allocating new resources (e.g., 2500 extra beds) with a 25% expansion cap to each existing NCI-CC, and (2) designating a given number (e.g., five) new NCI-CCs. The intersection of four different objective functions and two decision options forms 10 optimization problems. Since the studies are conducted at both the county and census tract

levels, we have solved 20 optimization problems. The statistics for the results are presented in Table 9.

Table 9 Basic statistics for accessibility before and after optimization of racial disparity

Variables	Current Accessibility	Allocating 2500 beds	Designating 5 centers
County level			
All population			
Average	11.14	11.95	12.01
Standard deviation	4.08	4.30	4.46
Race			
White	10.59	11.05	11.07
Black	13.00	13.03	13.03
Hispanic	11.43	11.45	11.45
Asian	13.90	13.91	13.92
Minorities	12.36	12.38	12.39
Tract level			
All population			
Average	11.14	11.95	12.01
Standard deviation	6.26	6.60	6.85
Race			
White	10.58	11.04	11.08
Black	12.99	13.03	13.04
Hispanic	11.51	11.52	11.52
Asian	13.92	13.94	13.95
Minorities	12.40	12.41	12.41

Due to the large number (20) of planning scenarios, we only present the results of four optimization problems for minimizing the disparity between Whites and non-White Minorities in maps (i.e., two decision options, each at two levels). Figures 16 and 17 show the results of allocating 2500 beds with 25% expansion cap at the county and census tract

levels, respectively. Since the solutions for designating five new centers are the same at both the county and census tract levels, only one map (Figure 18) is needed to show the results. These five new centers are listed in Table 10.
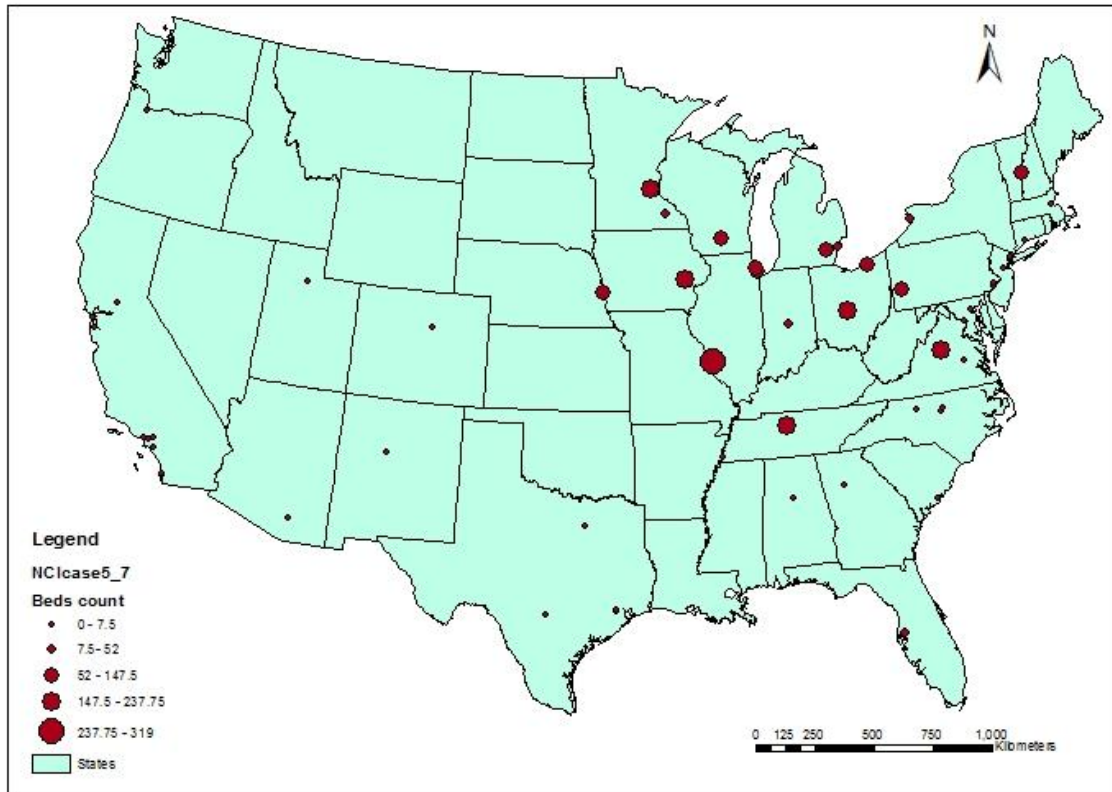
Figure 17 Allocation of 2500 additional beds towards maximum racial equality at the county level

As shown in Figures 16 and 17, the results at the county and tract levels are highly consistent. The cancer centers that receive new beds are mostly located in the middle northern regions (MN, WI, IL, OH, MI, VA and west of NY and PA, etc.). One NCI Cancer Center in TN also is allocated new beds at both the county and tract levels. Different from the county-level result, a center in Tampa, FL also receives increased beds. These areas have larger percentages of Whites. By increasing the bed sizes for centers in these areas, accessibility increases and helps improve the overall accessibility for Whites than for minorities. As a result, the gap between Whites and minorities is reduced. As reported in

Table 8, the weighted average accessibility for White people increases from 10.59 to 11.05 at the county level, and from 10.58 to 11.04 at the tract level. In the meantime, accessibility for the minorities also increases, but by very little, from 12.36 to 12.38 at the county level and from 12.40 to 12.42 at the tract level. The deviation coefficient of accessibility decreases 1.16% at the county level and 1.17% at the tract level.
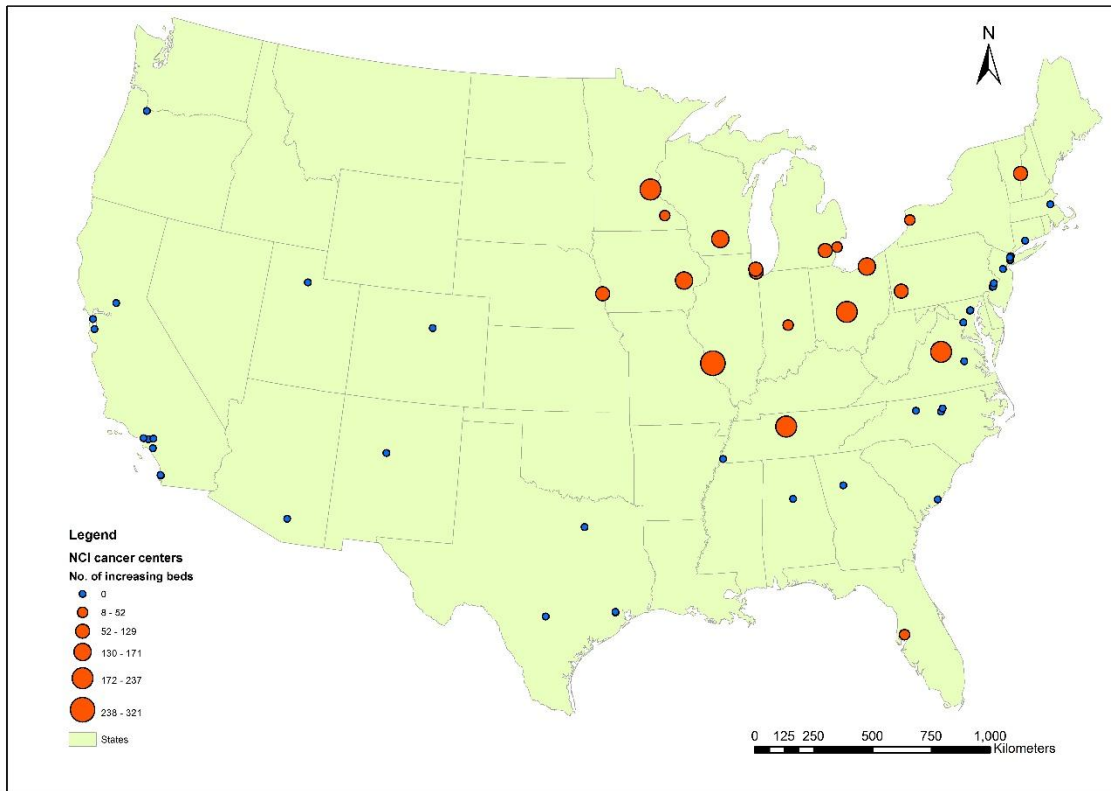
Figure 18 Allocation of 2500 additional beds towards maximum racial equality at the census tract level

As shown in Figure 18 and Table 10, the five newly designated NCI Cancer Centers are selected in order to minimize racial disparity between Whites and minorities. The results are identical at the county and tract levels. Two centers in West Virginia and one center in Kentucky fill the void in the central east of the study area, and two AMCs in Maine and Vermont are selected for new designations to meet the shortage in the northeast

corner. Table 7 shows that this scenario also improves racial equality of accessibility. The weighted average accessibility for Whites improves by 0.48 to 11.07 at the county level, and increases by 0.50 to 11.08 at the tract level. The weighted average accessibility scores for individual minority group and the combined category "minorities" have changed very little after the optimization. As a result, the racial disparities are reduced.
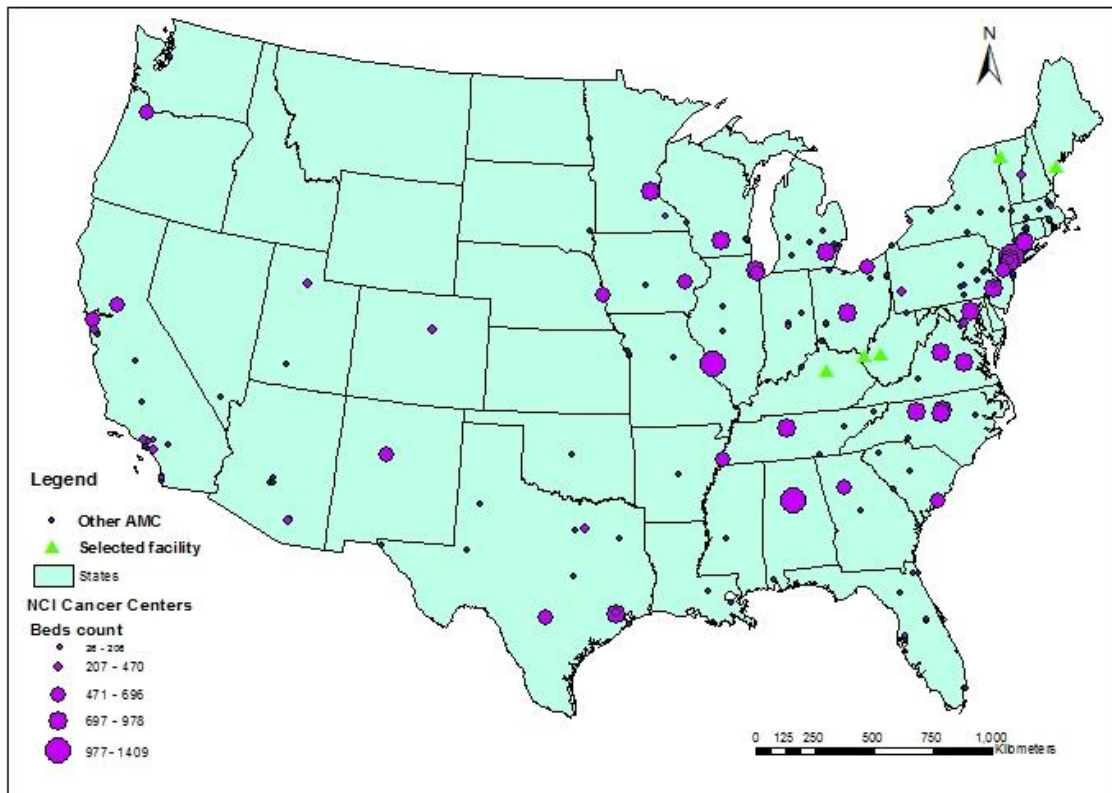
Figure 19 Designating five new NCI Cancer Centers towards maximum racial equality at the county (census tract) level

Table 10 Five new designated NCI Cancer Centers for minimal geographic disparity

Hospital	Address	No. of beds
UK Albert B. Chandler Hospital	800 Rose Street, Lexington, KY, 40536	643
Charleston Area Medical Center General Hospital	501 Morris Street, Charleston, WV, 25326	313
Cabell Huntington Hospital	1340 Hal Greer Boulevard, Huntington, WV, 25701	509
Maine Medical Center	22 Bramhall Street, Portland, ME, 4102	644
Fletcher Allen Health Care - Medical Center Campus	111 Colchester Avenue, Burlington, VT, 5401	559

Chapter 8 Solving the Optimization Problems by QP and PSO

This chapter illustrates some technical details on how the optimization problems defined in Chapters 6 and 7 are solved, and compares the approaches to make some recommendation.

There is a rich body of methods in the literature, broadly grouped under the umbrella of operational research. Some methods seek the problem's mathematical solution(s) as they solve the problem mathematically and prove that the solution(s) is the only optimum (are either the exhaustive list of optimums or nonexistent). Some are heuristic, and most use some computational algorithms to find the approximate solution(s) and often cannot guarantee whether the answer is the true optimum. The former is more desirable in terms of the answer's accuracy, but may not be feasible for all problems. The latter is usually more practical and computationally efficient.

This chapter uses two methods to illustrate how the problems are solved: the Quadratic programming (QP) represents a mathematical optimization approach, and the Particle Swarm Optimization (PSO) is a heuristic approach.

8.1 The Quadratic Programming (QP) Method

For the optimization problem of minimal geographic disparity, its objective function in Equation (3) is rewritten here for convenience such as:

$$\text{Min} \sum_{i=1}^{m}[\sum_{j=1}^{n}\frac{S_j f(d_{ij})}{\sum_{k=1}^{m}D_k f(d_{kj})} - a]^2$$

It is a quadratic function. The constraints in Equations (4) and (5) or in Equation (8) are linear. Therefore, it is a classic quadratic programming problem (Nocedal and Wright, 2006).

In matrix notation, we introduce two matrices **q** and **G** defined as

$$\mathbf{q} = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ q_{m1} & q_{m2} & \cdots & q_{mn} \end{pmatrix}, \qquad \mathbf{G} = \begin{pmatrix} G_1 & 0 & \cdots & 0 \\ 0 & G_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & G_n \end{pmatrix}$$

where

$$q_{ij} = f(d_{ij})$$

and

$$G_j = 1/[\sum_{k=1}^{m} D_k f(d_{kj})].$$

We define matrix **P** as the product of **q** and **G**, i.e.,

$$\mathbf{P} = \mathbf{qG}$$

Using **S** to denote the supply vector ($\mathbf{S} = (S_1 \quad S_2 \quad \cdots \quad S_n)^T$) and **a** to represent

the constant vector of the average accessibility ($\mathbf{a} = (a \quad a \quad \cdots \quad a)^T$), we can write the

objective function in matrix form

$$(\mathbf{PS} - \mathbf{a})^\mathbf{T}(\mathbf{PS} - \mathbf{a}) = 2[(1/2)\mathbf{S}^\mathbf{T}(\mathbf{P}^\mathbf{T}\mathbf{P})\mathbf{S} - \mathbf{a}^\mathbf{T}\mathbf{PS}] + \mathbf{a}^\mathbf{T}\mathbf{a} \tag{12}$$

The quadratic programming (QP) problem in standard form is:

$$\min_{x} \left\{ \tfrac{1}{2} x^T H x + f^T x \right\} \tag{13}$$

where x and f are column vectors with n elements, x^T and f^T are their vector transposes,

and H is a symmetric $n \times n$ matrix.

Comparing our problem in matrix notation in Equation (12) to the standard QP form

in Equation (13) and dropping the constant term $\mathbf{a}^\mathbf{T}\mathbf{a}$ and the multiplier 2, the matrices

in standard QP are defined as

58

$$\mathbf{H} = \mathbf{P}^{\mathbf{T}}\mathbf{P}$$

$$\mathbf{f}^{\mathbf{T}} = -\mathbf{a}^{\mathbf{T}}\mathbf{P} \rhd \mathbf{f} = -\mathbf{P}^{\mathbf{T}}\mathbf{a}$$

The constraints are linear, and their definitions in matrix form are straightforward:

$$\mathbf{Cx} \leq \mathbf{b}$$

$$\mathbf{Ex} = \mathbf{d}$$

For the optimization problem of minimal racial disparity, its objective function in

Equation (12) is also rewritten here for convenience:

$$\min(\sum_{i=1}^{m} W_i A_i / W - \sum_{i=1}^{m} M_i A_i / M)^2$$

or simplified as:

$$\min(\sum_{i=1}^{m} ((\frac{W_i}{W} - \frac{M_i}{M}) \ A_i))^2$$

As $\frac{W_i}{W} - \frac{M_i}{M}$ are the constant terms, we define a vector **R** as:

$$\mathbf{R} = [(\frac{W_1}{W} - \frac{M_1}{M}), (\frac{W_2}{W} - \frac{M_2}{M}), ..., (\frac{W_i}{W} - \frac{M_i}{M})]$$

Applying the same matrix notations **q**, **G** and **S** as in Equation (12) here, we

define **J = RqG.** In matrix form, the objective function becomes

$$\underset{\text{Min}}{} (RA)^2 = (RqGS)^2 = \mathbf{(JS)^2}$$

It can be transformed to

$$\underset{\text{Min}}{} (JS)^{\mathbf{T}}(JS) = S^{\mathbf{T}}J^{\mathbf{T}}JS$$

The above formulation once again fits the description of quadratic programming

(QP). Matching its standard form in Equation (10), the related matrices are defined as

$$\mathbf{H} = 2J^{\mathbf{T}}J$$

$$\mathbf{f} = 0$$

There are various free and open source programs to solve the QP problem (e.g., www.numerical.rl.ac.uk/qp/qp.html). This study uses Matlab R2009a to solve decision option 1 which is to add new resource, in particular its "quadprog" routine, because of its flexibility in coding large matrices and reliability (www.mathworks.com). Lingo 13.0 is used to solve decision option 2 because designating new NCI-CCs contains integer constraint (www.lindo.com). See Coleman and Li (1996) among others for mathematical detail.

8.2 The Particle Swarm Optimization (PSO) Method

As stated previously, mathematic optimization such as the QP is usually inefficient and computationally intensive. This becomes a major challenge for us as our study deals with a large data set. Intelligence optimization algorithms or swarm intelligence (SI) have been developed as heuristic methods to solve optimization problems. One of the most popular algorithms is Particle Swarm Optimization (PSO).

The PSO is widely used in planning and operational research. The PSO was an evolutionary algorithm originally proposed by Eberhart and Kennedy to graphically simulate the movement of a flock of birds (Kennedy, 1995). The PSO has some the advantages of an easily understandable algorithm logic, a fast convergence speed, and its good number of fits for real number optimization. So, this method has been used in a wide range of applications (www.swarmintelligence.org) since its introduction. For additional references, see Shi (1998), Eberhart (2000), and Kim et al. (2009).

In the PSO, each particle is treated as a point in an n-dimensional bounded solution space. Any particle l has position x_l, velocity v_l, and a previous position *pbest* that was previously found to be the best solution. For all particles, let *gbest* record the position of

the global best particle. The performance of each particle is measured according to a predefined fitness function. Initially, each particle has a random position within a range of X_{min} and X_{max}, with a velocity $v_l = 0$. The particles are updated by the following formulas:

$$v_l = wv_l + c_1 r_1(pbest_i - x_l) + c_2 r_2(gbest - x_l)$$

$$x_l = x_l + v_l$$

where w is the inertia weight, and c_1 and c_2 are acceleration constants. w provides a balance between global and local explorations, while c_1 and c_2 represent the weights of the stochastic acceleration terms that pull each particle x_l toward $pbest_i$ and $gbest$. In order to avoid the velocities becoming too high and particles flying out of the usable field, it is also necessary to clamp velocities to a maximum v_l smaller than v_{max}.

The pseudo code of the PSO is shown below.

Step 1

Setting values of the control parameters of PSO:

Population size N_P, inertia w, learning factors c_1 and c_2

Step 2: Initialization

Set the generation number $k = 0$

Initialize a population of N_P individuals

Initialize velocities, v, of the particles:

Step 3

WHILE the stopping criterion is not satisfied

FOR $i = 1$ to N_P

Calculate P_{best} and G_{best}

Evaluate the fitness of particles

IF $J(d_i^{k+1}) > J(d_i^k)$ Then

$$P_{best,j}^k = d_i^{k+1}$$

ELSE $P_{best,j}^k = d_i^k$

ENDIF

$$G_{best}^k = \max(P_{best,i}^k)$$

Step 4: Update position and velocity

Calculate the velocities and positions of the particles in the following way:

$$v_i^{k+1} = w^* d_i^k + c_1^* (P_{best,i}^k, d_i^k) + c_2^* (G_{best}^k d_i^k)$$

$$d_i^{k+1} = d_i^k + v_i^{k+1}$$

END FOR

Step 5: Increase the generation count

$k = k + 1$

END WHILE

For minimal geographic disparity in Equation (3), we set the fitness function for the optimization problem as the total absolute difference between the accessibility score of each demand location and weighted average accessibility score:

$$fitness = \sum_{i=1}^{m}(D_i \sum_{j=1}^{n} \left| \frac{S_j f(d_{ij})}{V_j} - a \right|)$$

For the optimization problem for minimal racial disparity in Equation (9), we set the fitness function as the absolute sum of the gap between the weighted average accessibility of different racial groups:

62

$$fitness = \left| \sum_{i=1}^{m} ((\frac{W_i}{W} - \frac{M_i}{M}) \; A_i) \right|$$

The PSO parameters used in all functions are set as follows: $N_p = 58$, $c_1 = c_2 = 2$, $w=0.4$. The accuracy is set as 10^{-4} and the constriction factor is 0.729. The maximum evaluation for all the methods is set as 20,000.

There are various free and open source PSO toolboxes to solve PSO problems based on Matlab, R, and Lingo (e.g., http://psotoolbox.sourceforge.net). This study builds the PSO optimization program with the help of two PSO toolboxes based on Matlab 2009a (Birge, 2005; Sam, 2009).

8.3. Comparing the Performances of Solution Methods

Table 11 summarizes the eight optimization scenarios used in assessing the performance of solution approaches.

Table 11 Optimization Scenarios

Objective	Analysis unit	Decision option	Scenario index
Minimal geographic disparity	County	Allocating new resources	S1
		Designating new NCI-CCs	S2
	Census tract	Allocating new resources	S3
		Designating new NCI-CCs	S4
Minimal racial disparity (Whites vs. minorities)	County	Allocating new resources	S5
		Designating new NCI-CCs	S6
	Census tract	Allocating new resources	S7
		Designating new NCI-CCs	S8

We use three criteria to assess the performances of QP and PSO methods in solving the optimization problems. The first measure of performance is *success rate*. See Table 12. In some cases (optimization scenarios), a method cannot solve the problem after numerous attempts. For example, when the QP was used to find minimal inequality of accessibility between Whites and minorities by designating new NCI Cancer Centers at the county level, it failed to reach a solution. It was unclear whether it was a software or hardware problem. The PSO displayed a relatively high success rate, but it still frequently failed in local solutions.

The second measure is *complexity in implementation*. In our case, the QP was relatively difficult to code, particularly in representing the quadratic objective function in matrix format. We used Matlab to generate hundreds to thousands of loop lines of codes for the QP problems and imported to Lingo to solve the model. The PSO toolboxes available in Matlab and Lingo provided the user with a clear guidance to program the optimization problems, and were much easier to code.

The final measure is computation time. When successful, the PSO method solved the problems within reasonable time. The QP approach usually used a significantly higher amount of computation time (in one case, more than 1,000 hours). Note that when successful, QP and PSO yielded identical or very similar results.

Based on the above discussion, our recommendation is to first attempt the PSO method to solve an optimization problem as it uses shorter time and fairly easier to implement. When unsuccessful, one may then employ the QP method. It is also a good practice to use a small sample data set to test both methods to validate the programs before applying the methods to the full data set.

Table 12 Assessing Performances of QP and PSO in Optimization

Scenarios	Success rate of solution (%)		Complexity in implementation		Average computation time (hours)*	
County level	PSO	QP	PSO	QP	PSO	QP
S1	100	100	Low	High	1.63	11.47
S2	70	100	Low	High	3.99	72.70
S5	100	100	Low	High	7.15	130.28
S6	100	100	High	High	3.82	71.63
Tract level						
S3	80	100	Low	High	36.05	511
S4	80	100	Low	High	80.44	1144.68
S7	80	100	Low	High	44.2	810.63
S8		---	High	---	96.15	---

* Based on a PC with Intel Core(TM) i7-3770 CPU @3.40GHz and 16GB memory.

Chapter 9 Conclusions

In public health studies, spatial accessibility is widely used to assess the convenience of people seeking a type of health care. This research uses the NCI Cancer Centers (NCI-CCs) as a case study to examine both geographic and racial disparities in spatial accessibility. The weighted averaged accessibility for entire United States is 11.14 which means there are only 11.14 averaged staffed beds supplied in the highest quality cancer care service NCI-CCs for every 100,000 people. Comparing with over 340 men and 290 women suffer from cancer for every 100,000 people, huge demand for quality cancer care is left unfulfilled. The current designation for NCI-CCs focuses on the single criterion in quality of research and care in cancer prevention, diagnosis, and treatment, and does not consider any equality issue. Research indicates that disparity in health care accessibility is a major reason for the subpar health performance ranking for the U.S. given our high health care cost. This research examines possible policies for reducing disparities in spatial accessibility of NCI-CCs as a first step toward the larger goal of improving the overall health in the U.S.

Four major tasks are accomplished in this research:

1.   It employs the multiple catchment areas method to assess the geographic variability of spatial accessibility of NCI-CCs. The result shows that the accessibility generally declines with increment of travel time away from the centers, and clearly displays an urban advantage.

2.   Analysis on the racial disparity reveals that on average White have the lowest spatial accessibility, followed by Hispanic and Black, and the differences are statistically significant. This "reversed racial advantage"

seems counterintuitive, but is consistent with the finding reported in a
influential prior study.

3.   Several optimization problems are formulated to minimize disparity in spatial accessibility of NCI-CCs across geographic areas or demographic (e.g., racial-ethnic) groups. Possible decision options such as allocating new resources and designation of new centers are explored. The results demonstrate that public policy for promoting equal health care access can benefit from optimization research that is scientifically sound.

4.   Based on the results, the modifiable areal unit problem is not a major concern in most cases. The two different study levels have the similar optimization results. In the scenarios of adding new resource on existing NCI-CCs, the results of the census tract level always have more NCI-CCs receiving added staffed beds than the county level, but the numbers of added beds for them are small. In the scenarios of designating new NCI-CCs, the county level and the census tract level have the exactly same results as they pick up the same AMCs.

5.   Based on our research, it is recommended that the PSO be attempted first to solve the aforementioned optimization problems due to its computational efficiency and relative briefness in coding. When unsuccessful, the QP can be used to obtain the solutions.

Given our primary interest in geographic issues and our limited expertise, this research has a focus on the methodological issues related to spatial analysis, and is of

exploratory nature. There are several limitations in this research, each of which calls for improved work in the future.

First, the focus on a single health care system such as NCI-CCs is very limited in scope as many other hospitals outside of the system also provide quality cancer care. In addition, the spatial accessibility measure accounts for the reality of patients seeking cancer care across the borders of counties or states, but not countries. In other words, it assumes that the care of the NCI-CCs is limited to patients in the U.S. and the patients in the U.S. are also restricted to seek the care within the U.S. Future work needs to expand the scope beyond the NCI-CCs and also beyond the geographic border of the U.S.

Secondly, this research used the general population as the potential demand for the cancer care facilities by assuming each person has the same risk of becoming a cancer patient. However, in reality, the probability or risk to become a cancer patient varies greatly across different groups of people, in both geographic and demographic categories. For example, very large numbers of research prove that older people have a much higher risk of developing cancer in most different types of cancer. After age 40, the rate of different cancers (i.e. lung cancer, liver cancer, breast cancer and prostatic cancer) has an exponential growth. Different census tracts or counties contain different age components. For example, census tracts in a rural area tend to have much more older people living than the census tracts in a central urban area. As the age group data is available in the census dataset, future studies should consider different cancer rates for different demographic groups to estimate the potential demand for the NCI-CCs more accurately. The same applied to geographic areas. People living in some higher cancer risk locations might have higher cancer rates while the living environment like sun exposure, chemical

68

pollution, air pollution and radiation affects may cause more cancer. Future studies should also consider the variation of cancer rates across geographic areas.

Thirdly, in implementing the spatial accessibility measure, we used an arbitrarily defined distance decay function to characterize the patient behavior in seeking specialized cancer care. Future research needs to use data of actual hospital visits to capture the best fitting distance decay function.

Fourthly, in analysis of accessibility disparity between demographic groups, our finding of the so-called "reversed racial advantage" is preliminary and more importantly, ecological in nature. Future work may use data of individual patients to (in) validate such a preliminary observation.

Finally, the optimization problems formulated in this study are exploratory. In particularly, the proposed decision options such as allocation of new beds and designation of new centers are largely speculations. It calls for inputs from health care policy makers to design feasible strategies. Future research may also consider planning options such as relocation of patients, improving transportation networks, etc. More constraints may also be added in the optimization models to make them more realistic. On the methodological front, we also plan to experiment with other solution methods beyond the QP and PSO.

References

Angela B. Mariotto, K. Robin Yabroff, et al. (2010). "Projections of the Cost of Cancer Care in the United States: 2010–2020." Oxford Journals 103(2): 117-128.

Avis, N. E., Smith, K. W., Link, C. L., Hortobagyi, G. N., & Rivera, E. (2006). Factors associated with participation in breast cancer treatment clinical trials. Journal of Clinical Oncology, 24(12), 1860-1867.

Birge, B. Particle Swarm Optimization Toolbox. Available: http://www.mathworks.com/matlabcentral/fileexchange/7506-particle-swarm-optimization-toolbox

Celaya, M. O., Rees, J. R., Gibson, J. J., Riddle, B. L., & Greenberg, E. R. (2006). Travel distance and season of diagnosis affect treatment choices for women with early-stage breast cancer in a predominantly rural population (United States). Cancer Causes & Control, 17(6), 851-856.

Centers for Disease Control and Prevention (CDC). 2010. Deaths and mortality. http://www.cdc.gov/nchs/fastats/deaths.htm (last accessed 12-21-2013).

Church, R. L. (1999). Location modelling and GIS. Geographical information systems, 1, 293-303.

Dai, D. (2010). Black residential segregation, disparities in spatial access to health care facilities, and late-stage breast cancer diagnosis in metropolitan Detroit. Health & Place, 16(5), 1038-1052.

Delamater, P. L., Messina, J. P., Grady, S. C., WinklerPrins, V., & Shortridge, A. M. (2013). Do more hospital beds lead to higher hospitalization rates? a spatial examination of Roemer's Law. PloS one, 8(2), e54900.

Delmelle, E. M., Thill, J. C., Peeters, D., & Thomas, I. (2014). A multi-period capacitated school location problem with modular equipment and closest assignment considerations. Journal of Geographical Systems, 16(3), 263-286.

Eberhart, R. C., & Shi, Y. (2000). Comparing inertia weights and constriction factors in particle swarm optimization. In Evolutionary Computation, 2000. Proceedings of the 2000 Congress on (Vol. 1, pp. 84-88). IEEE.

Fried, C. (1976). Equality and rights in medical care. Hastings Center Report,6(1), 29-34. Culyer, A. J., & Wagstaff, A. (1993). Equity and equality in health and health care. Journal of health economics, 12(4), 431-457.

Geurs, K. T., & VAn Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: review and research directions. Journal of Transport geography, 12(2), 127-140.

Guagliardo, M. F. (2004). Spatial accessibility of primary care: concepts, methods and challenges. International journal of health geographics, 3(1), 3.

Gumpertz, M. L., Pickle, L. W., Miller, B. A., & Bell, B. S. (2006). Geographic patterns of advanced breast cancer in Los Angeles: associations with biological and sociodemographic factors (United States). Cancer Causes & Control, 17(3), 325-339.

Gu, W., Wang, X., & McGregor, S. E. (2010). Optimization of preventive health care facility locations. International Journal of Health Geographics, 9(1), 17.

Hansen, W. G. (1959). How accessibility shapes land use. Journal of the American Institute of Planners, 25(2), 73-76.

Ikram, S. Z., Y. Hu and F. Wang. 2015. Disparities in spatial accessibility of pharmacies in Baton Rouge, Louisiana. Geographical Review 105 (forthcoming)

Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., Murray, T., & Thun, M. J. (2008). Cancer statistics, 2008. CA: a cancer journal for clinicians, 58(2), 71-96.

Jemal, A., Siegel, R., Xu, J., & Ward, E. (2010). Cancer statistics, 2010. CA: a cancer journal for clinicians, 60(5), 277-300.

Jennifer L. Malin, Eric C. Schneider, et al. (2006). "Results of the National Initiative for Cancer Care Quality: How Can We Improve the Quality of Cancer Care in the United States?" Journal of Clinical Oncology 24(4): 626-634.

Kim, J. J., Lee, J. W., & Lee, J. J. (2009). Central pattern generator parameter search for a biped walking robot using nonparametric estimation based particle swarm optimization. International Journal of Control, Automation and Systems, 7(3), 447-457.

Kennedy J. and R. C. Eberhart. (1995). "Particle Swarm Optimization". Proceeding of IEEE International Conference on Neural Networks 4: 1942-1948.

Kwan, M. P. (1998). Space-time and integral measures of individual accessibility: a comparative analysis using a point-based framework. Geographical analysis, 30(3), 191-216.

Kwan, M. P. (1999). Gender and individual access to urban opportunities: a study using space–time measures. The Professional Geographer, 51(2), 210-227.

Landis, S. H., Murray, T., Bolden, S., & Wingo, P. A. (1999). Cancer statistics, 1999. CA: A cancer Journal for Clinicians, 49(1), 8-31.

Luo, W., & Wang, F. (2003). Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region.Environment and Planning B Planning and Design, (30), 865-884.

National Cancer Institute. 2008. Cancer Health Disparities. http://www.cancer.gov/cancertopics/factsheet/disparities/cancer-health-disparities (last accessed 12-21-2013).

National Cancer Institute. 2013. National Cancer Institute-Designated Cancer Centers. http://www.cancer.gov/researchandfunding/extramural/cancercenters (last accessed 12-21-2013).

National Center for Health Statistics (NCHS). 2006. NCHS Urban–Rural Classification Scheme for Counties. (http://www.cdc.gov.libezp.lib.lsu.edu/nchs/data_access/urban_rural.htm) (last accessed 01-30-2015)

Oliver, A., & Mossialos, E. (2004). Equity of access to health care: outlining the foundations for action. Journal of Epidemiology and Community Health,58(8), 655-658.

Onega, T., Duell, E. J., Shi, X., Wang, D., Demidenko, E., & Goodman, D. (2008). Geographic access to cancer care in the US. Cancer, 112(4), 909-918.

Onega, T., Duell, E. J., Shi, X., Demidenko, E., Gottlieb, D., & Goodman, D. C. (2009). Influence of NCI cancer center attendance on mortality in lung, breast, colorectal, and prostate cancer patients. Medical Care Research and Review, 66(5), 542-560.

Oster, E. (2009). Does increased access increase equality? Gender and child health investments in India. Journal of development Economics, 89(1), 62-76.

Pacheco, J. A., & Casado, S. (2005). Solving two location models with few facilities by using a hybrid heuristic: a real health resources case. Computers & Operations Research, 32(12), 3075-3091.

Páez, A., Scott, D. M., & Morency, C. (2012). Measuring accessibility: positive and normative implementations of various accessibility indicators. Journal of Transport Geography, 25, 141-153.

Perry, B., & Gesler, W. (2000). Physical access to primary health care in Andean Bolivia. Social Science & Medicine, 50(9), 1177-1188.

Sam. Another Particle Swarm Optimization Toolbox. Available: http://www.mathworks.com/matlabcentral/fileexchange/25986-another-particle-swarm-toolbox

Schrage, L. E. (1987). User's Manual for Linear, Integer, and Quadratic Programming with LINDO: Supplement. Scientific Press.

Shavandi, H., & Mahlooji, H. (2006). A fuzzy queuing location model with a genetic algorithm for congested systems. Applied Mathematics and Computation, 181(1), 440-456.

Shavandi, H., & Mahlooji, H. (2008). Fuzzy hierarchical queuing models for the location set covering problem in congested systems. Scientia Iranica, 15(3), 378-388.

Shen, Q. (1998). Spatial technologies, accessibility, and the social construction of urban space. Computers, environment and urban systems, 22(5), 447-464.

Sherri de Coronado, Margaret W. Haber, et al. (2004). "NCI Thesaurus: Using Science-Based Terminology to Integrate Cancer Research Results." Studies In Health Technology And Informatics 107 (1): 33-37.

Shi, L., & Singh, D. A. (2012). Essentials of the US health care system. Jones & Bartlett Publishers.

Shi, X., Alford-Teaster, J., Onega, T., & Wang, D. (2012). Spatial access and local demand for major cancer care facilities in the United States. Annals of the Association of American Geographers, 102(5), 1125-1134.

Tao, Z., Cheng, Y., Dai, T., & Rosenberg, M. W. (2014). Spatial optimization of residential care facility locations in Beijing, China: maximum equity in accessibility. International journal of health geographics, 13(1), 33.

U.S. Census Bureau, 2015 （https://www.census.gov/geo/reference/ua/urban-rural-2010.html)

Wang, F. (2006). Quantitative methods and applications in GIS. CRC Press.

Wang, F. (2012). Measurement, optimization, and impact of health care accessibility: a methodological review. Annals of the Association of American Geographers, 102(5), 1104-1112.

Wang, F., & Tang, Q. (2013). Planning toward equal accessibility to services: a quadratic programming approach. Environment and Planning B: Planning and Design 40(2) 195 – 212

Weibull, J. W. (1976). An axiomatic approach to the measurement of accessibility. Regional science and urban economics, 6(4), 357-379.

Xu, Y., & Wang, F. (2015). Built environment and obesity by urbanicity in the US. Health & place, 34, 19-29.

Vita

Cong Fu was born in Jianyang City, Sichuan Province, China. He went to college in 2004 and received his Bachelor of Engineering degree in Geographic Information Systems (GIS) at the China University of Geosciences, Wuhan, China, in 2008. Since the fall of 2008, he began his graduate study and earned the Master of Science degree in Cartography and GIS in the Department of Remote Sensing at Wuhan University in 2010. In the spring of 2011, he has been a Ph.D. student in the Department of Geography and Anthropology, Louisiana State University, supported by research assistantships by Department of Geography & Anthropology, Department of Parking and Transportation and Office of Facility Services. His research focuses on human geography, public health, operational research, GIS and spatial analysis.