12-31-2018

17-06 Public vs Private Transportation Network Accessibility and Maternal-Infant Health Outcomes Across the Urban-Rural Boundary in Kalamazoo County, Michigan

Kathleen Baker
Western Michigan University, kathleen.baker@wmich.edu

TRCLC 17-06
December 31, 2018

Public vs Private Transportation Network Accessibility and Maternal-Infant Health Outcomes Across the Urban-Rural Boundary in Kalamazoo County, Michigan

FINAL REPORT

Kathleen M. Baker
Department of Geography
Health Data Research, Analysis and Mapping Center
W.E. Upjohn Center for the Study of Geographical Change
Western Michigan University
Kalamazoo, MI 49008

Transportation Research Center
for Livable Communities

Western Michigan University | University of Texas at Arlington | Utah State University | Wayne State University | Tennessee State University

Technical Report
Documentation Page

1. Report No. TRCLC 17-06	**2. Government Accession No.** N/A	**3. Recipient's Catalog No.** N/A

4. Title and Subtitle Public vs Private Transportation Network Accessibility and Maternal-Infant Health Outcomes Across the Urban-Rural Boundary in Kalamazoo County, Michigan	**5. Report Date** December 31, 2018 **6. Performing Organization Code** N/A
7. Author(s) Kathleen M. Baker	**8. Performing Org. Report No.** N/A
9. Performing Organization Name and Address Western Michigan University 1903 West Michigan Avenue Kalamazoo, MI 49008	**10. Work Unit No. (TRAIS)** N/A **11. Contract No.** TRCLC 17-06
12. Sponsoring Agency Name and Address Transportation Research Center for Livable Communities (TRCLC) 1903 W. Michigan Ave., Kalamazoo, MI 49008-5316	**13. Type of Report & Period Covered** Final Report 8/15/2017 – 12/31/2018 **14. Sponsoring Agency Code** N/A

15. Supplementary Notes

16. Abstract

This project is twofold. The first part of the project deals specifically with the benefits of multimodal transportation modeling to understand community structure of public health access in a community. Maternal risk and infant outcomes are examined in Kalamazoo County, MI with respect to transportation network accessibility by public transit and private vehicle across the urban-rural continuum. Infants born to mothers just outside the urban core had a higher rate of poor outcomes. Maternal risk factors, by contrast, were associated with the accessible rural areas – areas outside the city proper, but within 30 minutes by car to services. When as much variability as possible (departure time, routes, modes, time of day) was included in the model, very detailed community structure information emerged. This structural information is not specifically causal, but differences in behaviors and use of services, as well as differences in urban poverty and rural poverty were apparent. The second part of the project considered raster-based methods to provide insights into siting intervention locations at efficient and equitable locations for repeating cases of risk – in this case, repeating cases of sexually transmitted infections. The results provide metrics for decision makers to compare intervention locations by efficiency and equity across multimodal optimization.

17. Key Words Public health, geographic information systems, accessibility modeling, multimodal	**18. Distribution Statement** No restrictions.

19. Security Classification - report Unclassified	**20. Security Classification - page** Unclassified	**21. No. of Pages** 49	**22. Price** N/A

Disclaimer

The contents of this report reflect the views of the authors, who are solely responsible for the facts and the accuracy of the information presented herein. This publication is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. This report does not necessarily reflect the official views or policies of the U.S. government, or the Transportation Research Center for Livable Communities, who assume no liability for the contents or use thereof. This report does not represent standards, specifications, or regulations.

Acknowledgments

This research was funded by the US Department of Transportation through the Transportation Research Center for Livable Communities (TRCLC), a Tier 1 University Transportation Center at Western Michigan University. This research was conducted in collaboration with the W.E. Upjohn Center for the Study of Geographical Change and the Health Data Research, Analysis and Mapping Center at Western Michigan University.

Contents

Overall Introduction: Studies in Health Accessibility 1

Project Goals 2

Relevance to Specified Themes 3

References 3

Multimodal Accessibility and Maternal-Infant Health: An Urban-Rural Continuum in Southwest Michigan 4

1. Introduction 4

2. Methods 5

2.1 Maternal-Infant Population 5

2.2 Accessibility modeling 5

2.3 Travel cost metrics 7

2.4 Statistical Methods 9

3. Results 10

3.1 Principal component analysis 10

3.2 Association with maternal risk 12

3.3 Association with infant outcomes 14

4. Discussion 15

References 19

Project 2: Pareto optimality for assessing multi-modal transportation accessibility: balancing equity and efficiency when siting interventions 22

1. Introduction 22

2. Related Work 24

3. Methods 26

3.1 Simulated Dataset of Infected Individuals 26

3.2 Modelling Accessibility 26

3.2.1 Drive Time Model... 28

3.2.2 Transit Time Model 29

3.3 Siting Intervention Centers 30

3.3.1 Bi–objective Optimization of Single Mode 30

3.3.2 Multimodal Optimization.. 32

4. Results 32

4.1 Optimization of Individual Parameters 32

4.2 Bi-optimization of Parameters 34

4.3 Multimodal Optimization.. 36

5. Discussion.. 38

References 40

List of Tables

Table 1. Assumptions, model specifications and relevant output variables of the accessibility models 9
Table 2. Summary of transit related components 11
Table 3. Factor loadings for principal components analysis of variables related to travel cost by private car to OB/GYNs in Kalamazoo County. 12
Table 4. The association between mom's race and accessibility components 14

List of Figures

Figure 1. Schematic diagram of the accessibility model used in this study 6
Figure 2. Study area map showing the OB/GYN locations with respect to major road network ... 7
Figure 3. Standard deviation of travel time to reach to OB/GYNs from respective census blocks (results shown separately for six principal components) 13
Figure 4. Conversion of vector data to raster cells. Original roads with superimposed grids (on left) are converted to a speed based cost raster surface (middle) which govern the movement through cells in the raster model (right). Conversion of vector data to raster cells. Original roads with superimposed grids (on left) are converted to a speed based cost raster surface (middle) which govern the movement through cells in the raster model (right). 27
Figure 5. (a) Drive time and (b) transit time map showing the location corresponding to (i) minimum average (DA, TA) and (ii) minimum standard deviation (Ds, Ts) of travel time required (for all individuals to reach to that location from their respective household. 33
Figure 6. The Pareto frontiers of the drive time based bi-objective optimization. 35
Figure 7.The Pareto frontiers of the transit time based bi-objective optimization. 35
Figure 8. Locations or road sections optimally accessible by (i) drive, (ii) transit and (iii) both (iv) Blow-up of multimodal accessible road sections that are equitably and efficiently optimized.... 37

Overall Introduction: Studies in Health Accessibility

Studies conducted since the establishment of the Affordable Care Act have continued to find significant differences in health care services, access and health status of residents in rural versus urban areas (Weinhold, 2014; Douthit, 2015). The largest differences concern limited access to high quality providers and scarcity of healthcare technology in rural areas (Ricketts, 2000; Hart, 2005; Douthit, 2015). Residents of rural areas have longer travel times to access basic health care screening services leading, for example, to higher rates of late stage cancer (Williams, 2015). Women without appropriate access to healthcare during particular stages of pregnancy also have poorer outcomes than those women with regular access to care (Evans and Lein, 2005).

What many urban-rural studies fail to consider is the urban-to-rural gradient in counties of mixed urban and rural populations. Such population density gradients are not an isolated phenomenon in the United States. In fact, 77 percent of US counties have been designated both urban and rural by the US Census Bureau. Public transportation routes, such as buses, often cover only portions of the county which can then be treated as either 'urban' or 'rural' county in many datasets, based on a 50% population cutoff. Specialized van services for the elderly, and the like, can be time prohibitive in rural areas of on a county whose urban population has adequate access to services.

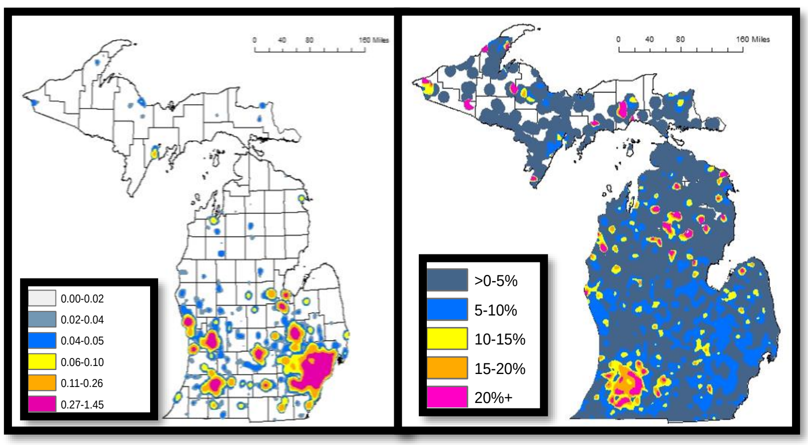


Figure 2. Density of gestational diabetes births per 10 kilometer radius (left) and gestational diabetes pregnancy rate per 10 kilometer radius, state of Michigan, 2013.

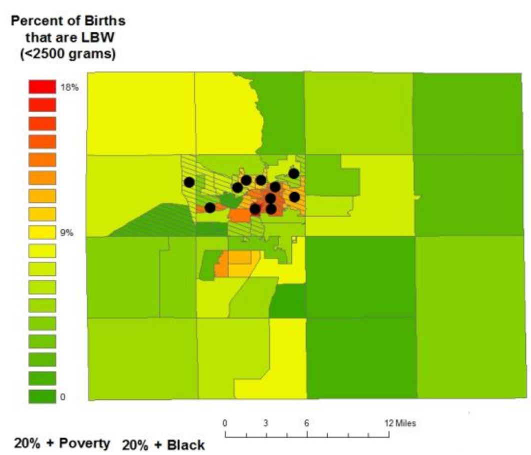


Figure 1. Spatial Distribution of Concentrated Poverty, Black Race and Low Birth Weight across Census Tracts in Kalamazoo County MI. Birth records and census tract datasets from 2010.

Kalamazoo County, MI is one such urban-rural county, with exceptionally high rates of gestational diabetes among expectant mothers and high disparities in the quality of birth outcomes among various segments of its population (Figure 2). Kalamazoo County is the only region of Michigan with both high densities and high rates of gestational diabetes (Figure 1) (MacQuillan 2017). Stratified bivariate modeling of birth outcomes (Kothari et al., 2016) revealed that race and SES independently contribute to birth outcomes, and racial congruity is associated with mitigated health outcomes in Kalamazoo (Figure 2).

The unusually high rates of gestational diabetes and disparities in birth outcomes in Kalamazoo County, MI provide substantial rationale for study of potential disparities in accessibility to maternal and infant services.

Project Goals

Goal 1: An overarching goal of this project is to improve maternal and infant health outcomes through analysis of public and private transportation network accessibility, particularly in counties with a strong urban-rural gradient. Project 1 describes results of an in-depth analysis of multimodal accessibility to maternal health services by mothers in Kalamazoo County, MI.

Goal 2: A secondary goal of this project is to use non-traditional raster methods, combined with Pareto optimality, to develop bi-objective optimization models that balance both efficiency and equity when siting intervention locations. Project 2 describes results of use of this method to examine siting an hypothetical intervention clinic for repeat sexually transmitted infection cases in Kalamazoo County, MI. STIs are one of the risk factors associated with poor maternal and infant health outcomes.

Common frameworks used in health care location-allocation studies focus on efficient allocation of services and usually disregard equity issues as well as transit accessibility. In contrast, this study proposes a heuristic approach to recommend locations that are multimodal accessible and allow equitable and efficient access to services.

As part of WMU Health Data Research, Analysis and Mapping (HDReAM) Center's efforts to provide a template for how universities and health departments can work collaboratively to analyze and disseminate information, publically available transportation data was also integrated into the Kalamazoo community's interactive mapping website. This data will enhance decision maker's understanding of accessibility as a key component in the understanding of spatial patterns in community assets, services, infrastructure, outcomes and interventions.

Relevance to Specified Themes

This research is primarily related to TRCLC themes #3 and #4. We focus on the ability of decision makers to use available, timely and accurate data when making public health decisions. Understanding accessibility to services via public and private transportation modes is critical to the design and implementation of intervention strategies. The research also examines a behaviorally and culturally specific type of individuals - women of child-bearing age and their infants - whose needs may differ from those in the general population.

References

Douthit, N., Kiv, S., Dwolatzky, T., & Biswas, S. (2015). Exposing some important barriers to health care access in the rural USA. *Public Health*, 129 (6):611-620.

Evans, W.N., & Lein, D.S. (2005). The benefits of prenatal care: evidence from the PAT bus strike. Journal of Econometrics, 125(1–2): 207-239.

Hart, L. G., Larson, E. H., & Lishner D. M. (2005). Rural definitions for health policy and research. *American Journal of Public Health*, 95(11):49–55.

Kotelchuck, M. (1994). An evaluation of the Kessner adequacy of prenatal care index and proposed adequacy of prenatal care utilization index. *American Journal of Public of Health, 84* (1994), pp. 1414–1420.

MacQuillan, E.L. (2017). Birth record analysis of gestational diabetes: Applications for intervention planning. Dissertation completed at Western Michigan University. Papers in review.

Ricketts, T.C. (2000). The changing nature of rural health care. *Annual Review of Public Health*. 21:639–57.

Weinhold, I. Gurtner, S. (2014). Understanding shortages of sufficient health care in rural areas. *Health Policy, 118*: 201-214.

Williams, F., Jeanetta, S., Obrien, D., Fresen, J. (2015). Rural-urban difference in female breast cancer diagnosis in Missouri. *Rural and Remote Health, 15* (3).

Multimodal Accessibility and Maternal-Infant Health: An Urban-Rural Continuum in Southwest Michigan

1. Introduction

Studies conducted since the establishment of the Affordable Care Act have continued to find significant differences in health care services, access, and health status of residents in rural versus urban areas (Weinhold, 2014; Douthit, 2015). The largest differences concern limited access to high quality providers and scarcity of healthcare technology in rural areas (Ricketts, 2000; Hart, 2005; Douthit, 2015). Residents of rural areas have longer travel times to access basic health care screening services, for example, leading to higher rates of late stage cancer (Williams, 2015). Women without appropriate access to healthcare during particular stages of pregnancy have poorer outcomes than women with regular access to care (Evans and Lein, 2005). What many urban-rural studies fail to consider is the urban-to-rural gradient in counties of mixed urban and rural populations. Such population density gradients are not an isolated phenomenon in the United States. Seventy-seven percent of US counties are designated both urban and rural by the US Census Bureau. Public transportation routes, such as buses, often cover only portions of counties that are treated as either 'urban' or 'rural' in federal databases. Specialized van services for the elderly, rideshare systems and the like, can be time prohibitive in rural areas of a county whose urban population has adequate access to services.

The overarching goal of this project is to improve maternal and infant health outcomes through analysis of public and private transportation network accessibility, particularly in counties with a strong urban-rural gradient. Kalamazoo County, Michigan is a mixed urban and rural county with high rates of maternal risk factors including gestational diabetes among expectant mothers (MacQuillan, 2017) and sexually transmitted infection rates nearly twice the state average (Owusu et al., 2018). In an examination of the high disparities in the quality of birth outcomes among various segments of its population, Kothari et al. (2017) found that race and socioeconomic status independently contribute to birth outcomes and neighborhood racial congruity mitigates health outcomes. In essence, these problems speak to the structural factors in the community that perpetuate inequities in health. The unusually high rates of maternal risk and disparities in birth outcomes provide a substantial rationale for focus on the county in a study of potential disparities in accessibility to maternal and infant services. In this context, key research questions include: a) what are available open source methods for quantifying transportation accessibility? b) can variability in accessibility be quantified in a meaningful way? c) is the urban-rural continuum adequately described by multimodal accessibility measures? and d) what insights can be gained into community structure through analysis of multimodal accessibility?

2. Methods

2.1 Maternal-Infant Population

Reported, confirmed cases of maternal risk factors and infant outcomes for Kalamazoo County, were accessed from 2009-2012 Michigan birth records. For each birth, the dataset included mother's home address. Batch geocoding was supplemented with extensive manual placement, resulting in an overall address match accuracy of over 90 percent of cases. Cases were assigned the census block centroid for the block in which the residential address was contained. The use of census block centroids allows for data aggregation, preserves some degree of anonymity regarding the personal address of each mother, and provides a method that for easy application across varying spatial and temporal scales. The 5,785 census blocks in the county provide an excellent sub-neighborhood scale breakdown of the region in a standard manner while introducing minimal travel time error because of their relatively small size. Outcomes selected for analysis included three maternal risk factors: sexually transmitted infection during pregnancy, gestational diabetes, and hypertension; and three birth outcomes: prematurity, low birth weight (LBW) and neonatal intensive care unit (NICU) admission. The occurrence of each type of risk factor and outcome confirmed for 2009-2012 births were aggregated for each census block.

2.2 Accessibility modeling

Accessibility to any service involves both the spatial and non-spatial aspects of travel cost. In general, travel cost is a surrogate for the relative ease by which services can be reached from a client location (Wang and Lou, 2005). Researchers widely use travel time or distance to study spatial accessibility (Apparicio et al., 2008; Ayon et al., 2018) because it is quantifiable through network modeling. However, most of these travel cost (either time or distance) based analysis primarily focus on private car and usually disregard public transit when quantifying accessibility (Martin et al., 2002, Agbenyo et al. 2017). The accessibility framework for our model is shown in Figure 1.

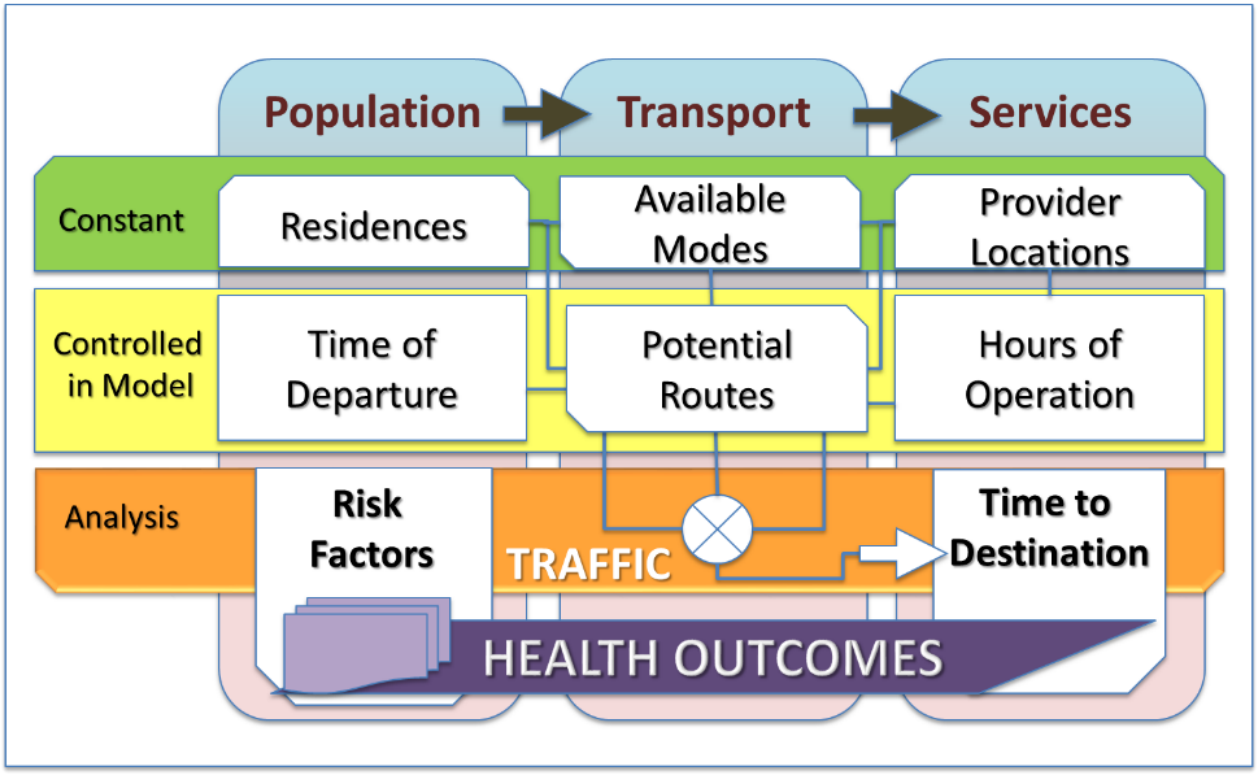


Figure 3. Schematic diagram of the accessibility model used in this study

As is typical of location-based service utilization models, a population is expected to access service providers from their primary residence through available modes of transportation. Residences of mothers who gave birth to live infants, available transportation modes and provider locations were held constant during analysis. Residence was associated with closest census block centroids, available transportation modes included riding the public bus or traveling in a private vehicle, and service providers included all obstetric and gynecological providers (OB/GYNs) in Kalamazoo County (Figure 2).

Times of departure were controlled in the model to provide estimates of variability in transit time, such that individuals were modeled to depart from each census block centroid every 10 minutes in the public transit model and every 15 minutes in the private vehicle model. Both models considered departures from 7:00 am to 4:00 pm to arrive at OB/GYNs during standard hours of operation from 8:00 am to 5:00 pm using any potential routes available as valid to the appropriate transportation mode. The model resulted in estimated times that are required to reach

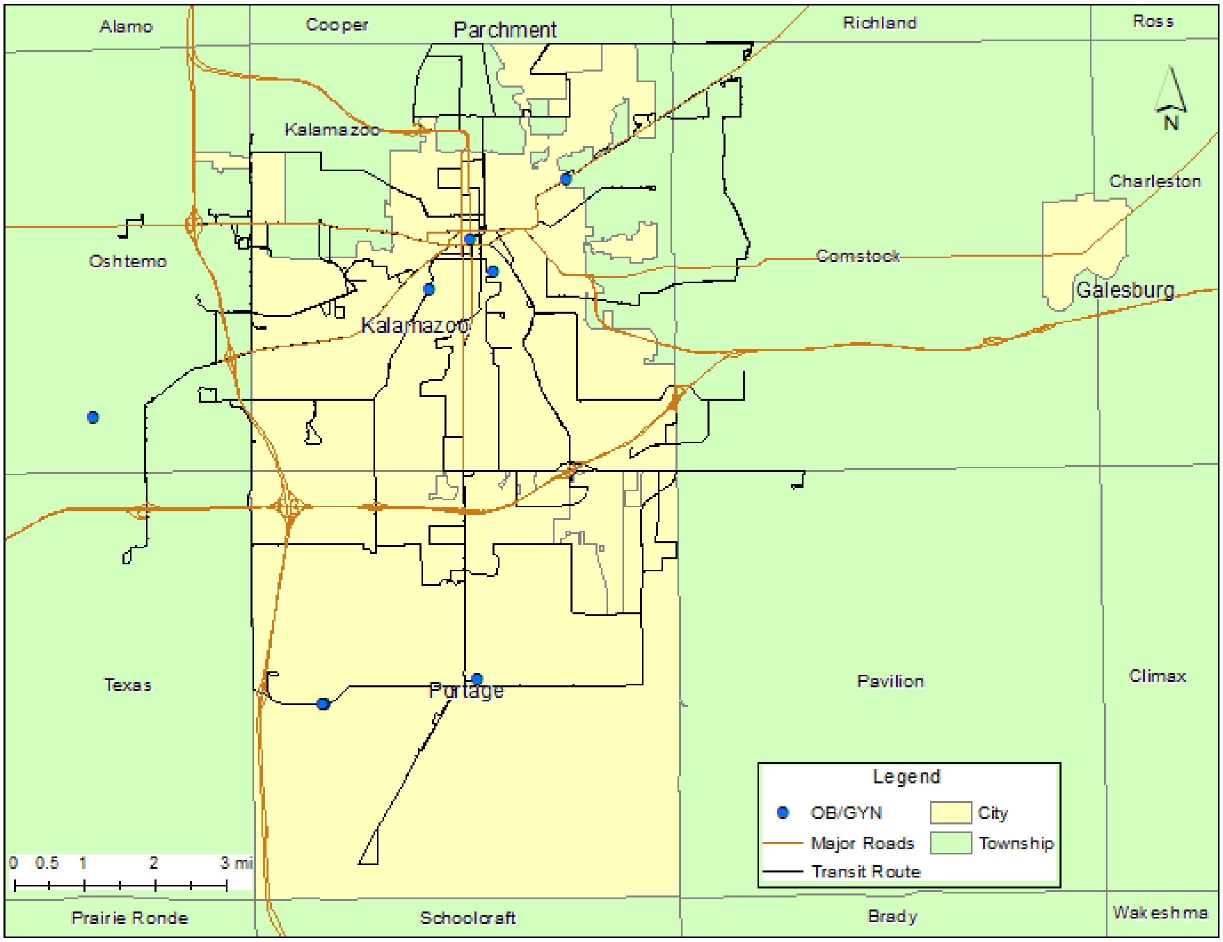


Figure 4. Study area map showing the OB/GYN locations with respect to major road network

to any available OB/GYN providers from each census block centroid within a maximum of 30 and 60 minutes for transit riders and a maximum of 15, 30 and 45 minutes for those traveling in private cars. These time thresholds were a function of county size and typical travel times associated with the major subregions. A regular weekday (April 5, 2012) was used to estimate the required travel time.

2.3 Travel cost metrics

To estimate the required travel time between unique origin to destination pairs, this study utilized OpenTripPlanner (OTP), an open source tool for multi-modal trip cost estimation. OTP exploits OpenStreetMap (OSM) for street network data and General Transit Feed Specification (GTFS) for transit modeling. It goes beyond conventional one-to-one trip prediction and enables users to estimate travel time for one-to-many and many-to-many origin-destinations. OTP does not revert to a payment schedule when a high number of pairs are necessary for analysis. The use of OTP has been successfully substituted for the traditional approach of creating arbitrary access links for walking and cycling (e.g., Delamater et al. 2012, Djurhuus et al. 2016). Due to a widespread community mapping effort, OSM is characterized by continual updates, improving its relative completeness and attribute accuracy in much of the U.S. These features lead to a wide acceptability of OSM in different domains such as Geocoding, 3D city modeling, and trip planning and analysis (Smith and Oh 2017). Another advantage is the user's ability to use past dates for trip planning.

The use of GTFS data particularly facilitates the transit time estimation by providing information pertaining to bus schedules, routes, and stop/station location. Over 800 agencies in the U.S. have stored transit specifications in a standard file format and published the data for integration particularly into dynamic mapping systems (Smith and Oh 2017). Though the GTFS data are static, a variety of applications such as the multimodal trip planning and analysis tool (Hillsman and Barbeau 2011), travel assistance (Barbeau et al. 2010), real-time transit tracking (Dailey and MacLean 2000, Ferris et al. 2010), timetable publication (Wessel and Widener 2017), mobile apps (Schweiger 2011), accessibility (Puchalsky et al. 2012), and interactive voice response (Windmiller at al. 2014) have all used these data. This study uses two of the six comma-separated text files common to the GTFS data structure (stops.txt; trip.txt) that contain information regarding passengers' pick up or drop off location and estimated travel time between stops, respectively (Smith and Oh 2017).

Travel time models for each transportation mode – private vehicle and public transit – were developed (Table 1). Exploiting a multimodal network graph, OTP identified the most efficient route at each time for each origin to destination pair (each census block centroid to each OB/GYN) and calculated the required time to traverse the corresponding network distance. During routing the private vehicle model considered one-way streets and posted speed limit when assessing efficiency; the public transit model considered only designated bus routes and schedule. Both models were constrained by standard intersection characteristics including turning time, traffic signals, and so on (Chien 2017). Each expectant mother was constrained to walking a maximum distance of 0.5 miles to, between, and from bus stops at a walking speed of 1.34 m/s or 3 mph. Application programming interface (API) tools were utilized through Python scripting to implement routing requests and batch processing. The Python scripts were also used to automate the accessibility analysis to accept both travel modes and walking limitations. A number of aggregate variables were calculated from the multiple travel times estimated by the model for each mode. Blocks that were more than 45 minutes by car from all OB/GYNs were excluded from the analysis as there is a greater chance that individuals in these areas are seeking health care from surrounding counties.

Table 1. Assumptions, model specifications and relevant output variables of the accessibility models

Model	Assumptions and model specifications	Relevant output variables
Private vehicle	shortest choice among alternative routes on street network; driving speed governed by posted limits; travel time estimated at fifteen-minute intervals over an eight-hour period (7am-4pm) and four travel time thresholds--15, 30, 45 minutes	• Number of OB/GYNs accessible • Average travel time during transit
Transit	Designated bus route with a static service schedule; limited walk speed with a distance threshold; transit time is estimated for departure at every ten-minute over an eight-hour period (7am-4pm) and two travel time thresholds—30 and 60-minutes.	• Number of OB/GYNs accessible • Average time riding public transit • Average time spent walking during transit • Standard deviation of time riding public transit • Standard deviation of time spent walking

2.4 Statistical Methods

From travel time model outputs, the average and standard deviation of destinations reached, travel time to destinations within standardized time thresholds and time spent walking were calculated for each block centroid. Factor analysis with Varimax rotation was performed using principal components as an extraction method to reduce the sixteen accessibility variables to orthogonal factors relating to the accessibility of OB/GYNs for each census block. Factor scores were then assigned to each block centroid and joined to presence/absence data for each selected maternal risk and infant outcome. Only blocks with at least one birth during the four years and with accessibility to the particular transportation network being analyzed were included in each analysis. Of the 5,785 census blocks in the county, 1,613 had births and access to both modes of transportation while 1,171 had births but no public transit access. Blocks without access to either transportation network were those associated with rivers, lakes or heavily industrialized areas. T-tests were performed to compare the factor loadings for blocks with and without 1) a mother who self-identified as non-white, maternal risk factors including 2) an STI during pregnancy, 3) gestational diabetes, 4) hypertension, and three poor birth outcomes including 5) prematurity, 6) low birth weight and 7) NICU admission.

3. Results

3.1 Principal component analysis

The principal components analysis yielded four transit specific factors, restricted to the portion of the county with transit access (Table 2), and two private vehicle factors for all blocks with at least one birth (Table 3) from 2009-2012 in the county. Transit 1 (T1) highly correlates with the number of public transit accessible destinations at 30 and 60 minute thresholds, and average time spent riding transit and walking for destinations within 30 minutes. Transit 2 (T2) highly correlates with standard deviation of the 30 minute variables: number of public transit accessible destinations, time spent riding public transit and time spent walking. Transit 3 (T3) highly correlates with time riding public transit and time spent walking for destinations within 60 minutes, and to a lesser degree with number of transit destinations within 60 minutes. Transit 4 (T4) highly correlates with standard deviation of the 30 minute variables: number of public transit accessible destinations, time spent riding public transit and time spent walking. These transit factors each had eigenvalues above one and together accounted for 79 percent of variance in the data.

Table 2. Summary of transit related components

Time Window	Variables	T1: 30 minute-accessibility	T2: 30-minute variability	T3: 60-minute accessibility	T4: 60-minute variability
30-minutes	Accessible destinations	**0.875**	0.091	-0.122	-0.208
	Average time riding public transit	**0.852**	0.266	-0.095	-0.215
	Average time spent walking	**0.901**	0.052	-0.075	-0.129
	St.Dev. of destinations	0.151	**0.832**	-0.014	-0.038
	St.Dev. of time riding public transit	0.003	**0.927**	-0.016	-0.174
	St.Dev. of time spent walking	0.295	**0.753**	-0.050	-0.088
60-minutes	Accessible destinations	**0.623**	0.387	0.502	0.136
	Average time riding public transit	-0.120	0.076	**0.889**	-0.234
	Average time spent walking	-0.103	-0.228	**0.886**	-0.068
	St.Dev. of destinations	-0.079	-0.016	-0.148	**0.915**
	St.Dev. of time riding public transit	-0.462	-0.281	-0.421	**0.641**
	St.Dev. of time spent walking	-0.387	-0.230	-0.018	**0.520**

Table 3 exhibits the results of principal component analysis for variables associated with travel by private car. Car 1 (C1) is highly correlates, positively, with destinations accessible within 15 and 30 minutes, travel time to destinations within 15 minutes and, negatively, with travel time to destinations within 45 minutes. It represents rapid access to OB/GYNs. Car 2 (C2) highly correlates with destinations accessible within 45 minutes and travel time required to reach destinations accessible in 30 minutes.

Table 3. Factor loadings for principal components analysis of variables related to travel cost by private car to OB/GYNs in Kalamazoo County.

Time Window	Variables	C1: Rapid Accessibility	C2: Accessible Rural
15-minutes	Accessible destinations	**0.874**	-0.093
	Average travel time	**0.831**	0.078
30-minutes	Accessible destinations	**0.791**	0.463
	Average travel time	-0.089	**0.895**
45-minutes	Accessible destinations	0.219	**0.795**
	Average travel time	**-0.889**	-0.096

As figure 3 shows, this factor is the most difficult to interpret. Census blocks that load highly on this factor constitute the accessible rural or areas of sprawl in the county. Blocks that load low on this factor have either extremely poor accessibility overall or quite high vehicle access at 15-30 minutes. Both car factors had eigenvalues over 1 and together accounted for 76 percent of variance in the data.

3.2 Association with maternal risk

All four of the transit factors and one of the private car factors were significantly associated with blocks in which at least one mother self-identified as non-white (Table 4). Presence of non-white mothers was associated with more transit destinations in 30 minutes (T1), higher transit travel time variability in 30 minutes (T2), fewer transit locations within 60 minutes (T3), less variability in transit at 60 minutes (T4), and more rapid access by private vehicle (C1).

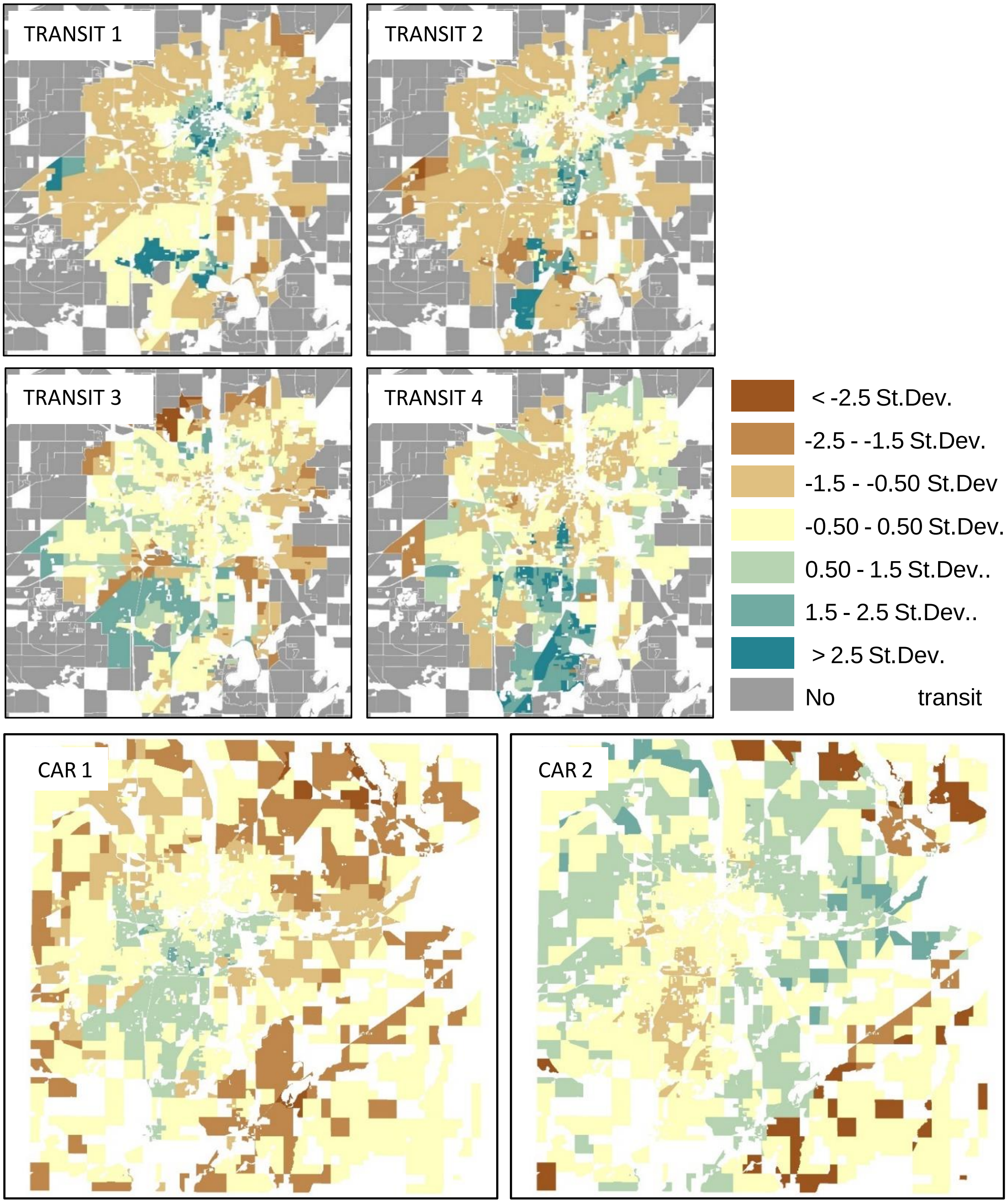


Figure 5. Standard deviation of travel time to reach to OB/GYNs from respective census blocks (results shown separately for six principal components)

Table 4. The association between mom's race and accessibility components

Components	Self-identified White/Non-white moms		
Transit	White Only (3,437) Mean (St.Dev.)	Non-white (1,821) Mean (St.Dev.)	t-value
T1	-0.095 (0.914)	0.221 (0.917)	11.902*
T2	0.061 (1.054)	0.197 (0.953)	4.757*
T3	0.104 (1.144)	0.037 (0.901)	-2.36*
T4	0.158 (1.143)	-0.096 (0.921)	-8.716*
Car	White Only (4,670)	Non-white (1,927)	t-value
C1	0.122 (0.957)	0.463 (0.6)	17.436*
C2	0.094 (0.774)	0.116 (0.421)	1.518

* statistically significant at 95% confidence interval

The only consistency among the three maternal risk variables was that loading on the accessible rural factor (C2) was higher for census blocks with at least one mom with an STI, gestational diabetes and hypertension (Table 5). Additionally, blocks with STIs were significantly associated with more transit destinations in 30 minutes (T1), higher transit ride time variability for 30 minute destinations (T2), fewer transit locations within 60 minutes (T3), and high 15-30 minute private vehicle access (C1). Blocks with gestational diabetes were significantly associated with less transit destinations in 30 minutes (T1), more transit locations within 60 minutes (T3) and variability in transit time at 60 minutes (T4). Blocks with mothers with hypertension were significantly associated with less transit locations within 60 minutes (PCA3t).

3.3 Association with infant outcomes

Table 6 shows that the only consistency among the three infant outcomes was significantly associating with blocks with higher variability in 30 minutes transit travel time (PCA2t). Additionally, blocks with prematurity were significantly associated with more transit destinations in 30 minutes (PCA1t) and higher loadings on both private vehicle access factors. Blocks with low birth weight were significantly associated with more transit destinations in 30 minutes (PCA1t), less transit accessible locations within 60 minutes (PCA3t), and more private vehicle access to OB/GYN services within 15 minutes. Blocks with an infant admitted to the NICU were significantly associated with higher loadings on the accessible rural factor with higher private vehicle access at 30 and 45-minutes intervals.

4. Discussion

Results show that detailed community structure information can emerge from the quantification of transport accessibility, in such variables as the number of destinations, time to destinations and variability in time to destinations. Even when considering only one type of health service, in this case OB/GYN offices, a thorough transportation analysis can yield a number of principal components relating to accessibility for just one county. This accessibility information has important implications for future studies of structural and/or institutional disparities, as we have shown there are significant relationships between community spatial structure and race, risk, and health outcomes. Using community transportation structure in lieu of common socio-demographic or economic variables clarifies the role of location in determining the limits to access and resources within which different segments of the populations live.

Previous research on public health in Kalamazoo County has focused on examination of socioeconomic variables, as is common in the literature. Finding patterns similar to previous research, but without the inclusion of socioeconomic variables in the model, is a critical step in understanding the spatial dimensions of disparity. Previous work on sexually transmitted infections in the county, for example, have shown a strong linkage to urbanization (Owusu et al. 2018) that is also evident in the significant relationships with transportation principal components that relate to the urbanized core of the county. Previous work on gestational diabetes has shown a relationship outside the urbanized core (Macquillan et al. 2018) and that too is clear in the significantly higher association of gestation diabetes with 60-minutes transit time, 60-minutes transit variability and the accessible rural private vehicle component.

Table 5. Statistical relations between maternal risk factors and accessibility components

Components	Sexually Transmitted Infections (STI)			Gestational Diabetes (GD)			Hypertension (HYP)		
Transit	No (4,430)[1]	Yes (828)	t-value	No (3,648)	Yes (1,610)	t-value	No (5,104)	Yes (154)	t-value
T1	-0.001 (0.93)[2]	0.095 (0.93)	2.75*	0.036 (0.93)	-0.035 (0.92)	2.58*	0.014 (0.93)	0.026 (0.93)	-0.17
T2	0.098 (1.03)	0.16 (0.98)	1.66*	0.106 (1.01)	0.114 (1.06)	-0.26	0.109 (1.01)	0.095 (0.91)	0.18
T3	0.094 (1.08)	0.013 (0.99)	-2.12*	0.05 (1.03)	0.151 (1.14)	-3.05*	0.087 (1.07)	-0.118 (0.96)	2.60*
T4	0.073 (1.09)	0.054 (1.03)	-0.47	0.046 (1.07)	0.122 (1.10)	-2.33*	0.07 (1.08)	0.064 (0.96)	0.07
Car	No (5,639)	Yes (958)	t-value	No (4,607)	Yes (1,990)	t-value	No (6,414)	Yes (183)	t-value
C1	0.202 (0.90)	0.34 (0.78)	5.00*	0.21 (0.88)	0.25 (0.89)	-1.64	0.22 (0.89)	0.288 (0.759)	-1.19
C2	0.095 (0.72)	0.133 (0.49)	2.06*	0.09 (0.73)	0.12 (0.6)	-1.81*	0.098 (0.70)	0.196 (0.409)	-3.13*

* statistically significant at 95% confidence interval

[1] number (n) of included census blocks follows no/yes designation for each maternal risk factor. [2] mean(standard deviation) are provided for each accessibility component and each maternal risk category.

Table 6. Statistical relations between infant birth outcomes and accessibility components

		Birth Time			Birth Weight			Neonatal Intensive Care		
Components	Transit	Full term n=4,251	Premature n=1,007	t-value	Normal n=4,486	Low weight n=772	t-value	None n=4,463	Admission n=795	t-value
	T1	0.004 (0.93)[1]	0.058 (0.94)	-1.66*	-0.003 (0.92)	0.117 (0.95)	-3.26*	0.008 (0.93)	0.05 (0.89)	-1.18
	T2	0.094 (1.02)	0.166 (1.03)	-2.01*	0.092 (1.03)	0.2 (1.01)	-2.71*	0.09 (1.02)	0.209 (1.03)	-3.01*
	T3	0.091 (1.08)	0.039 (1.01)	1.45	0.096 (1.09)	-0.008 (0.93)	2.78*	0.085 (1.07)	0.057 (1.00)	0.69
	T4	0.064 (1.09)	0.094 (1.04)	-0.8	0.07 (1.09)	0.066 (1.03)	0.09	0.07 (1.09)	0.071 (1.03)	-0.04
	Car	Full term n=5403	Premature n=1194	t-value	Normal -5,711	Low Birth Weight (887)	t-value	No -5,634	Yes -963	t-value
	C1	0.205 (0.90)	0.299 (0.80)	-3.61*	0.203 (0.90)	0.344 (0.77)	-4.99*	0.217 (0.89)	0.247 (0.84)	-0.99
	C2	0.093 (0.71)	0.133 (0.58)	-2.08*	0.097 (0.71)	0.125 (0.56)	-1.34	0.092 (0.70)	0.148 (0.60)	-2.57*

* statistically significant at 95% confidence interval

[1] mean and standard deviation of factor loading for appropriate census blocks

In the context of our key research questions, open source methods can be used to quantify multimodal transportation accessibility, and the travel time variability associated with those modes, in a way that reveals significant relationships between community structure and public health. However, naïve assumptions regarding travel times to public services and population health to not hold up. There is no direct correlation between travel time to service and health outcome. Instead, it becomes clear that different population segments (socioeconomically, culturally, etc.) with varying risk factors and outcomes live in different situations with respect to multimodal transportation accessibility. Quantifying the situations, then, clarifies structural disparities that can often be addressed through political and institutional will, making this type of analysis critical for long term social change that benefits public health.

References

Agbenyo, F., Marshall Nunbogu, A., & Dongzagla, A. (2017). Accessibility mapping of health facilities in rural Ghana. Journal of Transport and Health. https://doi.org/10.1016/j.jth.2017.04.010

Apparicio, P., Abdelmajid, M., Riva, M., & Shearmur, R. (2008). Comparing alternative approaches to measuring the geographical accessibility of urban health services: Distance types and aggregation-error issues. International Journal of Health Geographics. https://doi.org/10.1186/1476-072X-7-7

Ayon, B. D., Owusu, C., Oh, J. S., & Baker, K. (2018). Pareto Optimality for Assessing Multimodal Transportation Accessibility: Balancing Equity and Efficiency When Siting Interventions. Journal of Transport & Health. https://doi.org/10.1016/j.jth.2018.05.070

Barbeau, S. J., Winters, P. L., Georggi, N. L., Labrador, M. A., & Perez, R. (2010). Travel assistance device: utilising global positioning system-enabled mobile phones to aid transit riders with special needs. IET Intelligent Transport Systems. https://doi.org/10.1049/iet-its.2009.0028

Chien, R. (2017). The Components of Transit Time and Their Effects on Trip Mode Choice (Doctoral dissertation, University of Minnesota).

Dailey, D., & Maclean, S. (2000). Wireless Internet Access to Real-Time Transit Information. Transportation Research Record: Journal of the Transportation Research Board. https://doi.org/10.3141/1791-14

Delamater, P. L., Messina, J. P., Shortridge, A. M., & Grady, S. C. (2012). Measuring geographic access to health care: raster and network-based methods. International Journal of Health Geographics. https://doi.org/10.1186/1476-072X-11-15

Djurhuus, S., Sten Hansen, H., Aadahl, M., & Glümer, C. (2016). Building a multimodal network and determining individual accessibility by public transportation. Environment and Planning B: Planning and Design. https://doi.org/10.1177/0265813515602594

Douthit, N., Kiv, S., Dwolatzky, T., & Biswas, S. (2015). Exposing some important barriers to health care access in the rural USA. Public Health. https://doi.org/10.1016/j.puhe.2015.04.001

Evans, W. N., & Lien, D. S. (2005). The benefits of prenatal care: Evidence from the PAT bus strike. Journal of Econometrics. https://doi.org/10.1016/j.jeconom.2004.04.007

Ferris, B., Watkins, K., & Borning, A. (2010). OneBusAway: results from providing real-time arrival information for public transit. In Proceedings of the SIGCHI Conference on Human Computer Interaction. https://doi.org/10.1145/1753326.1753597

Hart, L. G., Larson, E. H., & Lishner, D. M. (2005). Rural definitions for health policy and research. American Journal of Public Health. https://doi.org/10.2105/AJPH.2004.042432

Hillsman, E. L., & Barbeau, S. J. (2011). Enabling cost-effective multimodal trip planners through open transit data(No. USF 21177926). National Center for Transit Research (US).

Kotelchuck, M. (1994). An evaluation of the Kessner Adequacy of Prenatal Care Index and a proposed Adequacy of Prenatal Care Utilization Index. American Journal of Public Health. https://doi.org/10.2105/AJPH.84.9.1414

Kothari, C. L., Romph, C., Bautista, T., & Lenz, D. (2017). Perinatal Periods of Risk Analysis: Disentangling Race and Socioeconomic Status to Inform a Black Infant Mortality Community Action Initiative. Maternal and Child Health Journal. https://doi.org/10.1007/s10995-017-2383-z

MacQuillan, Elizabeth L. (2017). Birth Record Analysis of Gestational Diabetes: Applications for Intervention Planning. Dissertations. 3132. https://scholarworks.wmich.edu/dissertations/3132

Martin, D., Wrigley, H., Barnett, S., & Roderick, P. (2002). Increasing the sophistication of access measurement in a rural healthcare study. Health and Place. https://doi.org/10.1016/S1353-8292(01)00031-4

Oh, J.-S., Smith, C. S., Qatra, R., Al-Akash, M., Communities, T. R. C. for L., & Technology, O. of the A. S. for R. and. (2017). Estimating and Enhancing Public Transit Accessibility for People with Mobility Limitations. Transportation Research Center for Livable Communities.

Owusu, C., Baker, K. M., Paul, R., & Curtis, A. B. (2018). Modelling individual vulnerability to sexually transmitted infections to optimise intervention strategies: analysis of surveillance data from Kalamazoo County, Michigan, USA. Sexually Transmitted Infections. http://dx.doi.org/10.1136/sextrans-2017-053350

Puchalsky, C. M., Joshi, D., & Scherr, W. (2012). Development of a Regional Forecasting Model Based on Google Transit Feed. In Transportation Research Board 91st Annual Meeting.

Ricketts, T. C. (2000). The Changing Nature of Rural Health Care. Annual Review of Public Health. https://doi.org/10.1146/annurev.publhealth.21.1.639

Schweiger, C. L. (2011). Mobile Device Technology for Real-Time Transit Information: Use and Deployment. In 18th ITS World CongressTransCoreITS AmericaERTICO-ITS EuropeITS Asia-Pacific.

Wang, F., & Luo, W. (2005). Assessing spatial and nonspatial factors for healthcare access: Towards an integrated approach to defining health professional shortage areas. Health and Place. https://doi.org/10.1016/j.healthplace.2004.02.003

Weinhold, I., & Gurtner, S. (2014). Understanding shortages of sufficient health care in rural areas. Health Policy. https://doi.org/10.1016/j.healthpol.2014.07.018

Wessel, N., & Widener, M. J. (2017). Discovering the space–time dimensions of schedule padding and delay from GTFS and real-time transit data. Journal of Geographical Systems. https://doi.org/10.1007/s10109-016-0244-8

Williams, F., Jeanetta, S., O'Brien, D., & Fresen, J. (2015). Rural-urban difference in female breast cancer diagnosis in Missouri. Rural and Remote Health. https://doi.org/10.1021/ef301981f

Project 2: Pareto optimality for assessing multi-modal transportation accessibility: balancing equity and efficiency when siting interventions

1. Introduction

A large body of research has contributed to understanding the complexities of health access (Cromley & McLafferty, 2011; McLafferty, 2003; Wang, 2012; Wang & Luo, 2005). These studies have generated knowledge on issues such as geographic accessibility, availability of services to meet needs, affordability of services provided, the organization of services to meet clients' needs and acceptability of the services provided (Cromley & McLafferty, 2011; McLafferty, 2003; Wang, 2012; Wang & Luo, 2005). Geographers have contributed enormous literature on geographic accessibility issues on when and where barriers in transportation, distance, travel time, and cost impede health services delivery (Cromley & McLafferty, 2011; McLafferty, 2003; Wang & Luo, 2005).

Geographic accessibility denotes the relative ease by which services can be reached from a client location and can include spatial and non-spatial characteristics (Cromley & McLafferty, 2011; Wang & Luo, 2005). Travel cost, in terms of distance or time, is frequently used as a proxy for geographic accessibility (Apparicio, Abdelmajid, Riva, & Shearmur, 2008; Schuurman, Fiedler, Grzybowski, & Grund, 2006; Tanser, Gijsbertsen, & Herbst, 2006). Travel time can be particularly relevant when core spatially concentrated populations are known to have a repeating pattern of infections and, thus, a shorter return interval for the use of services. For this study, we examine the travel time access of a population identified having repeat sexually transmitted infections (STIs) over three years. Previous examination of this population has shown risk of STI to be strongly associated with individual racial group and neighborhood-level low socioeconomic status (Owusu, Baker, Paul, & Curtis, 2018). In general, low-income households are greatly dependent on public transit. However, very few health care literatures consider public transit when quantifying accessibility (Mavoa, Witten, McCreanor, & O'Sullivan, 2012; Neutens, 2015).

Geographic approaches that require solving the P median problem, location set covering problem (LSCP), and maximum covering location problem (MCLP) all account for the total travel time, the number of facilities and maximize the population demand for the health facilities (Rahman & Smith, 2000; Wang, 2012). While these conventional techniques address population demands for health care facilities within a specified distance/time threshold during their measurement of geographic accessibility, they are limited in incorporating remote users (Rahman & Smith, 2000; Wang, 2012). An improved spatial accessibility measurement can offer more equitable resource configuration by paying attention to those remote users. Such a measurement technique not only aims to minimize the cumulative travel time of service users but also maximize the coverage by diminishing the gap between closest and farthest user groups (Wang & Tang, 2010). In contrast, the solution for p-median problem often is used to highlight opportunities to improve facilities in high-density population centers by minimizing end-user travel costs and maximize profits for the service providers (Drezner, 1995).

An alternative approach to ensure equality of access among the population being served in high-density population areas and remote areas is to optimize facility locations in such a way that it maximizes service coverage, minimizes travel needs of users and limit number of facilities. However, such an application should not be limited to only homogenous road network analysis in healthcare location-allocation studies where each road has the same speed-limit or a two-dimensional Euclidean plane is used to determine accessibility in terms of travel time or distance (Jia et al., 2014). This is because the transportation network with uniform speed-limit may lead to an unvaried spatial distribution of facilities whereas various speed limits presumably may produce a heterogeneous and more practical facility distribution (Jia et al., 2014).

However, location-allocation studies such as those implemented in (Gu, Wang, & McGregor, 2010; Jia et al., 2014; Mestre, Oliveira, & Barbosa-Póvoa, 2015; Mitropoulos, Mitropoulos, Giannikos, & Sissouras, 2006) that use P median or a similar technique consider some exsiting or hypothesized candidate locations to optimize. Sometimes the assumptions behind choosing candidate locations are applicable in particular situations, but that are impractical in other scenarios (Galindo & Batta, 2013). This has been demonstrated by studies that compute aggregated or weighted travel time from demand centers (e.g., centroids of census block) to a point location of a service provider, and hence disregards the detailed spatial distribution of individuals (Hewko, Smoyer-Tomic, & Hodgson, 2002; Huang & Wei, 2002; Schuurman et al., 2006). In this study, we propose a model that integrates dynamic travel time into geospatial models considering precise location of individual household along the street network. Additionally, we evaluate and predict intervention placements where the candidate locations are not pre-specified, but identified by the model.

2. Related Work

Common frameworks for solving public health intervention problems focus on efficient allocation of service centers, but the results cannot be easily adjusted to address equity issues. These methods focus on efficient allocation of service centers based on different objectives (e.g., minimal travel, minimal resources, maximal coverage), but discount health equity concerns on accessibility for different populations, utilization and service quality. For example, given a set of population centers, a *p-median* solver typically is used to choose the optimal facility site by minimizing end-user travel costs (e.g., distance, time) (Drezner, 1995). However, this method often selects locations that favor users living in high-density areas, thus perpetuating inequities in the burden of travel to such locations by remote users. *P-median* solutions also fail to address scenarios in which users do not always travel to their closest facility (Rahman & Smith, 2000). From a service point of view increase in travel cost may decrease facility usage. Recognizing that, the location set covering problem (LSCP) method recommends a minimum number of service locations such that each population center is covered by at least one facility within a given threshold (e.g., maximal service distance or time) (Shavandi & Mahlouji, 2008). However, inadequate resources may limit the number of facilities that can be maintained, regardless of the number suggested by LSCP methods (Rahman & Smith, 2000). An alternative model called the maximal covering location problem (MCLP) maximizes the coverage within a desired service distance or time threshold by locating a fixed number of facilities (Haghani, 1996; Shariff, Moin, & Omar, 2012; Verter & Lapierre, 2002).

Health equity is a multidimensional concept that focuses on addressing fairness in health services by taking into consideration social determinants of health such as household conditions, neighborhood factors (income, infrastructure) in formulation of policies and programs that benefit different populations (Braveman & Gruskin, 2003; Heiman & Artiga, 2015; Marmot, Friel, Bell, Houweling, & Taylor, 2008). In the United States, the need for a policy that incorporates health equity led to the introduction of the Affordable Care Act in 2010 (Heiman & Artiga, 2015). This paper perceives health equity through a transport geography lens, and mainly focuses on modeling geographic accessibility of health facilities that equitably incorporates different time spent to access services using different transportation modes in areas with high-risk of STIs. This approach was used in developing an equity model with an objective to minimize the accessibility gaps across all population locations by redistributing the total amount of supply among healthcare facilities (Wang & Tang, 2010). A bi-objective covering location model for locating ambulances at preexisting stations that balances efficiency in expected coverage and considers health equity by minimizing the number of uncovered demand zones have also been implemented (Chanta, Mayorga, & McLay, 2014). A similar study to improve the operational shortfalls in locations of health centers in Greece suggest the need for equitable distribution of health facilities to minimize travel distance between patients and the facilities; these studies all highlight optimal site for intervention placement on existing locations. However, these studies ignore the multimodal transportation options available to the user in the geographic accessibility modeling.

A multimodal geographic accessibility study to understand the population demand and health service locations using both car and public transportation in England developed a metric that incorporates the measurement of spatial weights (Martin, Wrigley, Barnett, & Roderick, 2002). However, weighted solutions are more appropriate to analyze aggregate level health data where for example the proportion of car ownership data can be used to create the weighted combination of travel time (Martin et al., 2002). Such single or combined travel time model may not be appropriate when/where different modes have different accessibility measures (Martin et al., 2002). Therefore, such a weighted model may lead to multimodal accessible locations which are not optimal when a particular mode is considered. This study proposes a bi-objective model to optimize the locations of health facilities which are accessible using different transportation modes to address this research gap. Specifically, the purpose of this study is to find optimal intervention locations based on transit time and drive time allowing for both equitable and efficient access to services across a multimodal transportation network. Using Pareto optimality this study develops bi-objective optimization models that minimize (i) total travel time for a target population to reach to an intervention location and (ii) the variations of travel time for repeat STI patients to reach to the locations of health facilities from a set of households

3. Methods

3.1 Simulated Dataset of Infected Individuals

Kalamazoo County, Michigan has high rates of STIs and four core areas of individuals with repeat infections and multiple types of infections as identified by Owusu et al. (2018). To protect the anonymity of individuals while simulating accurate patterns for analysis, a hypothetical set of individuals was modeled for this study by randomly placing households (n = 64) within the confines of these core areas of STIs.

3.2 Modelling Accessibility

Theoretical drive time and transit time model were developed using ArcGIS *Cost Distance* tool. ESRI's *cost distance* is a raster-based accumulated distance calculator that calculates the distance to the nearest source for each cell in the raster, based on the least-accumulative cost over a cost surface. Drive time models are typically vector based, but the raster data model allows for easier analysis across many layers, and its output is not limited to street nodes. For these reasons the raster data model was chosen for this analysis, although vector models do have the advantage of allowing for one-way streets and non-planar infrastructure that are essential in other types of analysis. The raster data model is composed of a matrix of regularly spaced square grid cells (or pixels) organized into rows and columns. In this analysis, the rows of the matrix are parallel to the X-axis and the columns to the Y-axis of the Cartesian plane in the Hotine Oblique Mercator projection system (NAD 1983, Michigan Georef). Speed-based raster surfaces, as described in more detail below, for drive time and transit time transportation scenarios, were generated and used as the input source raster to define the impedance when moving planimetrically through each cell. The relevant dataset is published by Ayon, Owusu, Oh, and Baker (2018). The Cost Distance tool utilizes the node/link cell representation common in graph theory, where the center of each cell represents a node and two adjacent nodes are connected to each other by links. Every link has an impedance (e.g., travel speed) which corresponds to the cost per unit distance for moving through the cell. The impedance value is multiplied by the cell resolution while taking into account travel direction through the cell to generate the final cost of traveling across the cell.

In traditional raster operations, cell-to-cell movement occurs either perpendicularly through or diagonally across cells..

Different researchers used different resolution to rasterize the road network. While Martin et al. (2002) used a cell size of 200 m, Tanser et al. (2006) used a raster grid of 30 m resolution. Higher resolution (i.e. smaller raster cells) helps to improve raster-based travel time estimation by decreasing the likelihood of multi-roads falling within one cell. Furthermore, reducing the cell size increase the probability of cells falling on or near the road network (Delamater, Messina, Shortridge, & Grady, 2012). Therefore, a finer resolution 25-meter raster cells are used to in this particular data model. Further reduction slows down the processing time and increase the data storage requirement and is beyond the scope of this study.

The accuracy of travel time calculation depends on the precise representation of both road segment length and travel speed. The road network database (Michigan Geographic Framework Version 14a) was acquired from the Michigan Center for Geographic Information and converted to a raster grid with cell resolution of 25m. Fig. 1 shows the hypothetical representation of converting vector road data to raster surface and assigning impedance values equivalent to speed limits to cells.

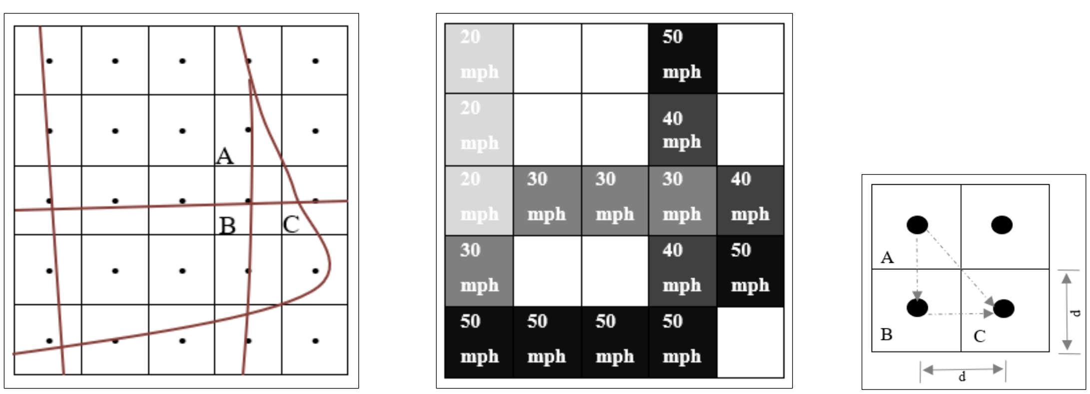


Figure 6. Conversion of vector data to raster cells. Original roads with superimposed grids (on left) are converted to a speed based cost raster surface (middle) which govern the movement through cells in the raster model (right). Conversion of vector data to raster cells. Original roads with superimposed grids (on left) are converted to a speed based cost raster surface (middle) which govern the movement through cells in the raster model (right).

The cell centroid is the hypothetical node of a particular cell. The distance (d) between two nodes is equal to the cell resolution, which is constant. Therefore, varied travel speed (e.g., S_A, S_B, S_C) were assigned to different cells to determine the time required to traverse the link. If the movement is perpendicular through raster cells, the travel time (t_{AC}) to move from cell A to C would be calculated such that

$$t_{AC} = \left(\frac{\frac{d}{2}}{S_A} + \frac{\frac{d}{2}}{S_B}\right) + \left(\frac{\frac{d}{2}}{S_B} + \frac{\frac{d}{2}}{S_C}\right) \qquad (1)$$

When moving diagonally, the travel time to move across the link would follow a direct route between the two nodes such that

$$t_{AC} = \left(\frac{\frac{\sqrt{2}}{2} * d}{S_A} + \frac{\frac{\sqrt{2}}{2} * d}{S_C}\right) \qquad (2)$$

This allows the accessibility model to create an individualized travel time-based raster surface for each at-risk household. For drive time and transit time scenarios, a full stack of travel time surfaces was analyzed to identify a set of potential intervention locations.

3.2.1 Drive Time Model

To model driving time for a personal vehicle, the travel speed assigned to each cell corresponded to the posted speed limit of the longest road segment falling inside the bounds of the cell. This study followed the hierarchical decision tree for assigning travel speed introduced by Delamater et al. (2012), using both Framework Classification Code (FCC) and National Functional Classification (NFC) as well as ownership data to assign travel speed to each road segment.

Estimating travel distance is complex, as it includes available network of streets, one-way/two-way streets, the shortest choice among alternative routes, etc. Travel time estimation becomes even more complicated because of several dynamic factors such as traffic congestion, speed limits, turning time, traffic signal and so on. The complexity is exponentially amplified when a modal split is considered.

These difficulties explain why straight-line distance is prevalent in literature. Travel times obtained by GoogleMaps are derived from independent source data and provides reasonable estimates which account for all dynamic variables (e.g., turn delay, signal time, etc.). Delamater et al. (2012) compared travel time estimates from Google MapsTM with travel time calculated using network cost distance and found that reducing the speed limit by 5mph produced results similar to those obtained by Google. Likewise, in this study, travel speed was assigned to each cell as 5 mph less than the specified by a corresponding code of a road segment. This reduced speed accounts for sub-optimal driving and traffic conditions due to congestion, stop signs, traffic lights, etc. Moreover, sample households were connected to the street network using a straight line which accounts for the driveway distance with a uniform travel speed of 10 mph.

3.2.2 Transit Time Model

The transit time model included both walking and ride time components. Theoretical walking time to the nearest bus stop was computed using Euclidean distance from each household to the nearest bus stop and background walking speeds were assigned to all cells connecting these paired locations. Because those with sexually transmitted infections are mostly between the ages of 18 and 35, a fairly brisk walking speed was assumed. If the household was within 400 m from a transit stop, walking speed was set at 4 km/h. In the U.S., 400 m or 0.25 miles is widely acceptable distance an average American will walk rather than drive (Yang, Y., & Diez-Roux, A. V. 2012). Walking speed was not changed with road infrastructure quality as in Tanser et al. (2006), but it was changed for individuals residing farther from bus stops. Distances from 401 to 800 m from bus stops are considered 'not directly connected to the bus stop', following the definition by Martin et al. (2002), and hence are assigned a background walking speed of 3 km/h as assigned in that paper. This decrease in speed also represents uncertainty in the length of most efficient and accessible walking path. To reiterate from above, our purpose is to examine the effectiveness of modeling transit in a raster data environment, but that necessarily reduces our ability to rely on traditional vector data model network concepts; thus, the generalization of walking habits with distance from bus stops. In this study, no sample households (randomly selected) were found beyond 800 m from nearest bus stop. Walking from the final bus stop to the intervention center was ignored.

Unlike car travel time estimation, the posted speed limit was not considered for bus travel time model. Kalamazoo Metro Transit's General Transit Feed Specification (GTFS) data are accessed to acquire bus schedule, routing, and bus stop information. These data are then used to compute a transit travel time for each segment of the transit route, resembling the approach followed by Mavoa et al. (2012). However, no transfer penalty was imposed. Kalamazoo Metro Transit scheduled its bus services in such a way that cross-connecting buses always meet each other at designated transfer locations. Even if a bus arrives at a transfer location before the other bus arrives, the preceding bus waits while following bus arrives. This wait time is included in the schedule and hence such ‘arrive to wait’ time are incorporated while calculating the transit time without imposing any further ‘transfer penalty’ time. Incorporation of bus schedules helps to address the limitation of ‘perfect world’ assumption for travel time estimations and provides a fair estimate to travel from one stop to another stop along the route and were cross-checked by personally traveling.

3.3 Siting Intervention Centers

Drive time and transit time raster surfaces for each simulated address of STI repeaters were created within the city area limit. These raster surfaces provide the estimated time to reach any location (j) along the road network from an individual household (i). The calculated travel time sets were then analyzed to compute the average and the standard deviation of travel times for each transportation network pixel, representing potential intervention locations.

For siting the intervention center, it is assumed that there are a finite number of potential facility locations and that demand for the facilities exists at a finite number of locations. In this study, the entire set of potential locations were represented by all the hypothetical nodes (J) of the raster cells, where $\forall j \in J$. Location modelers frequently use this assumption to solve mathematical intractability involving large-scale planar location problems (Church, Current, & Storbeck, 1991). Cells (or locations) with minimum average and the minimum standard deviation of travel time were then identified and compared against the existing intervention center location. Additionally, bi-objective optimization models were developed using Pareto optimality to optimize the potential health facility locations.

3.3.1 Bi–objective Optimization of Single Mode

For each pixel on the transportation network, the average and standard deviation travel time were used as input for bi–objective optimization models implemented in python script. Separate models were utilized for optimizing drive time and transit time-based intervention locations. The optimization was performed to find optimal health facility locations by balancing the benefit between the following two objectives.

i. Secure, efficient movement of service users by minimizing their total travel time.

ii. Ensure equity to incorporate remote users by minimizing the inherent variations in the travel time data sets.

Minimum average travel time was used as a proxy for efficiency; minimum standard deviation value of travel time was used as proxy equity. Standard deviation was chosen over simpler measures of spread, such as range, because it quantifies spread around a measure of central tendency, thus including the values of all elements in the set of modeled possibilities in the calculation. The bi-objective optimization problem was formulated as—

$$\min_{j} \mu_{T_j} = \min_{j} \sum_i \frac{t_{ij}}{I}$$

$$\min_{j} \sigma_{T_j} = \min_{j} \sqrt{\frac{\sum_I \left(t_{ij} - \mu_{T_j}\right)^2}{I-1}}$$

where t_{ij} refers to the time required to travel from a set of household locations, i to a potential intervention location j. I denotes the total number of households which correspond to the sample size and J is the set of cost raster cells which correspond to the total number of potential intervention locations.

Minimizing all related objective functions is challenging. Typically, such multi-criteria optimization does not offer a single solution, but rather suggests many alternative solutions. Pareto optimality offers a set of allocations or Pareto frontiers that are all Pareto efficient in such a manner that no objective can be improved without sacrificing at least one other objective. In this study, each point on the Pareto frontier corresponds to a location of health facility which is impossible to relocate for improving one objective without making the other criterion worse off. For example, a potential facility location A is said to (Pareto) dominate another location B, if $\mu_{T_A} \leq \mu_{T_B}$ and $\sigma_{T_A} < \sigma_{T_B}$ or vice versa., A Pareto optimal allocation results in, if no dominating solution exists.

3.3.2 Multimodal Optimization

It is tempting to argue that a single multimodal optimization, seeking to optimize both drive time and transit time simultaneously, would be practical for siting a facility. Any single model to search for coincident optimized solutions would constrain drivers of private vehicles only to the transit routes. This would occur because the solution would limit the optimal route to an intersection of acceptable paths open to both modes. This constraint is grossly unrealistic and renders the results of any such model unusable in a real situation. Therefore, drive time and transit time optimization were modeled separately throughout this analysis. Finally, the study extends to explore the coincident location(s) by analyzing optimized solutions resulted from both models. The coincidence of optimized locations is somewhat due to chance, as well as circumstances unique to a particular transportation network. In general, the area bounded by the minimum average and minimum standard deviation location for each frontier line would represent the constraints to intervention location. In this study, Pareto frontiers found from drive and transit time models were further analyzed to find the coincident geographical locations. Pareto optimality analysis yielded respective position (row and column number) of each frontier along with their associated values (i.e., average and standard deviation of travel time) so those frontiers could be mapped on the transportation network.

4. Results

4.1 Optimization of Individual Parameters

Drive time and transit time were calculated from each hypothetical STI repeating address to each raster cell on the respective transportation network. Mean and standard deviations of drive and transit times were calculated from the raster stack of individual results. Drive time and transit time cost rasters are shown in fig. 2, along with the locations corresponding to minimum average and minimum standard deviation of travel time required for all individuals to reach to that location from their respective household – which is considered as the measure of optimality.

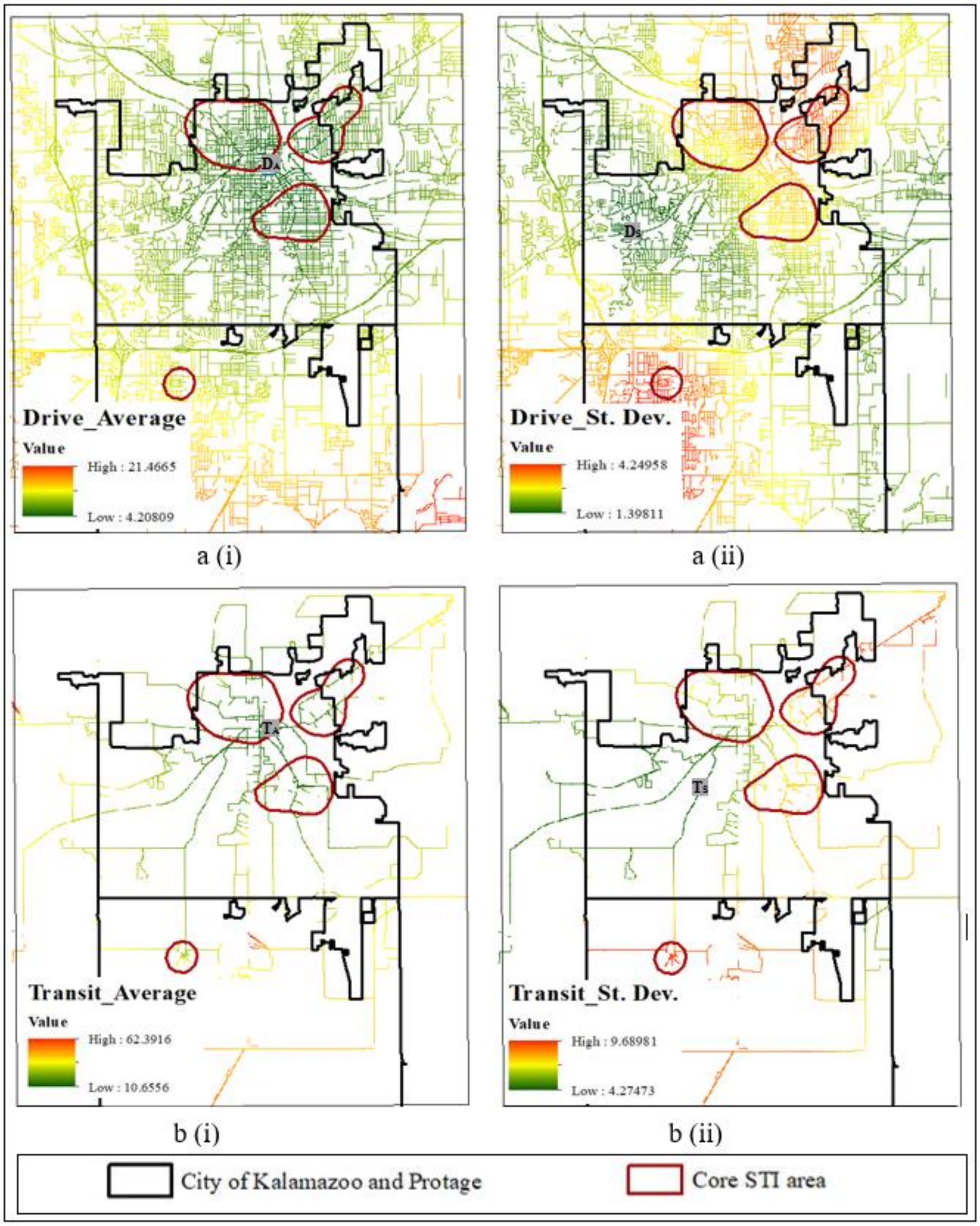


Figure 7. (a) Drive time and (b) transit time map showing the location corresponding to (i) minimum average (DA, TA) and (ii) minimum standard deviation (Ds, Ts) of travel time required (for all individuals to reach to that location from their respective household.

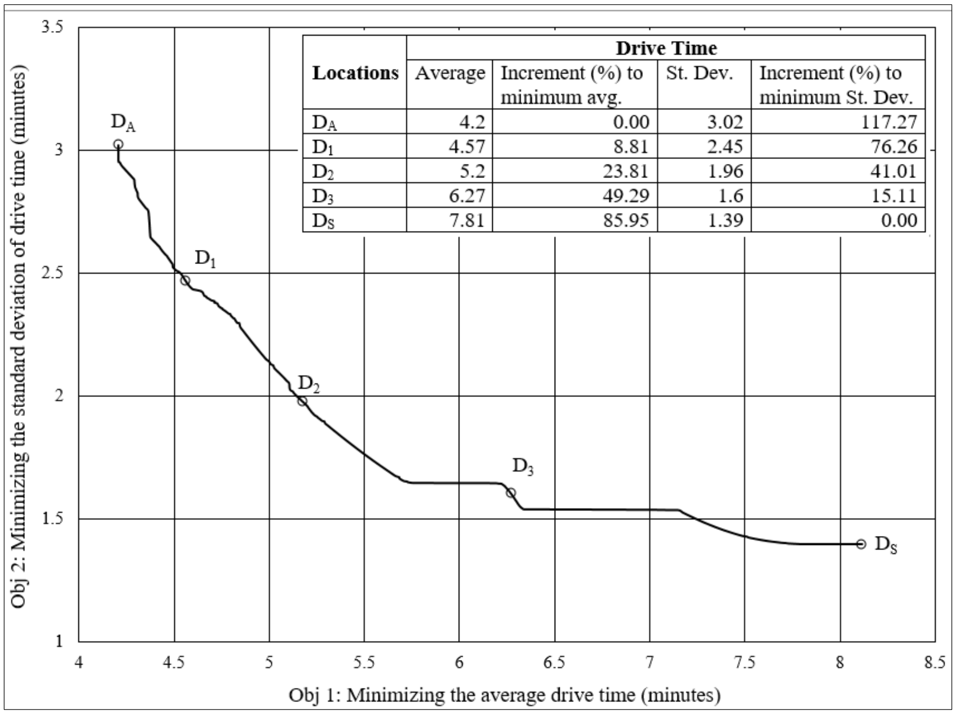
Location D_A and T_A are identified as the potential site for health facility with the lowest average drive and transit time respectively. These would be chosen as optimal locations for siting the facility according to the most popular method for location-allocation problem (i.e. *P-median*). However, *p-median* may fail to incorporate remote users and solution location may not be equitable.

4.2 Bi-optimization of Parameters

Bi-objective optimization model was utilized to minimize both the average and standard deviation of travel time. Two separate models were developed to gain Pareto frontiers corresponding to drive time and ride time-based optimal locations. Minimizing the standard deviation requires reducing the variability in travel time dataset and hence facilitating remote users. Similarly, locating an intervention center by minimizing the average time ensures efficiency by decreasing the total travel time needed for patients to reach that facility. The bi-objective model suggests only solution points that are Pareto optimal. These solution points are called Pareto frontiers and characterize the bounds of what can be considered bi-optimal in the siting of a health facility. Each frontier indicates a location from which it is impossible to reallocate the intervention center in a way that improves one objective without reducing the acceptability of the alternate criterion.

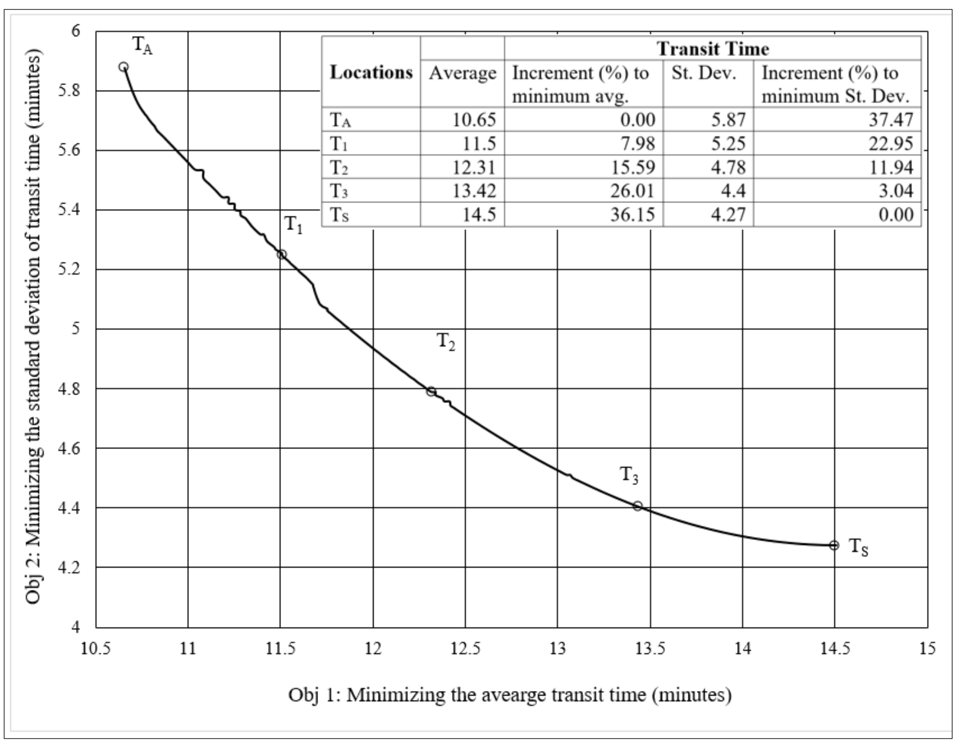
A line of Pareto frontiers can be established by connecting all solution points. Each point along that line represents a unique model parameterization. As Pareto optimality identifies multiple optimal solutions, it allows the decision makers to investigate differences among the solutions and make an informed choice among varying combinations of assessment criteria.

Fig. 3 and 4 show each model derived Pareto frontier. The drive time and transit time-based optimization yielded a set of 235 and 275 pixels on the transportation network, respectively. The minimum average (D_A and T_A) and minimum standard deviation (i.e., D_S and T_S) values bound the Pareto frontier lines obtained from two different models. Three other Pareto frontiers from each model are shown for discussion purposes. D_1, D_2, and D_3 are three compromised solutions at the median and quartile values between D_A and D_S; T_1, T_2, and T_3 are three other compromised solutions at the median and quartile values between T_A and T_S.



Locations	Drive Time Average	Increment (%) to minimum avg.	St. Dev.	Increment (%) to minimum St. Dev.
D_A	4.2	0.00	3.02	117.27
D_1	4.57	8.81	2.45	76.26
D_2	5.2	23.81	1.96	41.01
D_3	6.27	49.29	1.6	15.11
D_S	7.81	85.95	1.39	0.00

Figure 8. The Pareto frontiers of the drive time based bi-objective optimization.



Locations	Transit Time Average	Increment (%) to minimum avg.	St. Dev.	Increment (%) to minimum St. Dev.
T_A	10.65	0.00	5.87	37.47
T_1	11.5	7.98	5.25	22.95
T_2	12.31	15.59	4.78	11.94
T_3	13.42	26.01	4.4	3.04
T_S	14.5	36.15	4.27	0.00

Figure 9.The Pareto frontiers of the transit time based bi-objective optimization.

For example, location D_2 can be reached by an average travel time of 5.20 minutes with a standard deviation value of 1.96 minutes. Although the average travel time is increased by 23.81% when compared to location D_A, the standard deviation is decreased by 1.06. When compared to location D_S, the standard deviation of travel time is increased by 41.01%, but the average travel time is reduced by 2.61 minutes. Similarly, location T_2 reduces the standard deviation of transit time by 1.09 minutes when compared to location T_A by increasing the average travel time by 15.59%. T_2 lessens the average travel time by 2.19 minutes by conceding only 11.94% increase in standard deviation when compared to location T_S.

4.3 Multimodal Optimization

Multimodal optimization of the locations that have already been optimized for a single mode of transportation ensures a balance of equity and efficiency among a combined client set of transit riders and drivers of personal vehicles. The spatial bounding box of frontier solutions of each transportation mode is shown in fig. 5 (i, ii). Fig. 5 (iii) exhibits the common area between optimal drive time and transit time bounds. Fig. 5 (iv) provides a larger scale view of the road sections that are equitably and efficiently optimized for both travel modes. Siting an STI intervention facility along any of these road sections would ensure better access to remote repeat users as well as users in particular high-density areas irrespective of their modal share.

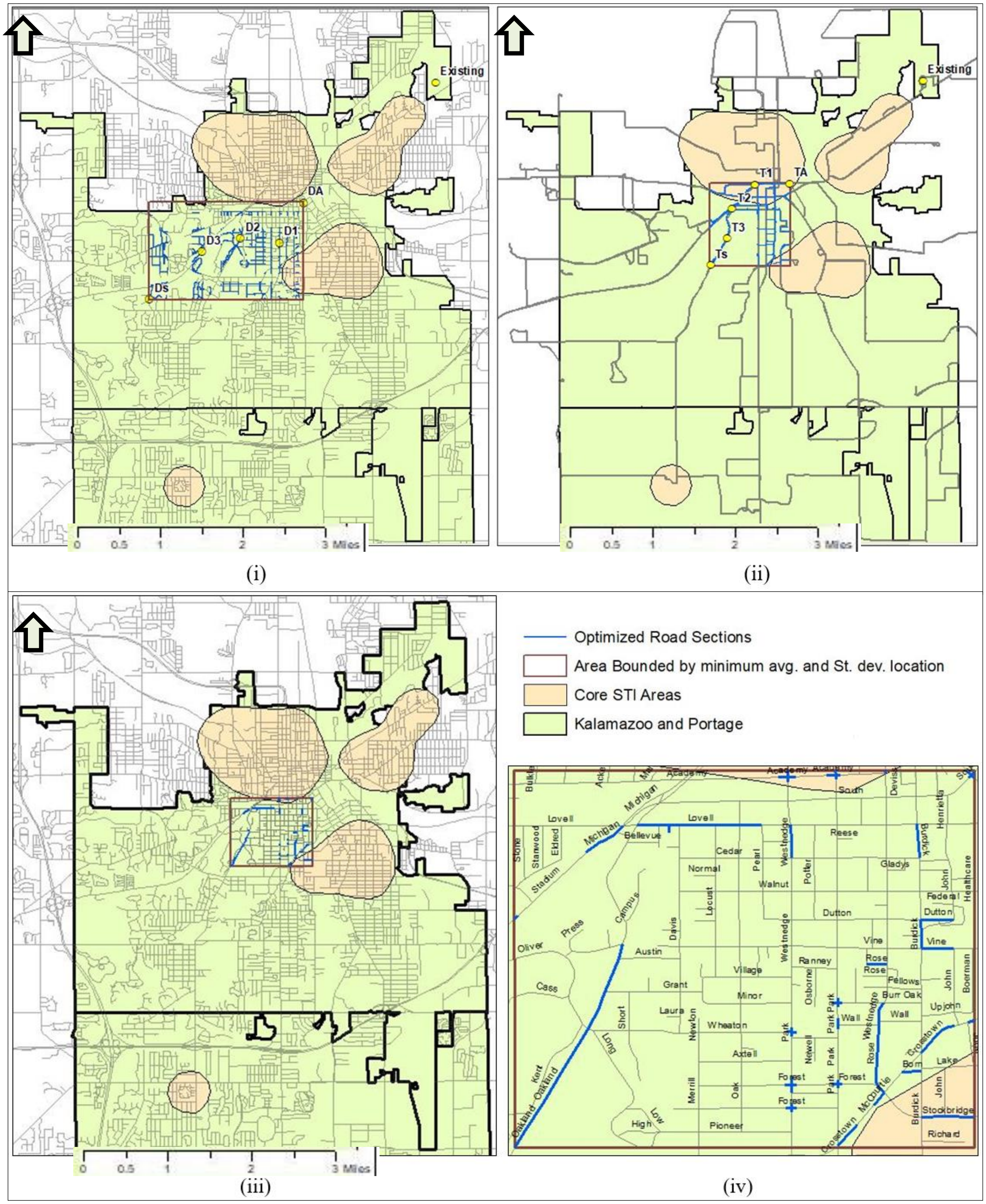


Figure 10. Locations or road sections optimally accessible by (i) drive, (ii) transit and (iii) both (iv) Blow-up of multimodal accessible road sections that are equitably and efficiently optimized.

5. Discussion

Improving spatial access to health facilities is recognized as an important component of reducing the prevalence of disease and achieving better health outcomes. Multimodal accessibility estimation and optimization can play a vital role in this respect. The novel and disaggregated nature of this study allow to consider individual's travel time from distinct household to facility locations, thus helps to address the inherent mismatch between popular statistical methods of significant density detection and the reality of individuals located on a street network or constrained by a particular transportation modality.

Cost raster based optimization not only offers the opportunity to compare different solutions but also paves the way for understanding how this approach may help identify potential locations that could provide better accessibility than the current facility location. For example, the existing facility in Kalamazoo County (see fig. 5) is located an average of 7.44 minutes away from drivers living at hypothetical household locations, with a standard deviation of 3.23 minutes, but location D_2 offers a more accessible location by minimizing the average drive time by 2.24 minutes and standard deviation by 1.27 minutes. This facility houses many programs and services quite apart from STI testing, so the purpose of this paper is not to recommend the relocation of the current facility. Instead, we present a case study of how equity and efficiency of facility placement can be quantified and compared for any number of at-risk populations. The advantage of Pareto solutions is that the analysis can be tailored to a range of populations and objectives. Decision makers with experience in a particular area with a predetermined client base may have specific objectives that are dependent on the geographical distribution of targeted population, socio-economic characteristics, the magnitude of travel time variability and so on, which vary in space and time. This study does not focus on quantifying the preference based on the aforementioned factors, rather offers a set of geocomputational tools to the decision makers for assessing multiple locations.

Network problems are generally considered to be better represented and modeled in vector data models. Common vector network modeling characteristics such as constraints on intersections, non-planar roadways (overpasses), one-way streets, and the like, are difficult or impossible to consider in raster analysis and certainly pose some limitations to the results of this study. However, raster analysis of network problems enable the researcher to make use of techniques such as stacked approach, which are not easily duplicated in vector analysis. Vector analysis also does not lend itself to paired optimality in a setting with virtually no limitations on candidate locations. It is the goal of this paper to present options outside of the standard regimen of vector solutions to network problems.

Another limitation of this case study is that the coincidence of optimization is due to chance, as well as circumstances unique to this particular transportation network for this local area. In general, the area bounded by the minimum average and minimum standard deviation location for each frontier would represent the constraints to intervention location. The extremum frontier values represent the bounds of the 'spatial frontier' or area of potential locations. The size of this area becomes, then, a usable metric by which to measure transit accessibility with respect to accessibility by private vehicle as it will vary by proximity of optimal accessibility and not with city size. By extension, the relative size of this area with respect to the total area of the jurisdiction or total population served can be used by decision makers to quantitatively assess the determinants of intervention site selection within this region.

From a public health policy perspective, equal access to health care is considered one of the most important parameters to address health equity (Oliver & Mossialos, 2004). At a time when socioeconomic disparities are prevalent, multi-modal transportation models can provide insight into the constraints and challenges met by individuals across a spectrum of transportation options including dial-a-ride services, light rail, city bus, a personal vehicle and active transportation options such as cycling and walking. This heuristic approach increases the sophistication of accessibility measurement by quantifying the spatial scope of optimization for specific public health problems and at-risk populations. Additionally, by presenting temporally-aware and spatially disaggregated accessibility metrics, this paper introduces a set of tools that offer efficient as well as more equitable solutions.

References

Apparicio, P., Abdelmajid, M., Riva, M., & Shearmur, R. (2008). Comparing alternative approaches to measuring the geographical accessibility of urban health services: Distance types and aggregation-error issues. *International journal of health geographics, 7*(1), 7.

Ayon, B. D., Owusu, C., Oh, J.-S., & Baker, K. (2018). Pareto Optimality for Assessing Multimodal Transportation Accessibility: Balancing Equity and Efficiency When Sitting Interventions. *Journal of Transport & Health, 9,* S16-S17.

Braveman, P., & Gruskin, S. (2003). Defining equity in health. *Journal of Epidemiology and Community Health, 57*(4), 254-258. doi: 10.1136/jech.57.4.254

Chanta, S., Mayorga, M. E., & McLay, L. A. (2014). Improving emergency service in rural areas: a bi-objective covering location model for EMS systems. *Annals of Operations Research, 221*(1), 133-159.

Church, R., Current, J., & Storbeck, J. (1991). A bicriterion maximal covering location formulation which considers the satisfaction of uncovered demand. *Decision Sciences, 22*(1), 38-52.

Cromley, E. K., & McLafferty, S. L. (2011). Access to Health Services *GIS and Public Health* (Second ed.). New York: The Guilford Press.

Delamater, P. L., Messina, J. P., Shortridge, A. M., & Grady, S. C. (2012). Measuring geographic access to health care: raster and network-based methods. *International journal of health geographics, 11*(1), 15.

Drezner, Z. (1995). Dynamic facility location: The progressive p-median problem. *Location Science, 3*(1), 1-7.

Galindo, G., & Batta, R. (2013). Review of recent developments in OR/MS research in disaster operations management. *European Journal of Operational Research, 230*(2), 201-211.

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

Haghani, A. (1996). Capacitated maximum covering location models: Formulations and solution procedures. *Journal of Advanced Transportation, 30*(3), 101-136. doi: 10.1002/atr.5670300308

Heiman, H. J., & Artiga, S. (2015). Beyond health care: the role of social determinants in promoting health and health equity. *Health*, *20*(10), 1-10.

Hewko, J., Smoyer-Tomic, K. E., & Hodgson, M. J. (2002). Measuring Neighbourhood Spatial Accessibility to Urban Amenities: Does Aggregation Error Matter? *Environment and Planning A: Economy and Space*, *34*(7), 1185-1206. doi: 10.1068/a34171

Huang, R., & Wei, Y. D. (2002). Analyzing Neighborhood Accessibility via Transit in a GIS Environment. *Geographic Information Sciences, 8*(1), 39-47. doi: 10.1080/10824000209480572

Jia, T., Tao, H., Qin, K., Wang, Y., Liu, C., & Gao, Q. (2014). Selecting the optimal healthcare centers with a modified P-median model: a visual analytic perspective. *International journal of health geographics, 13*(1), 42. doi: 10.1186/1476-072X-13-42

Marmot, M., Friel, S., Bell, R., Houweling, T. A. J., & Taylor, S. (2008). Closing the gap in a generation: health equity through action on the social determinants of health. *The Lancet, 372*(9650), 1661-1669. doi: https://doi.org/10.1016/S0140-6736(08)61690-6

Martin, D., Wrigley, H., Barnett, S., & Roderick, P. (2002). Increasing the sophistication of access measurement in a rural healthcare study. *Health & Place, 8*(1), 3-13. doi: https://doi.org/10.1016/S1353-8292(01)00031-4

Mavoa, S., Witten, K., McCreanor, T., & O’Sullivan, D. (2012). GIS based destination accessibility via public transit and walking in Auckland, New Zealand. *Journal of Transport Geography*, *20*(1), 15-22. doi: https://doi.org/10.1016/j.jtrangeo.2011.10.001

McLafferty, S. L. (2003). GIS and Health Care. *Annual Review of Public Health, 24*(1), 25-42. doi: 10.1146/annurev.publhealth.24.012902.141012

Mestre, A. M., Oliveira, M. D., & Barbosa-Póvoa, A. P. (2015). Location–allocation approaches for hospital network planning under uncertainty. *European Journal of Operational Research, 240*(3), 791-806. doi: https://doi.org/10.1016/j.ejor.2014.07.024

Mitropoulos, P., Mitropoulos, I., Giannikos, I., & Sissouras, A. (2006). A biobjective model for the locational planning of hospitals and health centers. *Health Care Management Science, 9*(2), 171-179. doi: 10.1007/s10729-006-7664-9

Neutens, T. (2015). Accessibility, equity and health care: review and research directions for transport geographers. *Journal of Transport Geography, 43*, 14-27. doi: 10.1016/j.jtrangeo.2014.12.006

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. doi: 10.1136/jech.2003.017731

Owusu, C., Baker, K. M., Paul, R., & Curtis, A. B. (2018). Modelling individual vulnerability to sexually transmitted infections to optimise intervention strategies: analysis of surveillance data from Kalamazoo County, Michigan, USA. *Sexually Transmitted Infections, 94*(5), 353-358. doi: 10.1136/sextrans-2017-053350

Rahman, S.-u., & Smith, D. K. (2000). Use of location-allocation models in health service development planning in developing nations. *European Journal of Operational Research, 123*(3), 437-452. doi: https://doi.org/10.1016/S0377-2217(99)00289-1

Schuurman, N., Fiedler, R. S., Grzybowski, S. C. W., & Grund, D. (2006). Defining rational hospital catchments for non-urban areas based on travel-time. *International journal of health geographics*, *5*(1), 43. doi: 10.1186/1476-072X-5-43

Shariff, S. S. R., Moin, N. H., & Omar, M. (2012). Location allocation modeling for healthcare facility planning in Malaysia. *Computers & Industrial Engineering, 62*(4), 1000-1010. doi: https://doi.org/10.1016/j.cie.2011.12.026

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

Tanser, F., Gijsbertsen, B., & Herbst, K. (2006). Modelling and understanding primary health care accessibility and utilization in rural South Africa: An exploration using a geographical information system. *Social Science & Medicine, 63*(3), 691-705. doi: https://doi.org/10.1016/j.socscimed.2006.01.015

Verter, V., & Lapierre, S. (2002). Location of Preventive Health Care Facilities. *Annals of Operations Research, 110*(1), 123-132. doi: 10.1023/A:1020767501233

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. doi: 10.1080/00045608.2012.657146

Wang, F., & Luo, W. (2005). Assessing spatial and nonspatial factors for healthcare access: towards an integrated approach to defining health professional shortage areas. *Health & Place, 11*(2), 131-146. doi: https://doi.org/10.1016/j.healthplace.2004.02.003

Wang, F., & Tang, Q. (2010). *Towards equal accessibility to services: A quadratic programming approach*. Paper presented at the 18th International Conference on GeoInformatics, Beijing, China.