Spatial accessibility to the COVID-19 testing sites and the driven factors behind in NYC

Anran Zheng

Abstract

During the COVID-19 pandemic, the low access to medical resource have brought people inconvenience and health challenges. Given the unequal distribution of healthcare facilities, it is crucial to understand how people access the medical resources in both urban and rural regions, especially from a transportation perspective. This study measured the spatial accessibility to COVID-19 testing sites and its influencing factors in NYC. The clustering patterns of testing sites were identified through spatial autocorrelation and kernel density analysis. Integrating walking, car driving, bus and subway modes, a multimodal network was built to present spatial accessibility scores. Several demographic and socioeconomic variables influencing the spatial accessibility were analyzed based on the Geodetetor model. Accessibility to testing sites was found to be heterogeneous across NYC and among different travel modes. The results of Geodetector showed that the COVID-19 positive rate, population density and the testing sites density were strong indicators in affecting the spatial accessibility. There is evident trend of enforcing interaction between various risk factors. These results provide urban and health planners suggestions on how to ensure the adequate and equitable access to COVID-19 testing sites.

Key words

Spatial accessibility, spatial equity, COVID-19 testing sites, multimodal network, Geodetector.

1. Introduction

With emerging new variants, the COVID-19 has been spreading among the Unites States and posing health and socioeconomic threatens. Since the outbreak of COVID-19, New York City (NYC) became the pandemic epicenter (Cordes & Castro, 2020). Fueled by the Omicron variant, a single-day peak of 50, 803 COVID-19 cases were reported in NYC on January 3rd (Kekatos, 2022). The surge of confirmed cases has exacerbated some challenges that many cities like NYC are facing with, such as the unequal distribution of medical resources and the insufficient supply of COVID testing packages and vaccination. The increased demand and the limited supply of healthcare and medical resources have reflected and even worsen racial and socioeconomic disparities.

Testing site is one important public facility in the prevention and protection of COVID-19. To access the testing sites, public transit provides people a convenient and effective given its large transportation volume, speed and punctuality. However, most transit facilities can also be accessed where population is more concentrated and socioeconomic activities are more active (Chen et al., 2017). Consequently, people living far away from the public transport will access the testing sites restrictively. Considering vulnerable groups, especially elderly, who suffer limited mobility or financial resources, health and policies planners should ensure that the access to COVID-19 testing sites is adequate and equitable across all socioeconomic factors (Duffy, Newing & Gorska, 2021; Tao et al., 2020).

A plethora of studies have focused on the measurement of spatial accessibility to health care services. One commonly used method for measuring the spatial accessibility is the floating catchment (FCA) method proposed by Luo and Wang in examining the spatial accessibility to primary health care in Chicago (2003). Under the context of COVID-19, there are some modifications of FCA methods in the examination of accessibility to healthcare resources. The three-step floating catchment area (3FSCA) method is used to identify the spatial accessibility of COVID-19 patients in Florida (Kim et al., 2021).

Considering the available hospital capacity and the average travel time to hospital, Escobar et al. (2020) used the enhanced two-step floating catchment area (E2SFCA) method to evaluate the current ICU supply in Manizales-Villamaría Metropolitan Area. One limitation of FCA-family method is that the demands, supplies, catchment size and spatial interaction functions are considered as static and fixed values. However, in the context of COVID-19, those variables are spatio-temporal changing. Without the consideration of fixed supply and demand capacities, another typical method that estimated the spatial accessibility is to establish the road network dataset to calculate the O-D travel time matrix (Wang & Wang, 2022). Silalahi et al (2020) created O-D Cost Matrix from the GIS-based network, where the nearest referral hospitals of the COVID-19 confirmed cases locations can be decided in Jakarta. Likewise, a study by Stentzel et al (2016) realized the accessibility to medical care facilities through O-D cost matrix.

There are still some problems with regards assessing the spatial access to the health care facilities during the COVID-19 pandemic, where the demand for medical resources increased dramatically (Ghorbanzadeh et al., 2021). Many research focuses on the county level of accessibility measurement of a whole state (Kim et al., 2021). However, investigating a micro-level region, such as census tracts or block groups, will be more meaningful because in reality, people tend to travel across tracts rather than counties to access the COVID-19 medical resources timely. In addition, most studies only considered one travel mode, especially driving, in the measurement of accessibility. But when accounting for a city, accessibility by different transit modes (personal vehicles, walking and public transit) should be analyzed and compared, since multimodal network is a fundamental component of a city that connects health facilities with people (Del Conte et al., 2022). Moreover, as many research has analyzed the relationship between confirmed cases and its influencing factors (Cordes & Castro, 2020), or between the medical resource distribution and demographic factors (Grigsby-Toussaint, Shin & Jones, 2021), very few studies explored the driven factors of spatial accessibility to medical resources, especially from a geographic perspective. Revealing the influencing factors of accessibility to healthcare facilities will help urban planners understand better about the cause of medical resources inequity and know how to utilize and allocate medical resources rationally.

Based on these challenges, this study focuses on the accessibility to the COVID-19 testing sites in NYC, which has a highly diverse population of 8.8 million people spread across five boroughs interconnected by bus and subway system (U.S. Census Bureau, 2021). Following three research issues are analyzed in this paper. First, the spatial clustering pattern of COVID-19 testing sites is identified through spatial autocorrelation and kernel density estimation methods. Second, both the transit and non-transit road network are developed and compared. By integrating and analyzing the transportation and geographical data, the GIS-based network can build a O-D cost matrix, which can evaluate the accessibility from origin to destinations through different travel modes. Finally, Geodetector method is applied to identify the influencing factors, from socioeconomic and demographic aspects, of spatial accessibilities of testing sites. The goal of this study is to identify the disparities in accessibility to testing sites by different travel modes and the causes of such disparities, then provides guidance for future efforts in allocating testing resources in an equitable manner.

2. Materials and Methods

2.1 Data collection and preparation

Three aspects of data are employed in this study: the COVID-19 testing sites, transit network dataset and the potential influencing factors. According to URISA's GISCorps (https://covid-19-giscorps.hub.arcgis.com/apps/locate-a-covid-19-testing-provider/explore), there are 734 testing sites in NYC. The transit network mainly consists of road network from Open Street Map (OSM)

(https://download.bbbike.org/osm/bbbike/NewYork/) and transit stations, including subways and buses in this study, from OpenMobilityData (https://transitfeeds.com/p/mta). Eleven potential influencing factors behind accessibility, in terms of transportation, socioeconomic and demographic were collected and prepared according to Table 1. Despite the spatial accessibility is measured on the tract level, the zip code tabulation area (ZCTA) level has greater size of samples and the accuracy of detecting influencing factors is greater than the tract level, therefore the collection and preparation of influencing factors are conducted on ZCTA level.

Influencing factor	Data source and prepare process
Population density	Gather population data from ACS first, then divide by area.
Road network density (categorized as transit and non-transit based on the road types)	Gather road network data from Open Street Map and OpenMobilityData. Based on NYC, create fishnet in ArcGIS, intersect this with road data and do statistical analysis.
Transit station density	Gather station data from OSM and spatial join with NYC.
Testing site density	Gather testing site data from URISA's GISCorps and spatial join with NYC.
COVID-19 positive rate	Gather COVID-19 data from NYC Open Data (https://data.cityofnewyork.us/browse?category=Health&q=covid)
Other demographic data: median income, median age, percentage of white, bachelor and public transit as mean of commuting	Gather from ACS

Table 1. data collection and preparation of influencing factors

NYC, the study area, covers 5 counties, 177 ZCTA blocks and 2167 census tracts in total (https://www1.nyc.gov/site/doh/data/data-sets/data-sets-and-tables.page). Some influencing factors of spatial accessibility in NYC are visualized in Figure 1. Population density and testing site density are highest in Manhattan, while lowest in Staten Island and eastern Queens county. Up to March 20th 2022, the COVID-19 positive rate was highest in Staten Island, as most ZCTA blocks have about 30% positive rate. Central and southern Manhattan has the lowest positive rate of COVID-19 that close to 10%. There is strong regional heterogeneity in COVID-19 pandemic situation, population and medical resource distribution. It is necessary to study the spatial accessibility and its influencing factors for urban planners to optimize the allocation of COVID-19 related healthcare resources.

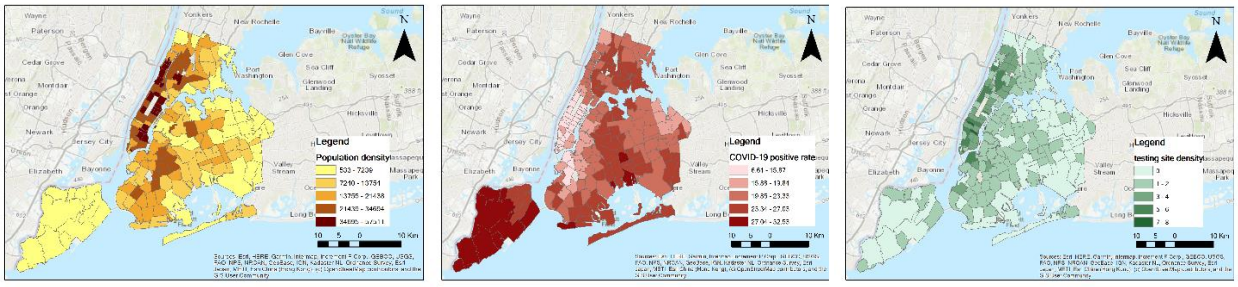

Figure 1. examples of accessibility influencing factors distribution: (a) population density. (b) COVID-19 positive rate. (c) testing site density

2.2 Method

2.2.1 Spatial distribution

Spatial autocorrelation and kernel density estimation methods are applied to investigate the presence of spatial clustering of testing sites. For spatial autocorrelation analysis, both global and local Moran's I were employed. Moran's I is a common measure of spatial autocorrelation in statistics. The possible values of Moran's I range from -1 to 1, where low negative values suggest strong negative spatial autocorrelation, high positive values suggest strong positive spatial autocorrelation, and values close to zero suggest complete spatial randomness (Cordes & Castro, 2020). Kernel density estimation is to estimate density from point-based data. By applying a kernel function on each point and spreading the observation over the kernel window, it usually results in a density surface that will cover the whole study area (Yin, 2020).

2.2.2 Accessibility measurement

In the measurement of spatial accessibilities by different travel modes, four most dominant travel patterns in NYC, including walking, driving, subway and bus, are compared based on the O-D cost matrix of ArcGIS Network Analyst. Walking and car driving are categorized into non-transit mode, while subways and buses are considered as transit mode. Network is a type of linear vector data that consists of edges, junctions, and nodes. The network dataset can model the spatial accessibility by calculating the distance from nodes (Silalahi et al., 2020). The O-D cost matrix, generated from the transportation network analysis, is the estimation of cost (e.g. travel time, distance) between a set of origins and destinations (Wang & Xu, 2011).

The key of network analysis is to define the origin and destination points for O-D cost matrix. Considering the flexibility and the wide range of non-transit takers' mobility, population weighted centroids were chosen as the origins of the non-transit road network. In contrast, the transit road network structure has bounded stations as origins. The destinations of both two types of road network are testing sites. Considering that for convenience and time savings, people tend to reach their nearest testing sites. This study hence only measures the accessibility to testing sites within 15 minutes of travelling by different modes from origins. After calculating the accessibility of each origin, the IDW interpolation can be utilized to estimate the accessibility of the whole study region.

2.2.2.1 Non-transit accessibility

In the non-transit road network, the calculation of census tract i's accessibility can be measured as the following equation:

$$\mathrm{Ai} = \mathrm{Pi} * \sum_{n}^{0} \frac{1}{Ti} \tag{1}$$

Where A_i is the accessibility of census tract i, P_i is the total population number of census tract i, n is number of testing sites that people can access within 15 minutes and T_i is the time people spend getting to each testing site.

The road network for walking and car driving were constructed from OSM dataset, which contains information for every road segment about speed limits, distances, directions (e.g. one-way streets) and turn restrictions. Considering that people typically walk and drive on different types of roads, this study extract common types of roads for pedestrians and drivers as table 2 shows. To calculate the accessibility as equation (1) shows, speeds of walking or driving on typical roads are decided in table 2, where 5km/h are set as the average speed people walk, while different speeds of driving are assigned according to the

different road types. Different from building the pedestrian road network, some restrictions are included in the driving road network, such as one-way driving, delay of turning and traffic lights. Since the elevation of NYC is relatively flat, its impact on the walking or driving can be ignored.

Non-transit	Road type	Speed (mph)
walk	Footway, living street, path, residential, service, etc.	3.5
drive	Motorway	50
	Primary road	40
	Secondary	30
	Tertiary	25
	Residential	15

Table 2. road types and speeds of walking and driving

2.2.2.2 Transit accessibility

To evaluate how transit improves the accessibility, subways and buses, two dominant transit systems, are chosen and their accessibilities are calculated differently from the non-transit modes, since transit routes and schedules are fixable with given stops that travelers should take. Unlike the non-transit road which is extracted directly from the OSM dataset, the transit road network for subway and bus are built based on the General Transit Feed Specification (GTFS) data using Conversion tools in ArcGIS Pro. The transit accessibility of each stop is calculated as the following equation:

$$Ai = \sum_{n}^{0} \frac{1}{Ti} \tag{2}$$

Where n stands for the count of testing sites that can be reached within 15 minutes of taking the subway or buses, while T_i is the time people spend getting to each testing site starting from stop i.

The speeds of taking the subway and bus are set to 28km/h and 13km/h respectively. To build a comprehensive transit road network system, pedestrian road network is combined with the subway or bus route system, since when getting off the nearest stops to the testing sites, people still need to walk for a while to reach their final destination, the testing sites. In the combination of these road networks, the connected junction between pedestrian lanes and subway or bus routes are built within the 50m buffer zones of stops so that people can switch from subway or bus modes to walk.

2.2.3 Influencing factors detection

To understand the main influencing factors of spatial accessibility, this study applied Geodetector to study the association between the potential risk factors and the overall spatial accessibility across ZCTA blocks, which combine the accessibility of four transit modes and its value within each ZCTA is calculated through zonal statistics in ArcGIS. The Geodetector method can measure the spatial differentiation and tested its significance, through the within strata variance less than the between strata variance (Xie et al., 2020). To study the influencing factors and reveal relationships among these factors, this study emploted factor detection and interaction detection from Geodetector.

The factor detection is expressed by q value with the following formulas (Wang & Xu, 2017):

$$q = 1 - \frac{\sum_{h=1}^{L} N_h \sigma_h^2}{N\sigma^2} = 1 - \frac{SSW}{SST}$$

$$SSW = \sum_{h=1}^{L} N_h \sigma_h^2 \, , \; SST = N\sigma^2 \tag{3}$$

Where q-statistic represents the explanatory power of factor X on the spatial heterogeneity of factor Y, the value of q ranges from 0 to 1; h = 1, …, L, which represents the stratification of the detector X and Y; the study area consists of N units and the h^{th} stratum consists of N^h units; σ^2_h and σ^2 are the variance of Y value for the h^{th} stratum and the whole study area. SSW and SST are the Within Sum of Squares of a layer and the Total Sum of Squares of New York State.

Interaction detection can identify the interaction relationship between two different factors. The q-statistics of X1 and X2, q(X1) and q(X2), are calculated first from equation (3). Then the q values of the interaction between X1 and X2, (q(X1 ∩ X2)), are calculated. The interaction type between X1 and X2 can be determined by comparing the values of q(X1), q(X2) and q(X1 ∩ X2) (Zheng et al., 2021).

In applying the Geodetector model, continuous risk factors should be transformed into discrete variables before their relationship with the accessibility is analyzed. Four typical methods of data classification, including natural breaks, equal interval, geometrical interval and quantile, are implemented to discretize the risk factors. By comparing the q-statistics and p-value, the optimal one is chosen.

3. Results

3.1 Spatial distribution pattern of testing sites

With spatial autocorrelation analysis and kernel density estimation methods, the spatial clustering pattern of COVID-19 testing sites will be identified in Figure 1. According to (a), the spatial distribution of testing sites is uneven, then (b) and (c) further confirmed the hotspots distribution in NYC. The testing sites mainly clustered at Manhattan and Southern Bronx, while numerous cold spots appeared in the Queen and Brooklyn county. Given the z-score of 16.52 from Global Moran's I Summary, the testing sites are spatially clustered in NYC.

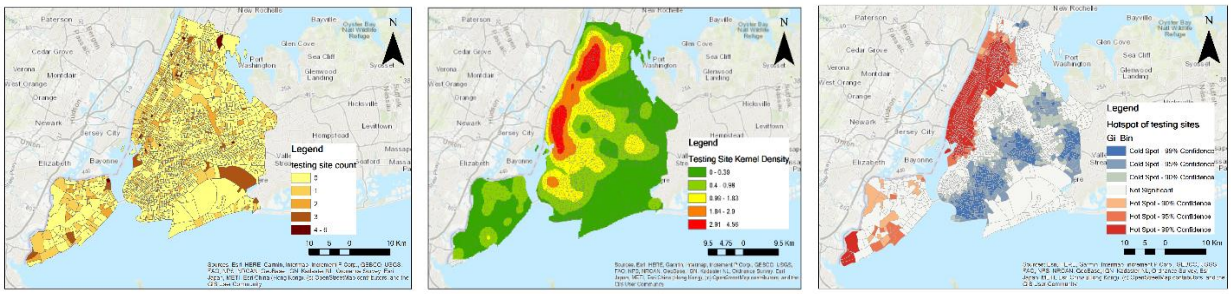

Figure 1. the spatial distribution of COVID-19 testing sites in NYC: (a) count of testing sites contained in each tract. (b) kernel density estimation of testing sites. (c) hotspots of testing sites

3.2 Geographic patterns of accessibility

Figure 2 shows the spatial accessibility of walking and driving. Walk accessibility indicates a tightly clustering distribution pattern, which is highest in Manhattan and decreases gradually to the edge. In contrast, the distribution of drive accessibility is random, with a few highest regions in Manhattan and Brooklyn, Queens and Staten Island. The accessibility enhances significantly from walking to driving, since ideally with driving, people can access to more testing sites within a shorter time. However, in reality the driving accessibility will be less than this study shows, considering the traffic congestion and parking issues.

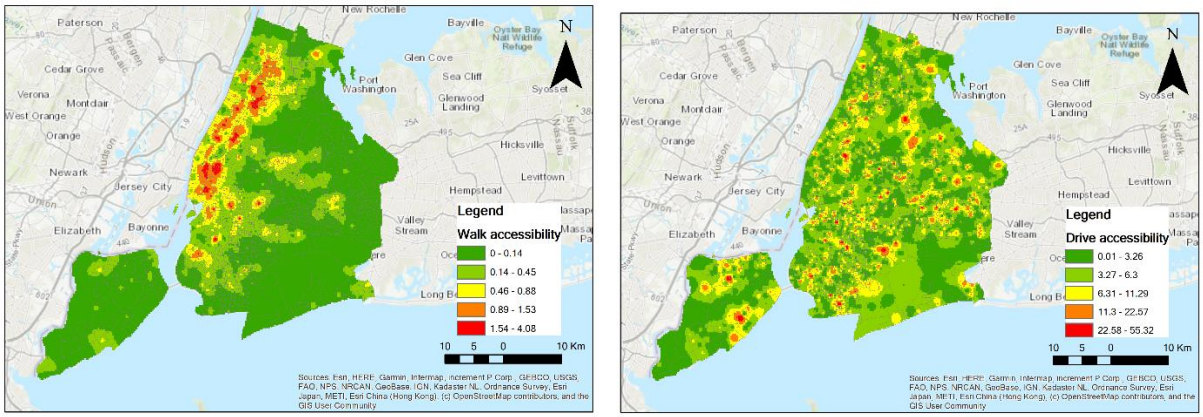

Figure 2. the non-transit accessibility to testing sites within 15 minutes: (a) walking, (b) driving

The transit accessibility is calculated in figure 4. Transit accessibility is highest in Manhattan and decay gradually to the boundary of NYC. Most regions are more accessible to subways than bus, since typically subways take shorter time to wait and hardly encounter traffic jam, despite that there are much more bus stations than subways in NYC.

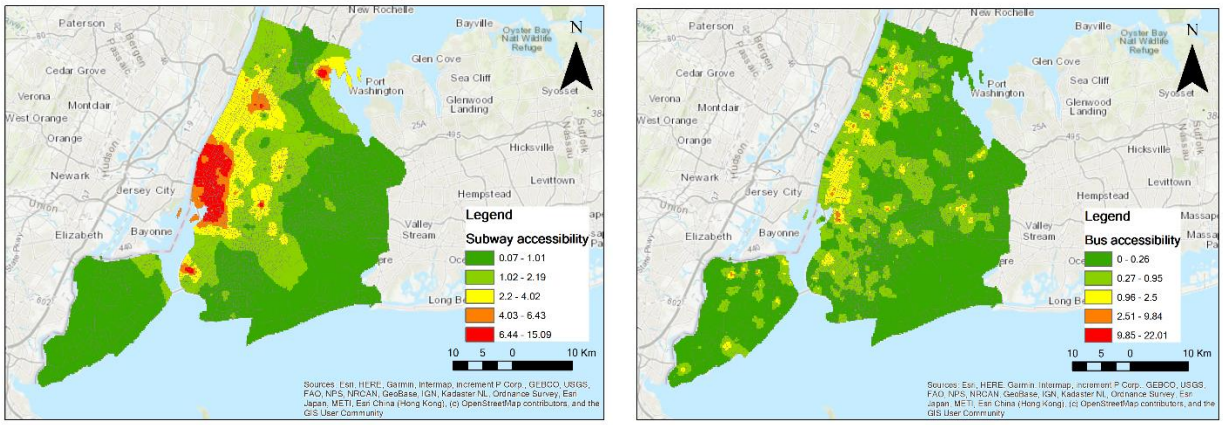

Figure 3. the transit accessibility to testing sites within 15 minutes: (a) subway, (b) bus

The overall accessibility layer can be generated by combining the accessibility of four different transit modes through raster calculator in ArcGIS. As figure 4 shows, Manhattan has the highest access to COVID-19 testing sites as expected because of the dense road network and the concentrated distribution of testing sites. Beyond downtown area, access values decreased quickly due to limited transportation infrastructure, however, with several tracts in Brooklyn, eastern Queens, Staten Island and Bronx still have relatively high access to testing sites. It is clear that the testing site accessibility varies greatly in NYC.

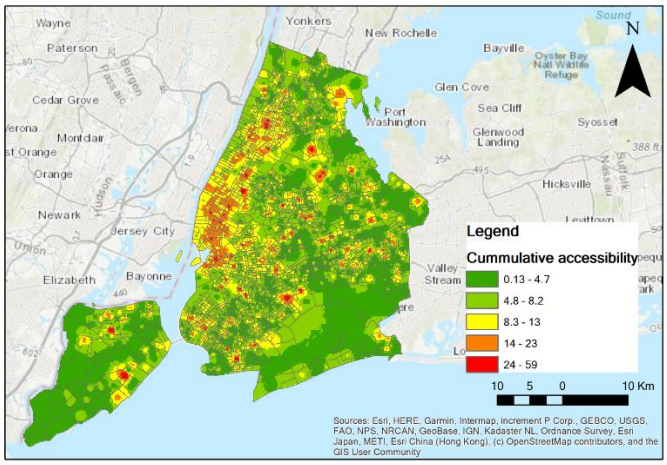

Figure 4. the overall accessibility to testing sites in NYC

3.3 Influencing factors of accessibility

There is obvious variation in accessibility to COVID-19 testing sites in NYC. Geodetector was applied to analyze the driving factors in this distribution pattern and their interaction. Four various methods of classification are utilized in the discretization of influencing factors as table S1 shows. Both the q-statistics and p-values generated from each method were compared to find the best discretization method. Quantile was chosen in discretizing the influencing factors.

As table 3 shows, the q values for all chose detection factors passed the significance test at 5% level, suggesting that all these factors have a significant determination ability of spatial accessibility. The q-statistics, the explanatory power, are not very high, as only four out of eleven factors' q statistics are close to 0.5: median income (0.4239), percentage of bachelor (0.4328), COVID-19 positive rate (0.5060) and testing site density (0.4994). The determination of other factors, such as median age (0.0744), percentage of white (0.0687) and transit stop density (0.1702), is relatively weak.

Influencing factor	q-statistic	p-value
Population density	0.4398	0.0000
Median Income	0.4239	0.0000
Median Age	0.0744	0.0128
White (%)	0.0687	0.0229
Bachelor (%)	0.4328	0.0000
Public Transport (%)	0.1400	0.0001
COVID Positive Rate	0.5060	0.0000
Testing site Density	0.4994	0.0000
Transit Road Density	0.2910	0.0000
Non-Transit Road Density	0.1591	0.0000
Transit stop density	0.1702	0.0000

Table 3. factor detection results of influencing factors

To solve the limitation of factor detector, which only consider the determination ability of single factor on the spatial accessibility, the interaction detector is used to identify the interaction between any two factors in table S2. In most cases, the q-statistics of two intersecting factors were greater than the q-statistic of

single factors, suggesting that the interaction probe of detection factors had a significant effect on the accessibility. In detail, the intersection between the COVID-19 positive rate (X7) and the testing sites density (X8) generates the most reinforcing effect, as q-statistics is about 0.6757. X7 also intersects with the population density (X1) and transit stop density (X11) to produce relatively strong enhancing effects. The intersection between X1 and bachelor proportion (X5) and median income (X2) generates reinforcing effect as well. In summary, the influence of most detection factors on the accessibility was not independent buy showed mutual or nonlinear enhancement.

4. Discussion

Based on the multimodal network, the results reveal strong spatial variation in testing site accessibility among different travel modes. By personal vehicles or taxi, people have the highest access to testing sites in NYC. In contrast, public transportation tends to be less accessible to destination, while pedestrians have the lowest access to testing sites. The spatial accessibility to testing sites by walking and public transit are found to share similar trends, as the accessibility by these modes is higher in downtown area but decay gradually to the edge of NYC. On the other hand, there is not clear clustering pattern in car accessibility to testing sites. Such differences might owe to the variation of road network. The sidewalk, subway and bus routes all share similar geographic distribution, as most of them pass through the downtown area and hardly stretch to some tracts near the NYC boundary. The road system of car, however, displays relatively even spatial distribution. In addition, compared with walking, bus and subways, cars can drive with more flexibility and mobility. Therefore, even some outskirts districts are associated with less testing sites and relatively inconvenient public transit services, cars can still connect people in these regions with high accessibility to testing sites.

From the results of factor detector in Geodetector model, several demographic and socioeconomic variables, including the COVID-19 positive rate, testing site density and population density, are proven to have a positive and significant relationship with accessibility to testing sites. However, the explanatory power of these variables are not very high. After introducing the intersection detector to explore the mutual effect of any two factors, their influence on the accessibility exceeded that of single detection factors. The intersection between the COVID-19 positive rate and population density, transit stop density, and testing site density have a greater impact on the accessibility. This suggests that the spatial heterogeneity of accessibility is decided by multiple intersecting factors rather than single one.

There are some strengths in this study. First, this study is conducted on the micro-level, since the accessibility was measured across census tracts, while the influencing factors of this accessibility were analyzed on the ZCTA level for a more precise outcome. Secondly, four common travel modes in NYC were included in building a multimodal network with a comprehensive transit system. The disparities of accessibility through various travel behaviors were highlighted, which fills the gap of many previous research that only focus on a single transport mode in the measurement of accessibility. Lastly, when using Geodetector to explore the influencing factors, the performance of four various discretization methods were evaluated based on the q-statistics and p-values from the result. The most appropriate discretization method can adapt to the spatial heterogeneity of different influencing factors as possible.

Regarding the COVID-19 testing site accessibility, NYC displays a high degree of inequity in distribution. Based on the finding, the following suggestions are made for the decision of urban planners and policy makers of NYC. The testing capacity should be expanded by increasing capacity at existing ones and adding new ones, especially in Staten Island, eastern Queen and southern Brooklyn County. The testing sites should be distributed evenly as possible and their service range should be expanded, which can satisfy more needs of people. In addition, since transportation facilities are important in the connection of people with access to testing sites, new transit routes with optimally located stops and improved transit

frequency should be built strategically. This will provide people without car ownership or elderly more access to testing resources. Furthermore, for regions with low access to healthcare services, visiting healthcare programs should be implemented and more COVID-19 self-testing kits should be allocated in solve the severity of equity in accessibility.

A few limitations exist in our study. In building the non-transit road network, the chosen of population weighted centroid as network origin might be less accurate than the residential address. This approach might suffer from aggregation bias and the utilization of more disaggregated data can address such problem. Besides, it is somewhat arbitrary to consider a 15-minute travel time as a threshold in defining the accessibility. This travel time is not suitable for everyone, especially disabilities, elderly and people who have COVID-19 symptoms and want a urgent test immediately. Finally, the multimodal network does not take into account factors that cause potential travel time delay, such as the time waiting for the transit, the transfer between different transit lines and the traffic jam. Therefore, the accessibility value calculated from our research might be lower than the reality. More elements, such as the peak and off-peak driving speed, the bus and subway time schedule, should be incorporate in building a more accurate multimodal network.

5. Conclusion

This study first conducted both spatial autocorrelation and kernel density to explore the spatial clustering pattern of COVID-19 testing sites in NYC. Based on the OSM and GTFS dataset, a GIS-based multimodal network, which include travel by car, bus, subway and walking, was built to show the variations in spatial accessibility in NYC. The relationship between spatial accessibility with socioeconomic and demographic characteristics is explored through the Geodetector model. The results show strong spatial heterogeneity across NYC by different travel modes. Urban core, where population, well-developed transit facilities and testing sites concentrated, provides people with more access to testing sites than people living in the edge of NYC. Compared with walking and public transit, travelling with cars is more effective and efficient in expanding people's accessibility to testing sites. The COVID-19 positive rate, population density and the testing sites density contribute to such spatial variation in accessibility. To improve the equity of accessibility, the capacity of testing sites and transit facilities should be expanded in regions suffering from low accessibility to testing sites with a focus on people with low mobility. Future research should identify the gaps in building a more accurate multimodal network in accessibility measurement, such as trying different travel time thresholds, setting rational travel time delay based on the traffic situation and transit schedule.

References

[1] Cordes, J. and Castro, M.C. (2020) "Spatial analysis of COVID-19 clusters and contextual factors in New York City" Spat Spatiotemporal Epidemiol 34 pp.100355 DOI: 10.1016/j.sste.2020.100355.

[2] Kekatos, M. (2022) "COVID-19 cases in NYC show omicron infections may be plummeting", abcNewS. https://abcnews.go.com/Health/covid-19-cases-nyc-show-omicron-infections-plummeting/story?id=82271946

[3] Chen, J., Ni, J., Xi, C., Li, S. and Wang, J. (2017) "Determining intra-urban spatial accessibility disparities in multimodal public transport networks" Journal of Transport Geography 65 pp.123-133 DOI: 10.1016/j.jtrangeo.2017.10.015.

[4] Duffy, C., Newing, A. and Gorska, J. (2021) "Evaluating the Geographical Accessibility and Equity of COVID-19 Vaccination Sites in England" Vaccines (Basel) 10 (1) DOI: 10.3390/vaccines10010050.

[5] Tao, R., Downs, J., Beckie, T.M., Chen, Y. and McNelley, W. (2020) "Examining spatial accessibility to COVID-19 testing sites in Florida" Annals of GIS 26 (4) pp.319-327 DOI: 10.1080/19475683.2020.1833365.
[6] Luo, W., & Wang, F. (2003). “Measures of Spatial Accessibility to Healthcare in a GIS Environment: Synthesis and a Case Study in Chicago Region.” Environ Plann B Plann Des, 30(6), 865-884. https://doi.org/10.1068/b29120
[7] Kim, K., Ghorbanzadeh, M., Horner, M.W. and Ozguven, E.E. (2021) "Identifying areas of potential critical healthcare shortages: A case study of spatial accessibility to ICU beds during the COVID-19 pandemic in Florida" Transp Policy (Oxf) 110 pp.478-486 DOI: 10.1016/j.tranpol.2021.07.004.
[8] Escobar, D.A., Cardona, S. and Ruiz, S. (2020) "Planning of expansion of ICU hospital care in times of Covid-19 using the E2SFCA model" Revista Espacios, 10.48082/espacios-a20v41n42p03 DOI: 10.48082/espacios-a20v41n42p03.
[9] Wang, F. and C. Wang. (2022) GIS-Automated Delineation of Hospital Service Areas. Boca Raton, FL: CRC Press. ISBN 9780367202286
[10] Silalahi, F.E.S., Hidayat, F., Dewi, R.S., Purwono, N. and Oktaviani, N. (2020) "GIS-based approaches on the accessibility of referral hospital using network analysis and the spatial distribution model of the spreading case of COVID-19 in Jakarta, Indonesia" BMC Health Serv Res 20 (1) pp.1053 DOI: 10.1186/s12913-020-05896-x.
[11] Stentzel, U., Piegsa, J., Fredrich, D., Hoffmann, W. and van den Berg, N. (2016) "Accessibility of general practitioners and selected specialist physicians by car and by public transport in a rural region of Germany" *BMC Health Serv Res* 16 (1) pp.587 DOI: 10.1186/s12913-016-1839-y.
[12] Ghorbanzadeh, M., Kim, K., Erman Ozguven, E. and Horner, M.W. (2021) "Spatial accessibility assessment of COVID-19 patients to healthcare facilities: A case study of Florida" Travel Behav Soc 24 pp.95-101 DOI: 10.1016/j.tbs.2021.03.004.
[13] Del Conte, D.E., Locascio, A., Amoruso, J. and McNamara, M.L. (2022) "Modeling multimodal access to primary care in an urban environment" *Transportation Research Interdisciplinary Perspectives* 13 DOI: 10.1016/j.trip.2022.100550.
[14] Grigsby-Toussaint, D.S., Shin, J.C. and Jones, A. (2021) "Disparities in the distribution of COVID-19 testing sites in black and Latino areas in new York City" Prev Med 147 pp.106463 DOI: 10.1016/j.ypmed.2021.106463.
[15] U.S. Census Bureau QuickFacts: New York City, New York. (2022). Retrieved February 27th, 2022, from https://www.census.gov/quickfacts/newyorkcitynewyork .
[16] Yin, P. (2020). Kernels and Density Estimation. *The Geographic Information Science & Technology Body of Knowledge* (1st Quarter 2020 Edition), John P. Wilson (ed.). DOI: 10.22224/gistbok/2020.1.12
[17] Silalahi, F. E. S., Hidayat, F., Dewi, R. S., Purwono, N., & Oktaviani, N. (2020). GIS-based approaches on the accessibility of referral hospital using network analysis and the spatial CPLN 550 Final Exam Group B Q2 13 distribution model of the spreading case of COVID-19 in Jakarta, Indonesia. BMC Health Serv Res, 20(1), 1053. https://doi.org/10.1186/s12913-020-05896-x
[18] Wang, F., & Xu, Y. (2011). Estimating O–D travel time matrix by Google Maps API: implementation, advantages, and implications. Annals of GIS, 17(4), 199-209. https://doi.org/10.1080/19475683.2011.625977

[19] Xie, Z., Qin, Y., Li, Y., Shen, W., Zheng, Z. and Liu, S. (2020) "Spatial and temporal differentiation of COVID-19 epidemic spread in mainland China and its influencing factors" Sci Total Environ 744 pp.140929 DOI: 10.1016/j.scitotenv.2020.140929.
[20] Wang,J.;Xu,C. (2017) Geodetector: Principle and prospective. ActaGeograph. Sin. 72,116–134.
[21] Zheng, A.; Wang, T.; Li, X (2021). Spatiotemporal Characteristics and Risk Factors of the COVID-19 Pandemic in New York State: Implication of Future Policies. ISPRS Int. J. Geo-Inf. 10, 627. https://doi.org/10.3390/ijgi10090627
[22] Cheng, W.; Xi, H.; Celestin, S. (2020) Application of geodetector insensitivity analysis of reference crop evapotranspiration spatial changes in Northwest China. Sciences in Cold and Arid Regions. 13(4):314−325.DOI:10.3724/SP.J.1226.2021.20038.

Supplementary material

All the data, figure and coding scripts can be found on the author's Github: https://github.com/Anran0716/Anran-Zheng-AccessTestingSite

methods	quantile		natural jenk		geometric interval		equal interval	
factors	q-statistic	p-value	q-statistic	p-value	q-statistic	p-value	q-statistic	p-value
Population density	0.4398	0.0000	0.4346	0.0000	0.4219	0.0000	0.4002	0.0000
Median Income	0.4239	0.0000	0.4301	0.0000	0.3362	0.0000	0.4239	0.0000
Median Age	0.0744	0.0128	0.0612	0.0730	0.0682	0.0240	0.0744	0.0128
White(%)	0.0687	0.0229	0.0679	0.0216	0.0674	0.0218	0.0687	0.0229
Bachelor (%)	0.4328	0.0000	0.4094	0.0000	0.4024	0.0000	0.4328	0.0000
Public Transport (%)	0.1400	0.0001	0.1696	0.0000	0.1571	0.0000	0.1400	0.0001
COVID Positive Rate	0.5060	0.0000	0.4819	0.0000	0.4446	0.0000	0.4873	0.0000
Testing site Density	0.4994	0.0000	0.5361	0.0000	0.5403	0.0000	0.5053	0.0000
Transit Road Density	0.2910	0.0000	0.2895	0.0000	0.2864	0.0000	0.2254	0.0034
Non-Transit Road Density	0.1591	0.0000	0.1823	0.0059	0.1580	0.0099	0.2340	0.0186
Stop density	0.1702	0.0000	0.1128	0.0230	0.1365	0.0001	0.1087	0.0419

Table S1. The comparison of results from four different data classification methods

factor 1	factor 2	value	comparison	result
x1	x3	0.4781	(Min(x1,x3),Max(x1,x3))	Weakened, single factor nonlinear
x3	x10	0.2838	(Min(x3,x10),Max(x3,x10))	Weakened, single factor nonlinear
x3	x6	0.2626	(Min(x3,x6),Max(x3,x6))	Weakened, single factor nonlinear
x4	x10	0.3538	(Min(x4,x10),Max(x4,x10))	Weakened, single factor nonlinear
x4	x6	0.3472	(Min(x4,x6),Max(x4,x6))	Weakened, single factor nonlinear
x4	x9	0.4376	(Min(x4,x9),Max(x4,x9))	Weakened, single factor nonlinear
x3	x9	0.3964	<Min(x3,x9)	Weakened, nonlinear
x1	x10	0.521	>Max(x1,x10)	Enhanced, double factors
x1	x2	0.6661	>Max(x1,x2)	Enhanced, double factors
x1	x4	0.5775	>Max(x1,x4)	Enhanced, double factors
x1	x8	0.6235	>Max(x1,x8)	Enhanced, double factors
x1	x9	0.5663	>Max(x1,x9)	Enhanced, double factors
x2	x10	0.529	>Max(x2,x10)	Enhanced, double factors

x2	x3	0.5334	>Max(x2,x3)	Enhanced, double factors
x2	x4	0.4787	>Max(x2,x4)	Enhanced, double factors
x2	x5	0.5359	>Max(x2,x5)	Enhanced, double factors
x2	x6	0.5988	>Max(x2,x6)	Enhanced, double factors
x3	x4	0.3173	>Max(x3,x4)	Enhanced, double factors
x3	x5	0.551	>Max(x3,x5)	Enhanced, double factors
x3	x7	0.615	>Max(x3,x7)	Enhanced, double factors
x3	x8	0.5437	>Max(x3,x8)	Enhanced, double factors
x4	x5	0.4979	>Max(x4,x5)	Enhanced, double factors
x4	x7	0.567	>Max(x4,x7)	Enhanced, double factors
x4	x8	0.6174	>Max(x4,x8)	Enhanced, double factors
x5	x10	0.5863	>Max(x5,x10)	Enhanced, double factors
x5	x6	0.6152	>Max(x5,x6)	Enhanced, double factors
x5	x7	0.563	>Max(x5,x7)	Enhanced, double factors
x6	x10	0.3433	>Max(x6,x10)	Enhanced, double factors
x6	x7	0.6233	>Max(x6,x7)	Enhanced, double factors
x6	x8	0.6097	>Max(x6,x8)	Enhanced, double factors
x6	x9	0.5522	>Max(x6,x9)	Enhanced, double factors
x7	x10	0.6359	>Max(x7,x10)	Enhanced, double factors
x7	x8	0.6757	>Max(x7,x8)	Enhanced, double factors
x7	x9	0.6166	>Max(x7,x9)	Enhanced, double factors
x8	x10	0.5618	>Max(x8,x10)	Enhanced, double factors
x8	x9	0.5506	>Max(x8,x9)	Enhanced, double factors
x9	x10	0.4256	>Max(x9,x10)	Enhanced, double factors
x1	x5	0.6546	>x1+x5	Enhanced, nonlinear
x1	x6	0.5675	>x1+x6	Enhanced, nonlinear
x1	x7	0.6709	>x1+x7	Enhanced, nonlinear
x2	x7	0.5579	>x2+x7	Enhanced, nonlinear
x2	x8	0.6704	>x2+x8	Enhanced, nonlinear
x2	x9	0.58	>x2+x9	Enhanced, nonlinear
x11	x1	0.5527	>Max(x1,x11)	Enhanced, double factors
x11	x10	0.3129	>Max(x10,x11)	Enhanced, double factors
x11	x2	0.5946	>Max(x2,x11)	Enhanced, double factors
x11	x3	0.3146	>Max(x3,x11)	Enhanced, double factors
x11	x4	0.3538	>x4+x11	Enhanced, nonlinear
x11	x5	0.5863	>Max(x5,x11)	Enhanced, double factors
x11	x6	0.3433	>x6+x11	Enhanced, nonlinear
x11	x7	0.6359	>Max(x7,x11)	Enhanced, double factors
x11	x8	0.5467	>Max(x8,x11)	Enhanced, double factors
x11	x9	0.4663	>Max(x9,x11)	Enhanced, double factors

Table S2. Results of interaction detection (X1 - Population density, X2 - Median Income, X3 - Median Age, X4 - White(%), X5 - Bachelor (%), X6 - Public Transport (%), X7 – COVID-19

Positive Rate, X8 - Testing site Density, X9 - Transit Road Density, X10 – Non-transit Road Density, X11 - Stop density)