AIM: Using AI to Improve Residential Racial and Economic Segregation for Mortality Analysis

Jinwei Liu
Department of Computer and Information Sciences
Florida A&M University
Tallahassee, FL 32307, USA
jinwei.liu@famu.edu

Rui Gong
Department of Informatics and Mathematics
Mercer University
Macon, GA 31207, USA
gong_r@mercer.edu

***Abstract*—Residential racial and economic segregation has been associated with an increased risk for main causes of death. Quintiles of the Index of Concentration and the Extremes (ICE) are commonly used to measure residential segregation because of their robust and comprehensiveness. However, the high correlation between ICE and the proportion of population in poverty can cause the collinearity, and this issue was avoided by ignoring poverty rate or transforming the continuous variable to the category variable in previous research, which induced the missing information because poverty was one significantly crucial covariant in different models for the association of residential segregation and health outcomes. In this paper we developed one new methodology named AIM by exploring artificial intelligence (AI) techniques to extract the integrated information from poverty. The original algorithm of this methodology utilized the high correlation between ICE and poverty rate to implement balanced clustering, by which ICE quintiles were improved and the stronger association of the AI based segregation with age-adjusted mortality rate was observed.**

***Index Terms*—Poverty Rate, Balanced Clustering, Age-Adjusted Mortality Rate, AI based Segregation, Health Outcome**

I. Introduction

In the United States (US), racial and ethnic minorities are particularly disadvantaged, and they fare significantly worse than Whites on almost all health outcomes such as cancer, cardiometabolic disease, infant mortality, and mental health [1]. The deep and pervasive history of structural racism in the US has been revealed as a key determinant of racial and ethnic health inequities [1], [2]. Structural racism refers to "the totality of ways in which societies foster [racial] discrimination, via mutually reinforcing [inequitable] systems (e.g., in housing, education, employment, earnings, benefits, credit, media, health care, economic opportunity, environment, and criminal justice, etc.) that in turn reinforce discriminatory beliefs, values, and distribution of resources", reflected in history, culture, and interconnected institutions, and it is the inherent mechanisms of society that preserve systems of White [2], [3]. The structural racism can promote racial discrimination through various systems (e.g., housing, employment and healthcare, etc.), and exacerbate pre-existing patterns of inequality [4].

Residential segregation refers to the degree to which two or more social groups live apart in different sections of a city or metropolitan statistical area [5]. Racial residential segregation is a type of structural racism, and it is a fundamental cause of racial health disparities [6]. Racial residential segregation shapes socioeconomic opportunity structures, determines access to health promoting resources and services, and constrains individual choices that affect health risks [6], [7]. Previous research reveals that racial residential segregation affects the health outcomes of minority racial groups, which leads to many negative health outcomes. Metropolitan areas with high black–white residential segregation are characterized by isolation of black residents from quality schools, employment, and healthy environments, as well as greater exposure to spatially concentrated poverty, violent crime, and related manifestations of social disorganization when compared to white residents [6], [8]–[12]. Economic segregation is defined as the degree to which people live among others of similar economic status, and it has grown in the US over the past 30 years [13]. Economic segregation influences health outcomes. For example, people living in concentrated poverty have fewer job prospects and worse overall health outcomes, and are exposed to more crime than those living in more economically heterogeneous communities. Economic segregation also influences education and patterns of policy responsiveness, which can affect health outcomes. In the US, richer students often attend one set of schools in richer neighborhoods, while poorer students attend a separate set of schools in poorer neighborhoods. Economic segregation creates a cascading set of effects that shapes how politicians view the effectiveness and efficiency of focusing their resources in a given community as well as their ability to recognize the interests and preferences of those residing in various neighborhoods [14]. Economic residential segregation is somehow a result of economic inequity and is often closely tied to racial residential segregation. Economic segregation operated as a community-level social determinant of health following disasters [15]. Residential racial and economic segregation remain upstream factors that influence access to equitable healthcare, independent of individual-level risk factors including for breast cancer care [16]–[18].

Because of the high correlation between residential segregation and the poverty rate, the poverty rate was transformed to the category variable [19] or analysis was conducted one by one [20]. However, little research has paid attention to leveraging AI in their high correlation to improve residential racial and economic segregation for mortality analysis in the US. To address this problem, we propose AIM, an AI-based method for improving the analysis of the relationship between residential racial and economic segregation and mortality in

979-8-3503-2914-8/24/$31.00 ©2024 IEEE
DOI 10.1109/WFPST58552.2024.00010

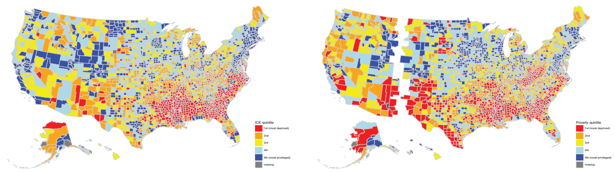
the US.

The outline of this paper is as follows. Section II reviews the related work. Section III presents a brief description of the multiple datasets. Section IV introduces the proposed methods and the design of AIM. Section V presents the results and findings based on the application of AIM on all datasets. Section VI provides concluding remarks.

II. Related Work

ICE was first proposed by Massey [21] to measure spatial social polarisation by creating one formula to categorize all items into five quintiles including the most deprived and the most privileged social groups. In the following two decades, different kinds of ICEs, such as ICE education and ICE income +race/ethnicity [22], were developed and became one of the most commonly used measurement for residential segregation. ICE was evaluated as a useful metric for public health monitoring because it captures the extremes of both privilege and deprivation [23]. Several studies have examined the relationship between residential racial and economic segregation and mortality or COVID-19 vulnerability. [19] examined the association of residential racial and economic segregation with cancer mortality at the US county level, and they found that age-adjusted mortality rates were statistically significantly higher for the most deprived counties for all cancers combined and for 12 of 13 selected cancer sites compared with the most privileged counties, with the largest magnitude occurring with lung and bronchus cancer. Seewaldt and Winn [24] studied the relationship between residential racial and economic segregation and cancer mortality. Brown *et al.* [25] investigated whether racial and economic residential segregation were associated with COVID-19 related factors in the nation's capital, Washington D.C., during the first year of the pandemic. They found that Washington D.C. neighborhoods with a higher concentration of African Americans, lower income residents, and African Americans with low income had a higher incidence of COVID-19 and greater per- cent positivity, but lower testing rates compared to their counterparts [25].

Scholars selected special ICE according to their research interest and the methodology they implemented. Zhang *et al.* [19] used ICE income + race/ethnicity to capture the synergistic effect of both racial and economic segregation and avoid the collineariry issue in the multivariable linear mixed model. Chambers *et al.* [26] applied generalized linear mixed models and discovered the association between ICE income and ICE income + race/ethnicity with preterm birth or with infant mortality experienced by Black women residing in California. Leapman *et al.* [27] utilized ICE race to reveal the racial disparity among older Black and White men with prostate cancer. Most previous research chose ICE quintile, while a few studies have used ICE quarter to analyse the impact of the extremes on health outcomes [28]. In this article we use ICE quintile, but AIM can also apply on ICE quarter to improve residential segregation. The formula of ICE income or ICE income + race/ethnicity requires the information about



(a) U.S. Map by ICE Quintile. (b) U.S. Map by Poverty Quintile.

Fig. 1: U.S. Maps by ICE Quintiles and Poverty Quintile at County Level.

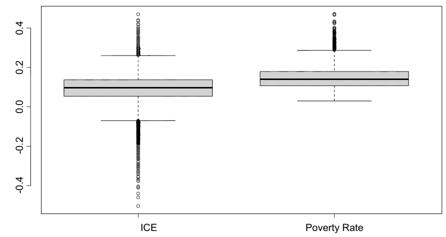


Fig. 2: Boxplot of ICE and Poverty Rate.

household income, and the computation of poverty level also relies on this information. The common calculate feature induces the high correlation between ICEs and poverty rate. As we can handle Big Data and Clouds, AI or ML technology developed some methodologies to overcome the collinearity issue and combine all data sets without losing the crucial information. However, no previous research proposed a good way to improve residential segregation by using the proper AI method based on the characteristics of ICEs and poverty rates. Motivated by the recent AI technology, we propose one new methodology AIM to improve residential racial and economic segregation by utilizing the high correlation between ICE and poverty level.

III. Data Description

A. Dataset 1

We collected the U.S. Census Bureau's 2016-2020 American Community Survey 5-year estimate data to obtain ICEs at county level. In this survey, the estimated number of White householders whose income was equal to or more than \$125,000 in the county i, the estimated number of Black householders whose income was equal to or less than \$25,000 in the county i, and the estimated total population with known income in the county, were explored as A_i, P_i, and T_i, respectively.

B. Dataset 2

The county-level poverty data was explored from Small Area Income and Poverty Estimation (SAIPE) Program by U.S. Census Bureau. We took the average of the estimated proportions of the population in poverty from 2016 to 2020 at county level.

C. Dataset 3

We collected county-level 2016-2020 age-adjusted mortality rate per 100,000 and demographic data from Centers for Disease Control and Prevention (CDC) Wonder online database. Demographic data includes the real number of males, the real number of females, the real number of persons with age 64 or older, the real number of White alone persons, the real number of Black alone persons and the real population in each county.

IV. Methods

ICEs and poverty rates were obtained first at county level, respectively. The cutting points for ICEs were taken based on the average 20^{th} percentile and 80^{th} percentile household incomes from 2016 to 2020 [29]. We calculated ICEs by

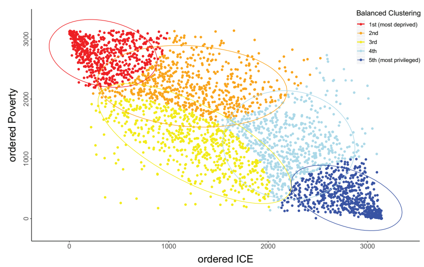


Fig. 3: Balanced Clustering on the Ordered ICE and the Ordered Poverty.

implementing its formula $(A_i - P_i)/T_i$, where A_i is the number of persons in Category A, P_i is the number of persons in Category P, and T_i is the total count of persons belonging to population categorized in relation to the measure. The poverty rates were estimated by SAIPE program directly. ICEs change from -1 to 1, -1 represents in the county i there are only Black persons with household income $\leq$ \$25,000 and 1 represents only White persons with household income $\geq$ \$125,000 live in the county i. The domain of poverty rate is $[0, 1]$ and U.S. Census Bureau uses money income before taxes to determine who is in poverty. It is reasonable to assume that there always exists the negative relationship between ICEs and poverty rate regardless different regions and time periods, and their correlation -0.8 at county level in our study substantiated this assumption. The goal of this study is to utilize their correlation or collinarity, extract the information from poverty, combine the information with the ICEs, and proposal better residential segregation.

After excluding missing values, we obtained 3221 ICEs and 3143 poverty rates. ICE for Valdez-Cordova in Alaska, poverty rates for Kalawao county in Hawaii and all counties in Puerto Rico, ICEs and poverty rates for Shannon county in South Dakota and Wade Hampton in Alaska were not available. Then we ordered ICEs and poverty rates to category each county into one quintile, respectively. In this study, the 1^{st} quintile represented the most deprived counties, and the 5^{th} quintile represented the most privileged counties. We let the 1^{st} quintile contain 645 counties, and each other quintile contain 644 counties by ICEs; we let each of the first three quintiles contain 628 counties, and each other quintile contain 629 counties by poverty rates. Two U.S. maps by ICE quintile and Property quintile were drawn in Figure 1. Compared with U.S. map by ICE quintile in Figure 1(a), U.S. map by Poverty quintile obtains the similar scatter pattern in Figure 1(b): most counties in the first quintile (most deprived) locate in the Southeast and most counties in the fifth quintile (most privileged) are from the Northeast and Northwest. The logic explanation behind this similarity is the high correlation or collinearity between ICEs and poverty rates. The main reason

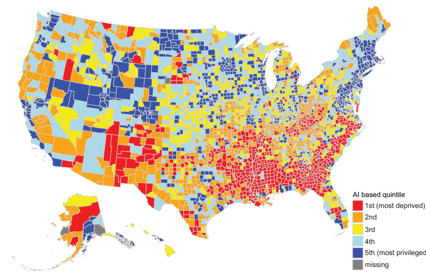


Fig. 4: U.S. Map by AIM Quintile.

TABLE I: Characteristics of U.S. counties by ICE quintiles, 2016-2020.

	1^{st} Quintile	2^{nd} Quintile	3^{rd} Quintile	4^{th} Quintile	5^{th} Quintile	P-value
% Male	49.71	50.22	50.31	50.23	50.03	< 0.0001
% Age 65+	18.61	20.24	20.08	19.77	17.72	< 0.0001
Race/Ethnicity						
% White	63.93	88.38	91.46	92.21	89.85	< 0.0001
% Black	32.64	6.18	4.39	4.19	5.08	< 0.0001
% Poverty	20.06	17.66	14.39	12.10	9.26	< 0.0001

we ordered ICEs and poverty rates instead of nominating them is that as information technology develops the racial and economic differences among counties become small. No ICEs were closed to the extreme number -1 or 1, and no poverty rates were near 1. Figure 2 displays this fact by the distributions of ICEs and poverty rates where most data points of ICEs and poverty rates are between 0 and 0.2. Moreover, all ICEs are between -0.5 and 0.5 and all poverty rates are between 0 and 0.5, so the order method is used for the centralization of data points to distict the small differences among ICEs and poverty rates, respectively.

One machine learning method, Balanced Clustering, was applied to combine ICE quintile and poverty quintile finally to make new AI based residential segregation, which takes the advantages of racial and economic segregation from ICEs and income information from poverty rates. Clustering groups the data points and guarantees the elements in each group have similar patters, but unbalanced clustering tends to form outlier cluster [30]. In addition, the sample size need to be large enough in each ICE quintile and proverty quintile for statistical inference, but clustering concentrates most data points in one or two quintiles, the uneven sample sizes can make the health

TABLE II: Characteristics of U.S. counties by Poverty quintiles, 2016-2020.

	1^{st} Quintile	2^{nd} Quintile	3^{rd} Quintile	4^{th} Quintile	5^{th} Quintile	P-value
% Male	50.39	49.98	50.05	50.07	50.07	0.011
% Age 65+	18.33	19.68	20.24	20.14	18.06	< 0.0001
Race/Ethnicity						
% White	70.06	85.86	90.11	91.89	90.54	< 0.0001
% Black	23.71	10.15	6.35	4.74	4.85	< 0.0001
% Poverty	24.02	17.06	13.98	11.35	8.20	< 0.0001

TABLE III: Characteristics of U.S. counties by AIM quintiles, 2016-2020.

	1^{st} Quintile	2^{nd} Quintile	3^{rd} Quintile	4^{th} Quintile	5^{th} Quintile	P-value
% Male	50.00	50.41	50.05	50.07	50.02	0.0069
% Age 65+	18.69	19.45	20.66	19.84	17.83	< 0.0001
Race/Ethnicity						
% White	67.64	89.19	89.45	91.88	90.32	< 0.0001
% Black	27.46	5.66	7.42	4.42	4.87	< 0.0001
% Poverty	23.25	17.66	12.57	12.70	8.42	< 0.0001

TABLE IV: Average Age-adjusted mortality rate of the U.S. counties by quintiles of ICE, Poverty, and AIM, 2016-2020.

ICE quintile	1^{st} quintile	2^{nd} quintile	3^{rd} quintile	4^{th} quintile	5^{th} quintile	P-value for trend
All cause	917.98 (907.59, 928.37)	829.10 (818.68, 839.53)	780.97 (771.98, 789.96)	744.22 (736.04, 752.41)	664.20 (658.42, 669.98)	< 0.0001
Alzheimer disease	35.19 (33.79, 36.59)	33.26 (31.85, 34.68)	31.68 (30.47, 32.89)	33.89 (32.78, 34.99)	30.17 (29.39, 30.94)	< 0.0001
Covid19	112.95 (107.48, 118.42)	115.45 (109.90, 121.00)	96.31 (91.55, 101.06)	80.46 (76.15, 84.77)	72.47 (69.47, 75.47)	< 0.0001
Poverty quintile	1^{st} quintile	2^{nd} quintile	3^{rd} quintile	4^{th} quintile	5^{th} quintile	P-value for trend
All cause	911.64 (898.68, 924.59)	818.85 (809.37, 828.33)	762.96 (755.10, 770.82)	716.86 (707.94, 725.78)	668.65 (661.31, 675.98)	< 0.0001
Alzheimer disease	35.39 (33.77, 37.02)	33.27 (32.10, 34.45)	34.02 (33.05, 34.99)	30.46 (28.35, 31.57)	29.90 (29.00, 30.81)	< 0.0001
Covid19	140.80 (134.57, 147.04)	100.93 (96.45,105.41)	88.18 (84.49, 91.87)	77.95 (73.74, 82.14)	70.41 (66.97, 73.85)	< 0.0001
AIM quintile	1^{st} quintile	2^{nd} quintile	3^{rd} quintile	4^{th} quintile	5^{th} quintile	P-value for trend
All cause	944.01 (931.91, 956.11)	811.31 (802.04, 820.58)	796.28 (786.68, 805.88)	725.66 (718.34, 732.98)	667.67 (661.38, 673.97)	< 0.0001
Alzheimer disease	35.56 (33.94, 37.18)	34.29 (33.06, 35.52)	32.48 (31.20, 33.76)	31.73 (30.76, 32.70)	30.38 (29.54, 31.21)	< 0.0001
Covid19	127.08 (120.77, 133.39)	107.97 (103.18, 112.76)	89.12 (84.18, 94.07)	83.66 (79.90, 87.41)	71.26 (68.06, 74.47)	< 0.0001

comparison biased. We use Balanced Clustering to prevent a too small or too large number of data points in a cluster and maintain good clustering performance simultaneously in this study [31]. Balanced Clustering requires no missing value, so we take 3142 counties which have both ICEs and poverty rates. The ordered number of ICEs and the ordered number of poverty rates are two dimensions to determine the new AI based quintile. By implementing Balanced Clustering we can overcome the collinearity issue between ICEs and poverty rates and avoid losing poverty information to improve ICE quintile. From Figure 3 we can see that Balanced Clustering considers the properties of both ICEs and poverty rates to regroup all 3142 data points, then we obtain AIM quintiles where the most privilege group represents counties with the relative highest ICEs and the lowest poverty rates and the most deprived group represents the opposite ones. Between the most privilege group and the most deprived group there are three groups, they are built up by minimizing the sum of squared distances to cluster centers as showed in Figure 3.

After implementing Balanced Clustering, we obtained AIM quintile which combined ICEs and Poverty rates to create one new residential segregation. AIM quintile can be viewed as the updated version of ICE quintile because the information of poverty was explored and inputted as one dimension in Balanced Clusterting to adjust ICE quintile. U.S. map by AIM quintile in Figure 4 keeps the most counties in the first quintile(most deprived) induced by Poverty quintile and the most counties in the fifth quintile (most privileged) induced by ICE quintile, which imples that poverty rates provides more information, especially for the fist quintile, and AIM improves ICE by integrating the information.

V. Experiments and Findings

In this section, we provide the experimental results based on multiple datasets, and analyze the relationship between residential racial and economic segregation and mortality in the US.

Table I shows that race/ethnicity, proportion of the population in poverty, gender, and proportion of the population with 65 or older are significantly different in ICE quintiles.

Table II displays the same results except for gender in Poverty quintiles. Race/ethnicity and income are main factors to calculate ICE and income is the threshold to obtain the poverty rate, so it is reasonable to observe p-values for race/ethnicity and poverty rate are significant.

Quintiles of AI based method are used to obtain Table III, which takes the advantages of Table I and Table II. Table III reflects that race/ethnicity and poverty rate are significantly different, but gender is statistically the same among quintiles. This is the result of considering both ICE and poverty rate when balanced clustering is applied in our new AI based method, so gender is not one variable for mortality rate disparity among quintiles by AI based method. To preclude the effect of age on this disparity we use age-adjusted mortality rate per 100,000 to analysis the association between some main causes of death and ICE quintile, Poverty quintile and AIM quintile, respectively.

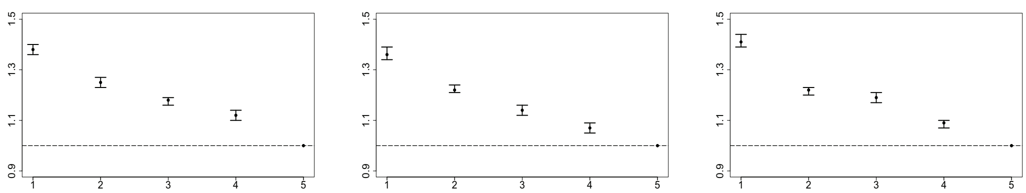
Average age-adjusted mortality rates were obtained at the county level, weighted by county population size for ICE quintile, Poverty quintile and AIM quintile in Table IV. We took the most privileged counties as the reference group to derive age-adjusted mortality rate ratios from average age-adjusted mortality rates in Table V. P values for trend were calculated in the linear regression model with each mortality rate and rate ratio as the dependent variable and continuous ICE quintiles, Poverty quintiles, and AIM quintiles as the independent variable, respectively. 95% confidence intervals are used in Table IV and Table V.

Table IV shows average age-adjusted mortality rates of all causes of death decrease from the most deprived counties to the most privileged counties by ICE quintile, Poverty quintile, and AIM quintile; for Alzheimer disease the decreasing trend of average age-adjusted mortality rates is not clear by ICE quintile and Poverty quintile; for Covid19 the decreasing trend is not clear by ICE quintile while all P values for trend are significant. Table V shows the same result as in Table IV. Figure 5, Figure 6, and Figure 7 exhibit the association between age-adjusted mortality rate ratios with ICE quintile,

TABLE V: Age-adjusted mortality rate ratio of the U.S. counties by quintiles of ICE, Poverty, and AIM, 2016-2020.

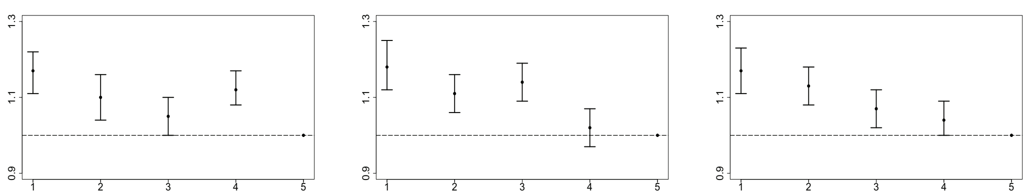
ICE quintile	1^{st} quintile	2^{nd} quintile	3^{rd} quintile	4^{th} quintile	P-value for trend
All cause	1.38 (1.36, 1.40)	1.25 (1.23, 1.27)	1.18 (1.16, 1.19)	1.12 (1.10, 1.14)	< 0.0001
Alzheimer disease	1.17 (1.11, 1.22)	1.10 (1.04, 1.16)	1.05 (1.00, 1.10)	1.12 (1.08, 1.17)	< 0.0001
Covid19	1.56 (1.46, 1.66)	1.59 (1.49, 1.70)	1.33 (1.24, 1.42)	1.11 (1.04, 1.19)	< 0.0001
Poverty quintile	1^{st} quintile	2^{nd} quintile	3^{rd} quintile	4^{th} quintile	P-value for trend
All cause	1.36 (1.34, 1.39)	1.22 (1.21, 1.24)	1.14 (1.12, 1.16)	1.07 (1.05, 1.09)	< 0.0001
Alzheimer disease	1.18 (1.12, 1.25)	1.11 (1.06, 1.16)	1.14 (1.09, 1.19)	1.02 (0.97, 1.07)	< 0.0001
Covid19	2.00 (1.87, 2.13)	1.43 (1.34,1.53)	1.25 (1.17, 1.33)	1.11 (1.03, 1.19)	< 0.0001
AIM quintile	1^{st} quintile	2^{nd} quintile	3^{rd} quintile	4^{th} quintile	P-value for trend
All cause	1.41 (1.39, 1.44)	1.22 (1.20, 1.23)	1.19 (1.17, 1.21)	1.09 (1.07, 1.10)	< 0.0001
Alzheimer disease	1.17 (1.11, 1.23)	1.13 (1.08, 1.18)	1.07 (1.02, 1.12)	1.04 (1.00, 1.09)	< 0.0001
Covid19	1.78 (1.67, 1.90)	1.52 (1.42, 1.61)	1.25 (1.16, 1.34)	1.17 (1.10, 1.25)	< 0.0001

*The 5th quintile (the most privileged counties) was used as the reference group.



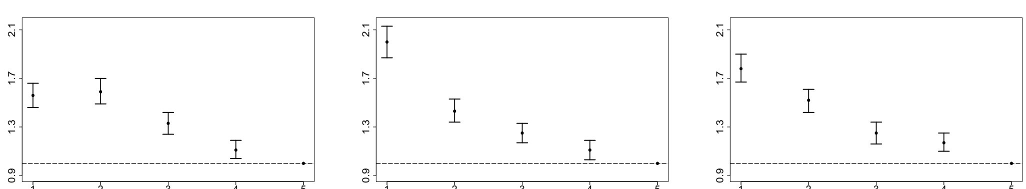
(a) (b) (c)

Fig. 5: (a)Age-adjusted Mortality Rate Ratio for All Causes of Death by ICE quintile; (b)Age-adjusted Mortality Rate Ratio for All Causes of Death by Poverty quintile; (c) Age-adjusted Mortality Rate Ratio for All Causes of Death by AIM quintile.



(a) (b) (c)

Fig. 6: (a)Age-adjusted Mortality Rate Ratio for Alzheimer Disease by ICE quintile; (b)Age-adjusted Mortality Rate Ratio for Alzheimer Disease by Poverty quintile; (c) Age-adjusted Mortality Rate Ratio for Alzheimer Disease by AIM quintile.



(a) (b) (c)

Fig. 7: (a)Age-adjusted Mortality Rate Ratio for Covid19 by ICE quintile; (b)Age-adjusted Mortality Rate Ratio for Covid19 by Poverty quintile; (c) Age-adjusted Mortality Rate Ratio for Covid19 by AIM quintile.

Poverty quintile, and AIM quintile, respectively. All causes of death, Alzheimer disease and Covid19 were chosen on purpose to compare ICE quintile and AIM quintile in this study. For all causes of death ICE quintile and Poverty quintile made the decreasing trend smoothly from the first quintile to the fifth quintile in Figure 5; for Alzheimer disease both ICE quintile and Poverty quintile presented the fluctuated relationship in Figure 6; for Covid19 ICE quintile made one curve while Poverty quintile showed the decreasing trend Figure 7. For those 3 cases AIM quintile kept the decreasing trend, which ascertains AIM improved ICE by extracting the information from poverty rates.

VI. Conclusion

In this paper, we propose AIM, an AI-based method for improving the analysis of the relationship between residential racial and economic segregation and mortality in the US. Instead of using ICE directly, we implement AIM to integrate the information from both ICE and poverty rates. Compared with ICE quintile, AIM quintile shows the clearer decreasing trend for the age-adjusted mortality rates ratio from the most deprived quintile to the most privileged quintile. In addition, the difference by AIM quintile between the most deprived quintile and the most privileged quintile is larger than by ICE quintile. We found a stronger association of the AI based segregation with main causes of death. The racial and ethnic minorities such as Black Americans living in concentrated poverty have a higher mortality rate. Residential racial and economic segregation remain upstream factors that influence access to equitable healthcare. The findings in this paper suggest that policymakers should give attention to initiatives that will protect the health of populations.

Acknowledgment

This research was supported in part by the Faculty Research Awards Program at Florida A&M University.

References

[1] A. Larrabee Sonderlund, M. Charifson, A. Schoenthaler, T. Carson, and N. J. Williams. Racialized economic segregation and health outcomes: a systematic review of studies that use the index of concentration at the extremes for race, income, and their interaction. *PloS one*, 17(1), 2022.

[2] Z. D. Bailey, N. Krieger, M. Agénor, J. Graves, N. Linos, and M. T. Bassett. Structural racism and health inequities in the usa: evidence and interventions. *The lancet*, 389(10077):1453–1463, 2017.

[3] J. R. Feagin. *The white racial frame: Centuries of racial framing and counter-framing*. Routledge, 2020.

[4] Y. Qiu, K. Liao, Y. Zou, and G. Huang. A bibliometric analysis on research regarding residential segregation and health based on citespace. *International Journal of Environmental Research and Public Health*, 19(6):10069, 2022.

[5] D. S. Massey and N. A. Denton. The dimensions of residential segregation. *Social forces*, 67(2):281–315, 1988.

[6] D. R. Williams and C. Collins. Racial residential segregation: a fundamental cause of racial disparities in health. *Public health reports*, 116(5):404–416, 2001.

[7] M. Goodman, S. Lyons, L. T. Dean, C. Arroyo, and J. A. Hipp. How segregation makes us fat: food behaviors and food environment as mediators of the relationship between residential segregation and individual body mass index. *Frontiers in public health*, 6:92, 2018.

[8] P. B. Bach, H. H. Pham, D. Schrag, R. C. Tate, and J. L. Hargraves. Primary care physicians who treat blacks and whites. *N Engl J Med*, 351(6):575–584, 2004.

[9] K. Baicker, A. Chandra, J. S. Skinner, and J. E. Wennberg. Who you are and where you live: How race and geography affect the treatment of medicare beneficiaries: There is no simple story that explains the regional patterns of racial disparities in health care. *Health affairs*, 23(Suppl2):VAR–33, 2004.

[10] T. C. Buchmueller, M. Jacobson, and C. Wold. How far to the hospital?: The effect of hospital closures on access to care. *Journal of health economics*, 25(4):740–761, 2006.

[11] C. A. Collins and D. R. Williams. Segregation and mortality: the deadly effects of racism? In *Sociological forum*, volume 14, pages 495–523. Springer, 1999.

[12] D. R. Williams and P. B. Jackson. Social sources of racial disparities in health. *Health affairs*, 24(2):325–334, 2005.

[13] P. Flavin and W. W. Franko. Economic segregation and unequal policy responsiveness. *Political Behavior*, 42:845–864, 2020.

[14] P. Flavin and W. W. Franko. Economic segregation and unequal policy responsiveness. *Political Behavior*, 42:845–864, 2020.

[15] M. M. Weden, V. Parks, A. M. Parker, L. Drakeford, and R. Ramchand. Health disparities in the us gulf coast: The interplay of environmental disaster, material loss, and residential segregation. *Environmental Justice*, 14(2)(110-123), 2021.

[16] K. White, J. S. Haas, and D. R. Williams. Elucidating the role of place in health care disparities: the example of racial/ethnic residential segregation. *Health Serv Res*, 47(3pt2):1278–1299, 2012.

[17] L. T. Dean, S. Gehlert, M. L. Neuhouser, A. Oh, K. Zanetti, M. Goodman, B. Thompson, K. Visvanathan, and K. H. Schmitz. Social factors matter in cancer risk and survivorship. *Cancer Causes Control*, 29:611–618, 2018.

[18] L. Dean, S. V. Subramanian, D. R. Williams, K. Armstrong, C. Z. Charles, and I. Kawachi. The role of social capital in african–american women's use of mammography. *Soc Sci Med*, 104(148-156), 2014.

[19] L. Zhang, R. Gong, L. Shi, M. Wen, X. Sun, K. R. Yabroff, and X. Han. Association of residential racial and economic segregation with cancer mortality in the us. *JAMA oncology*, 9(1):122–126, 2023.

[20] A. Larrabee Sonderlund, M. Charifson, A. Schoenthaler, T. Carson, and N. J. Williams. Racialized economic segregation and healthoutcomes: A systematic review of studies that usethe index of concentration at the extremes forrace, income, and their interaction. *PLoS ONE*, 17(1):33–36, 2022.

[21] D. S. Massey. *The prodigal paradigm returns: ecology comes back to sociology*. Psychology Press, Patent Country, 2001.

[22] J. M. Feldman, P. D. Waterman, B. A. Coull, and N. Krieger. Spatial social polarisation: using the index of concentration at the extremes jointly for income and race/ethnicity to analyse risk of hypertension. *J Epidemiol Community Health*, 69(12)(12):1199–1207, 2015.

[23] N. Krieger, P. D. Waterman, J. Spasojevic, W. Li, G. Maduro, and G. Van Wye. Public health monitoring of privilege and deprivation with the index of concentration at the extremes. *Am J Public Health*, 106(2):256–263, 2016.

[24] V. L. Seewaldt and R. A. Winn. Residential racial and economic segregation and cancer mortality in the us—speaking out on inequality and injustice. *JAMA oncology*, 9(1)(126-127), 2023.

[25] K. M. Brown, J. Y. Lewis, and S. K. Davis. An ecological study of the association between neighborhood racial and economic residential segregation with covid-19 vulnerability in the united states' capital city. *Annals of Epidemiology*, 59:33–36, 2021.

[26] B. D. Chambers, R. J. Baer, M. R. McLemore, and L. L. Jelliffe-Pawlowski. Using index of concentration at the extremes as indicators of structural racism to evaluate the association with preterm birth and infant mortality-california, 2011-2012. *J Urban Health*, 96(2):159–170, 2019.

[27] M. S. Leapman, M. Dinan, S. Pasha, J. Long, S. L. Washington, X. Ma, and C. P. Gross. Mediators of racial disparity in the use of prostate magnetic resonance imaging among patients with prostate cancer. *JAMA Oncology*, 8(5):687–696, 2022.

[28] M. M. Munir, S. Woldesenbet, L. Alaimo, Z. Moazzam, H. A. Lima, Y. Endo, J. Beane, A. Kim, M. Dillhoff, J. Cloyd, A. Ejaz, and T. M. Pawlik. Mediators of county-level racial and economic privilege in cancer screening. *J Surg Oncol*, 127(7):1212–1222, 2023.

[29] U.S. Census Bureau. Historical inceome tables: Households, 2021.

[30] S. Zhong and J. Ghosh. A unified framework for model-based clustering. *Journal of Machine Learning Research*, 4:1001–1037, 2003.

[31] H. Liu, J. Han, F. Nie, and X. Li. Balanced clustering with least square regression. *Proceedings of the AAAI Conference on Artificial Intelligence*, 31(1), 2017.