medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Segregated Patterns of Racial and Socioeconomic Inclusivity of Access to Hospital Care Among the Medicare Population

Vikas Saini, M.D.[1] and Kelsey Chalmers, PhD.[1,2]

[1] Lown Institute, Needham, MA USA

[2] Menzies Centre for Health Policy, Sydney School of Public Health, University of Sydney, Australia, 6 NSW 2006

Corresponding Author: Vikas Saini, M.D.

vsaini@lowninstitute.org

Lown Institute

163 Highland Ave

Needham, MA 02494

Abstract

Background: A hospital-level metric of inclusion is defined, which measures disparities in hospitalizations by race and socioeconomic status. This measure can help frame research agendas around inequities in health care and guide policy.

The primary objective was to quantify, for each US hospital in our sample, the difference in the patient population and the surrounding area measured by income level, education level, and racial groups. The secondary objective was to assess the associations of these inclusivity metrics with hospital characteristics.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Methods: Cross sectional observational study using MEDPAR claims data for calendar year 2018 involving inpatient admissions to non-specialty, non-federal hospitals of Medicare fee-for-service beneficiaries over the age of 65 years. A community area radius was calculated for each hospital based on their patient counts from surrounding zip codes. The main outcome was a difference metric between the patient demographics and the demographics of the zip codes in the community area. Hospital characteristics associated with the lowest and highest scores were reported using quantile regressions.

Results: There were 4,230,080 patients included in the sample at 3,548 hospitals. The median community area (CA) radius was 24.7 miles, contributing 90.1% of a hospital's total number of patients. Estimated median incomes for hospital patients' zip codes ranged from $26,805 less than their CA zip codes to $29,041 more. Education differences varied by approximately one level of attainment in each direction. The widest percentage differences in racialized groups between hospital patient zip codes and CA codes were for Black and African American populations, which varied from 39.2% greater than the CA to 30.6% less than the CA. Metro areas, larger hospitals and major teaching hospitals were more likely to be found in the highest inclusivity scores and also in the lowest. Some Hospital Referral Regions had wide differences in inclusivity scores, indicating patterns of localized segregation in hospital utilization among Medicare beneficiaries. Safety net hospitals consistently were more inclusive.

Conclusions and Relevance: Race and class inclusivity of US hospitals varies widely, with larger metro areas and teaching hospitals showing the widest patterns of segregation. Safety net hospitals overperform in both urban and rural settings, providing an essential social service of socioeconomic inclusivity.

Introduction

The impact of the United States' long history of socioeconomic exclusion and structural racism is particularly evident in medicine and health care delivery. There are clear differences in rates of insurance

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

coverage, access to facilities and diagnostic tests or treatments, which have impact on a broad range of health outcomes.[1–3] One third of health care spending in the United States is devoted to hospitals, which therefore represent a critical locus for assessing various disparities across populations that have been reported.[4–7]

There have been efforts to develop metrics of inequality in population health but to our knowledge, measures at the individual hospital level are absent.[8] The Index of Dissimilarity, originally applied to the measurement of residential segregation[9,10] has been applied to hospital segregation within geographic regions,[11] but not to individual hospitals. A reproducible and transparent hospital-level metric would shed light on existing disparities, guide policy actions, and monitor the effectiveness of efforts to eliminate them. We sought to develop such a measure, which we have called hospital *inclusivity*, based on readily accessible national administrative data.

Our inclusivity measure addresses disparities in the utilization of hospital care along the lines of class, measured as income level and educational achievement, and race. Structural racism can result in adverse health outcomes for racialized groups via pathways such as residential segregation and access to quality health care.[12] Among the white population, recent declines in life expectancy with a significant gradient based on level of education, have highlighted increasing class disparities in health status.[13]

We first defined a catchment area for individual hospitals (referred hereinafter as the hospital *community area* [CA]), which included areas with a high proportion of inpatient stays and also areas near a hospital with a low proportion of stays which were potentially under-serviced. The inclusivity metrics were designed as a direct comparison of the area demographics of a hospital's inpatients to the demographics of the hospital's CA. Our goal was to detect whether a hospital had lower patient numbers than expected from particular neighborhoods if all neighborhoods within the community had contributed patients

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

equally. We then sought to investigate the hospital characteristics associated with the highest and lowest inclusivity results.

Methods

Data sources and hospital population

We used a 100% sample from the Centers for Medicare & Medicaid Services' (CMS) Chronic Conditions Data Warehouse of Medicare fee-for-service data from the Medicare Provider Analysis and Review (MedPar) table, filed at short-term general or critical access hospitals from 2018. We excluded Medicare Advantage claims and Kaiser Permanente hospitals dominated by Medicare Advantage patients, hospitals not on the 2020 CMS Hospital Compare website,[11] federal hospitals and specialty hospitals. Specialty hospitals were defined as those with more than 40% of their annual admissions as either orthopedic Diagnosis Related Groups (DRG), cardiac DRGs, or admissions with a surgical International Classification of Disease 10th edition code and an elective procedure flag. Hospitals were also excluded if they had fewer than 50 admitted patients in 2018.

We obtained hospital addresses from CMS Hospital Compare[14] (reported using latitude and longitude coordinates).

We counted beneficiaries over the age of 65 from each zip code for every hospital in our sample. Table 1 gives the patient counts and characteristics; for race, we used the Research Triangle Institute race code reported in the Master Beneficiary Summary File.[15]

We used the 2018 American Community Survey (ACS) 5-year data (2009-2018) for the demographic zip code information.[16] For income, we used the zip code-level median household income in the past 12 months (in 2018 inflation-adjusted dollars), where householder is 65 years and over. For education, we

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

used the estimated number of persons where the head of the household is 65 years and older with the following levels: less than 9th grade; 9th to 12th grade; high school diploma; some college (no degree); associate's degree; bachelors' degree; graduate or professional degree. We assigned each of these education levels a value from 1 to 7, and then created an expected education value for each zip code (that is, the sum of each education value multiplied by proportion of persons in the zip code with the education level).

For race, we used the estimated number of persons over 65 in the following 8 categories: American Indian or Alaska Native, Asian, Black or African American, Hispanic or Latino, Native Hawaiian and other Pacific Islander, Other, Two or more races, and White (not Hispanic or Latino). We converted each of these counts to a proportion of zip code's population over 65.

There were 3,492 zip codes with missing median incomes for households where the head of the household was over 65. We replaced 84.5% of these with the county-level equivalent, and for the remaining 540 zip codes we used predictive mean matching to impute these median incomes based on the state, the percentage of people over 65 with a bachelors' degree, and the percentage of people in the state over 65.

Some zip codes in the inpatient claims did not match to an equivalent area in the ACS data (that is, the zip codes were non-residential areas). There was a small amount of zip codes that also had zero estimated persons over 65 in the ACS, despite there being Medicare claims from these areas (this might be due to the different time periods of the ACS and the claim). We excluded hospitals if more than 5% of their beneficiaries were from these zip codes with no ACS data (117 hospitals).

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Community area radius

We calculated the community area radius using cumulative patient counts from zip codes. For each hospital, we sorted zip codes by descending patient counts and then calculated the cumulative sum of patients by these ordered zip codes. Conceptually, we wanted to find the ordered zip code at which the addition of more patients made a negligible difference to the cumulative patient count. We first calculated the sequential difference of these cumulative counts, and then the rolling mean of these using a moving window of three zip codes. We then selected the zip code where this rolling mean estimate was first zero (referred to from here as zip code *A*).

We then calculated the geodesic distance from the hospital address to the centroid of all zip codes where the hospital had patients.[17] We examined all zip codes which had an equal or greater patient count compared to zip code *A*, and selected the maximum distance from the hospital as the community area radius (CA-R). To account for noise and small patient counts from some zip codes, we made one adjustment: we found the minimum distance across all zip codes that had patient counts within ±1 of zip code *A*'s patient count. This distance alone was compared to the distances of zip codes where the patient count was greater than zip code *A*'s patient count, plus one. This prevented a distant zip code with a small count of patients from being selected as the CA-R. Once we found this CA-R, we found the distance between each hospital address and all zip code centroids within this radius, including zip codes with zero patient counts.

Community area demographics

We weighted the demographic statistics from each zip code within the CA-R based on its distance to the hospital. We first found the minimum distance which included at least 50% of the hospital's patients (referred to from here as the 50th percentile radius). Demographic information within this radius was

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

weighted at 100% of the zip codes' demographics. We then found the proportion of the hospital's total patients that were outside the CA (p_{CA}), then calculated a decay rate per mile from 100% at the 50th percentile radius to a final weighting corresponding to the p_{CA} at the CA-R. This decay function gave the weights for the population counts from each zip code.

We calculated the CA income and education scores as the weighted average of income and education values of zip codes within the CA. Equation 1 shows the community area income score CI_H for hospital H, across N zip codes in the community area (each with weighting w_i, population over 65 P_i and median income inc_i).

$$CI_H = \frac{\sum_{i=1}^{N} w_i P_i inc_i}{\sum_{i=1}^{N} w_i P_i} \quad (1)$$

For the CA race scores (equation 2), we found the weighted proportion of each ACS race category (for $k = 1, \ldots, 8$) across all zip codes in the community area.

$$CR_{kH} = \frac{\sum_{i=1}^{N} w_i r_{ki}}{\sum_{i=1}^{N} w_i P_i} \quad (2)$$

Where r_{ki} is the number of persons over 65 in race category k in zip code i.

Hospital patient demographics

The hospital score for income (I_H), education (E_H) and each race category (R_{kH}) are the unweighted averages or proportions across all zip codes where the hospital had patient counts within the CA.

Inclusivity metrics

Our inclusivity metrics for education and income are the log of the ratio of the CA score to the hospital score. For the inclusivity race metric (equation 3), we used a measure similar to Chi-squared difference.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

We gave higher scores to hospitals with disproportionately higher patient counts from zip codes with greater proportions of Black, Native American, Asian, Pacific Islanders, Hispanic, other or more than two races compared to their CA. We did not include the proportion of white people in this equation to avoid the arithmetic cancelling of the differences across the other groups, as well as to give hospitals a higher score for higher proportions of racialized groups.

$$I_{R_H} = \sum_{k=1}^{7} d_{H_k} \frac{\left(R_{H_k} - CR_{H_k}\right)^2}{R_{H_k} + CR_{H_k}} \tag{3}$$

Where d_{H_k} is given by:

$$d_{H_k} = \begin{cases} -1, & R_{H_k} < CR_{H_k} \\ 1, & R_{H_k} \geq CR_{H_k} \end{cases}$$

Composite inclusivity score

The composite inclusivity score was the equally weighted sum of normalized scores for race, income and education. We first normalized these by min-max transformation, and then aligned the values so that zero in the original score mapped to zero in the transformed score. This was because zero in the original scores meant that the hospital and community area demographics are equal; we wanted to preserve this meaning in the composite.

We made one change to the composite score for a small group of hospitals. Within each hospital's community area, we calculated the probability that two randomly selected persons would be the same race (based on the density function of a multivariate hypergeometric distribution). If this probability was greater than 95%, we classed the hospital as being in a *homogeneous* area. For these hospitals (N = 373), we did not include their race score in the composite, and instead took the equally weighted sum of the normalized scores for income and education.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Hospital characteristics

We defined the following hospital characteristics for our comparative analysis: safety net, teaching and financial status, size, geographic region, and core-based statistical area (CBSA). Safety net hospitals were the top 20% of hospitals where the proportion of a hospital's patient stay days were billed as dual eligible. We derived the geographic region from the 2010 Census Regions and Divisions of the United States report. The remaining characteristics were defined using the American Hospital Association (AHA) 2017 dataset. Hospital size was based on bed counts. We labelled hospitals recorded as a Member of Council of Teaching Hospital of the Association of American Medical Colleges as major teaching hospitals; hospitals lacking these memberships, but having a medical school affiliation as reported to the American Medical Association were minor teaching hospitals. For the core-based statistical area (CBSA), metropolitan areas have 50,000 or more people, micropolitan 10,000 to 50,000, and all other areas are rural. Hospitals designated *government, nonfederal* and *nongovernment, not-for-profit* were labelled as non-profit; the remaining category *investor-owned (for-profit)* were for-profit hospitals.

Statistical analysis

We compared the CA-R and the radius of the 50th percentile perimeter across CBSA type and hospital bed-size using a rank sum test. We report the Kendall rank correlation between hospitals' racial, education and income scores within metro, micro and rural areas.

We investigated differences in best and worst scores by hospital characteristics using quantile regression for the 10th and 90th percentile inclusivity scores.[18,19] Rather than estimate a mean response for each hospital characteristic level like in a linear regression, quantile regressions predict the responses at given percentile for each hospital characteristic type, adjusted for other characteristics. For example, a quantile regression at the 50th percentile models the median response for each hospital characteristic type.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

We then inspected the range of inclusivity scores within each Hospital Referral Regions (HRR) and plotted these on a HRR map[20] (eFigure 1). We investigated the HRRs where a hospital's safety net status had the greatest impact on the mean inclusivity score within HRRs using multi-level modelling, with HRR as a random effect.[21]

All tests were two-sided and at the 5% significance level. The analysis took place between February to May, 2021 using SAS Enterprise Guide version 7.15 HF8 and R version 4.0.0 (with data manipulation using *tidyverse* package[22] and plotting using *ggplot*[23]).

This study was approved by the New England IRB, and a waiver of consent was granted as there was no risk to patients and we did not report identifiable data.

Results

Table 1 shows the characteristics of the 3,548 hospitals in our study, with 4,230,080 included patients. The mean age of patients was 77.6 (standard deviation of 8.2), with 3,474,579 white patients (82.1%) and 2,349,447 females (55.5%).

Community area radius

The median CA-R across all hospitals was 24.7 miles (IQR: 17.3 to 37.2). Micropolitan areas (with 10,000 to 50,000 people) had the largest median CA-R (29.1 miles [IQR: 21.7 to 41.6]) compared to metro and rural areas (Table 2)). Metro-area hospitals, however, had the largest 50th percentile radius of their hospital patients (median 6.3 miles [IQR: 4.0 to 9.6]), followed by micro- and rural area hospitals. Both of these radius measures also varied by hospital size, with extra-large hospitals having the greatest CA-R (median 34.8 miles [IQR: 21.6 to 55.1]) and extra small hospitals having the smallest (median 21.5 miles [IQR: 16.2 to 28.6]).

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

The mean proportion (± standard deviation) of hospital patients within the CA-R was 87.6% ± 8.7 (Median 90.1%, IQR: 84.9 to 93.2). There were similar results for hospitals in metro and rural areas, but hospitals in micro areas had CA-R containing higher proportions of total patients (eTable 1, eFigure 2).

Inclusivity scores

The composite inclusivity scores were designed so that zero represented hospital demographics that matched community demographics, with positive and negative values reflecting greater or lesser inclusivity. The median composite inclusivity score was 1.03 points (IQR: [-0.67, 3.49]). The median score for income inclusivity was 1.93 points (IQR [-1.23, 5.74]), education inclusivity was 1.05 points (IQR [-0.49, 3.24]) and for racial inclusivity was -0.07 points (IQR [-0.58, 0.04]).

Figure 1 shows the differences between the CA and hospital's patient demographics, as well as the hospitals' inclusivity metrics. The summary statistics of these demographics are given in the online supplement in eTable 2. The maximum positive difference between the median income estimate was $26,805 (a 100 times log-ratio of 66.99), and the maximum negative difference was -$29,041 (a 100 times log-ratio of -54.45). For education, the maximum positive and negative differences between the median education levels were roughly 0.95 and 1.0 (a 100 times log-ratio of 24.37), approximately equivalent to the difference between two adjacent ACS education levels (for example, *some college (no degree)* to *associate's degree*). Among racialized groups, Native Hawaiian or Pacific Islander populations had the smallest differences in hospital inclusivity ranging from 4.29% greater to 0.5% less than the community. The largest differences were seen for the Black and African American populations, where a hospital had a maximum difference of 39.2% greater than its CA (hospital was 84.2% while CA was 44.9%, meaning a contribution to its racial inclusivity score of 11.90 points) while another had 30.6% less than its CA (hospital was 9.0% while CA was 39.6%, the racial inclusivity score contribution being -19.3 points).

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Correlations of inclusivity scores

All three component inclusivity scores were significantly and positively correlated for hospitals within metro areas (Figure 2). The correlation between metro-area hospitals' income and education scores (rank correlation $R = 0.60$; $p < 0.001$) was notably higher than the correlation between income and race ($R = 0.24$; $p < 0.001$), and education and race ($R = 0.20$; $p < 0.001$). In micro and rural areas, racial scores were not correlated with either income or education, whereas income and education inclusivity remained correlated (micro $R = 0.47$; $p < 0.001$; rural $R = 0.29$; $p < 0.001$).

The association of inclusivity with hospital characteristics

eFigure 3 shows the distribution of inclusivity scores for each hospital characteristic. Figure 3 shows the unadjusted 10th to 90th percentile variances of the inclusivity scores for each hospital characteristic. Metro areas, major teaching hospitals and large hospitals had much wider ranges of inclusivity scores. Here, we report the results for the associations between hospital characteristics and the highest and lowest composite inclusivity scores (see online supplement for income, education and racial inclusivity score results). Quantile regression plots of the estimated coefficients are shown in Figure 4 and in the online supplement in eFigures 4 to 6, with the adjusted 10th, 50th and 90th percentile estimates shown in eTable 4.

Quantile regressions confirmed that the hospital characteristics associated with the highest composite inclusivity scores were metro-area hospitals, major teaching hospitals, and safety net hospitals, while the characteristics with the lowest inclusivity scores were metro-area hospitals, major teaching hospitals and non-safety net hospitals. Both of the 10th percentile coefficients for rural versus metro-area hospitals (2.77 95% CI: [1.78, 3.76]) and micro versus metro-area hospitals (3.46 [2.47, 4.45]) were large and positive, meaning that metro-area hospitals had lower estimated 10th percentile inclusivity scores compared to rural and micro-area hospitals. At the 90th percentile, metro-area hospitals had higher estimated coefficients in

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

comparison to rural (3.4 points greater than rural hospitals [adjusted 2.7, CI: 1.6 to 3.7; $p < 0.001$]) and micro-area hospitals. Similarly, major teaching hospitals were associated with a greater range of the composite score compared to minor teaching and non-teaching hospitals, with no significant differences across median scores (Figure 4). Safety net versus non-safety net hospitals had positive coefficients for all percentiles in the quantile regressions.

The online supplement also shows the quantile regression results for income (eFigure 4), education (eFigure 5) and the racial inclusivity scores (eFigure 6). Major teaching hospitals had the best and worst scores on income, the best scores for racial inclusivity, and the worst scores for education. The only difference between non-profit and for-profit hospitals was found at the 90th percentile of education inclusivity, where the for-profit hospital score was 1.3 points greater than non-profit hospitals (adjusted difference of 1.0, CI: 0.3 to 1.8, $p = 0.004$). Safety net hospitals outperformed at the low and high end of racial inclusivity scores, with no significant role in the middle.

Inclusivity by hospital referral regions

The difference between the maximum and minimum inclusivity within an HRR ranged from a minimum difference of 0.05 points (Wilkes-Barre, Pennsylvania) to a maximum of 81.5 points (Manhattan, New York). eFigure 1 shows a US map of these differences, with the largest ranges clearly visible in metropolitan areas.

From the multi-level regression of HRR and safety net status, there were 21 HRRs where the slope coefficient between non-safety net and safety net hospitals were significantly greater than zero (eFigure 7). Most hospitals within these HRRs were in metropolitan areas; 15 HRRs only had metro-area hospitals, while the rest ranged from 94.6% (Boston, MA; 2 non-metro hospitals out of 37) to 56.5% (Tulsa, OK; 10 non-metro hospitals out of 23).

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Within rural areas, safety net hospitals had higher inclusivity scores for income (median 0.7 [IQR: -1.7 to 3.6] for non-safety nets versus a median of 2.0 [IQR: -0.3 to 5.1] for safety nets) and education (median 0.5 [IQR -0.4 to 1.7] for non-safety nets versus a median of 0.9 [IQR -0.1 to 2.3]) (eTable 3). There was no difference between safety net and non-safety net hospitals in the racial inclusivity score (both had a median of 0 [IQR: -0.2 to 0.0]).

Discussion

In the midst of rising concerns about persistent injustices of race and class, there have been recent calls for metrics with greater clarity to ensure that measurement promotes accountability and lasting change.[24] With this novel metric of racial and economic inclusivity of hospitals, we sought to measure hospitals' service to their surrounding communities. Our approach was to measure the deviance between two populations: the theoretical counts from hospitals' community area, and the actual patient counts from these zip codes. The theoretical counts are based on the assumption that patients in surrounding areas are equally likely to be admitted to a hospital, with this likelihood decreasing only by distance from the hospital to the zip code. Hospitals are more *inclusive* if they have a higher patient population of racialized groups, lower income patients and lower educated patients compared to the community area population. We devised our metric so that, in an ideal world, all hospitals would have a patient population which matched their community area and would have an inclusivity score of zero.

The wide range of inclusivity results provide insight into the local distribution of racial and socioeconomic differences in hospital utilizations amongst Medicare beneficiaries. While the median inclusivity scores did not vary substantially by hospital characteristics (other than safety net status), the range of hospital inclusivity scores did, and quantile regressions showed that inclusivity at the extremes was associated with particular types of hospitals. The range of differences of hospital inclusivity within

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

HRRs also varied nationally, with the widest differences within HRRs reflecting patterns of local segregation in hospital utilization among Medicare beneficiaries.

Our measure of inclusivity is multidimensional and included education, income and race, because the intersections of these may be important in understanding patterns of exclusion. While income and education were treated as continuous variables reflecting economic class, race was categorical. Our racial inclusivity metric incorporates the difference in proportions for multiple racialized groups. For most hospitals receiving a negative inclusivity score, this meant they had higher counts of patients from zip codes with high proportions of white people. However, there were some hospitals that received a low inclusivity score despite having patient counts of one marginalized group higher than their community because their patient counts from other minority groups were lower than in their community. These differences are visible in Figure 1, where (unlike income and education) the gradient of racial inclusivity scores does not directly map from highest to lowest differences of white patient counts. To allow the greatest insights into local market profiles, we did not calculate a summary measure of inclusivity combining all racial minorities, though the method would permit it. Instead, we sought a level of granularity that would enable the use of this metric in studies of the crosscurrents of race and class – racial subgroups interacting with income or education. Our goal was to enable enough flexibility for analysis across subgroups and dimensions, so that it would be possible to assess, for example, the inclusivity of hospitals for poorer white populations, or higher income Hispanics.

The lower correlations between race and income/education inclusivity may arise from this granularity. Hospitals having one minority race in really high proportions, but not the others could have lower racial inclusivity without affecting income inclusivity. The low correlations could also reflect underlying income and educational heterogeneity within racial categories in America; that is, the varying educational and income inequality between Black, Asian, and the Hispanic population, versus the white population.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Our core finding that disparity of inclusivity scores is greatest in metro areas, where both the best and worst performers are found, is not surprising. Metro areas are far more diverse, creating more 'opportunity' for a dispersion of inclusivity. Of course, racial and economic diversity need not inherently lead to variation in hospital inclusivity, but the fact that it does underscores the manner in which health care utilization mirrors the many socioeconomic factors that determine it in the background.

Safety net status was associated with higher inclusivity scores than non-safety net, which one could consider tautological given our definition of safety net status was the hospitals with the highest proportion of dual eligible patients. However, this safety net definition is based on only one axis, income (or dual eligible status), as opposed to inclusivity which has a second dimension, geography. A safety net hospital would have an income inclusivity score of zero if they were serving solely an area with low-income zip codes. Safety net hospitals will have high inclusivity scores if they selectively serve the low-income zip codes within a CA that has both high- and low-income zip codes. Inclusivity thus measures biased selection from within a geographical area, measuring social stratification across areas. Our analysis indicates that safety net hospitals are providing a substantial social service of racial and socioeconomic inclusivity, in both more diverse metropolitan areas and (for income and education) in more racially homogeneous, rural areas. As such, safety net hospitals appear to be integral and complementary to the less inclusive hospitals in their communities as the two types together give the whole population, hopefully, adequate access to hospital care.

Inclusivity scores reflect the many current and historical factors contributing to these population divisions. These include residential segregation, itself coupled to the historical sequence of immigration, along wit labor mobility, linkages of insurance coverage to differing employment conditions and associated partitions in the labor and health insurance markets. Hospital-level measures of inclusivity as reported here may enable further investigation of these additional associations.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Our inclusivity scores are driven by the difference between total CA demographics and demographics from a hospital's patients' zip codes within this CA. As such they depend on the definition of the CA-R. Instead of a fixed radius method,[25] we used a method similar to the variable radius methods of previous analyses of hospital market areas.[26] Ours differed in that it was defined by the distribution characteristics of each hospital rather than by a single cut-off for a proportion of patients. Nevertheless, our CA-R encompassed on average 87.6% of a hospital's patients (median 90.1%) similar to the defined radius used in previous work.[27] Our finding that the radius did not vary significantly between metro, micro and rural hospitals is not consistent with the findings reported by Gresenz and colleagues.[27] This may be due to our use of CBSA definitions of population areas rather than quintiles of population density. Also, because radius covaried as a function of hospital size and since rural hospitals do not have large hospitals (eFigure 2), this would have the effect of reducing rural radius and negating its differences with metro areas. Micro areas on the other hand tended to have a larger CA-R as well as greater proportions of their total patients within that radius, suggesting that hospitals in these intermediate population zones may serve a particular function as magnets for rural area patients.

Since we used Medicare data alone, our results will not reflect patients' differing choice of hospitals driven by differing insurance coverage nor will they reflect revenue-seeking marketing behavior on the part of hospital managers towards patients with more lucrative commercial insurance coverage. If current insurance coverage is not a driver of the differences in hospital utilization, why would Medicare patients choose certain hospitals? The patterns we report are more likely reflections of embedded habits of patients and hospitals, driven by such issues as trust,[28] the communities a hospital seeks to engage, patterns of patient preference and utilization dating to an individual's pre-Medicare insurance status,[29] rates of supplemental insurance,[30] hospital financial aid and collection policies, other legacy behaviors, and other factors such as the number of Black or minority doctors.[31]

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Communities of color and working-class communities may prefer certain hospitals, so do hospitals control who walks in their door? The drivers of the disparities in inclusivity we report here are likely multifactorial. Factors to consider would include community engagement, staffing decisions, advertising, investment decisions, acquisition of provider networks, the placement of clinics, et cetera. Any policy initiatives to reduce disparities in inclusivity will likely need to be similarly multipronged at both local and systemic levels. Policy makers will have to address these challenges creatively and a metric of inclusivity may help to track trends and assess the effects of interventions.

Limitations

This study is restricted to Medicare fee-for-service patients only, therefore the application of these results to other populations such as Medicare Advantage patients cannot be made with confidence. Furthermore, the application of the method to commercial data might also yield different results.

We used CBSA definitions of population areas, which are relatively imprecise. Had we used finer gradations of population density, it is possible that our results, particularly for urban versus rural differences in radius calculations may have been altered.

Our calculations depend on the assumption that everyone in a particular zip code acts in the same way, with the same propensity to use a hospital as any other person within the zip code. This may not be valid. This also means that socioeconomic and cultural changes within zip codes, such as gentrification of poorer neighborhoods, are not considered.

Another limitation is that we did not use demographic data at the patient level, but at the zip code level (that is, the summary values from the American Community Survey). While patient-level data on race exists in the Medicare files, we decided to use zip code counts to be consistent with the income and education values (which were only available at a zip code level).

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

We did not use travel time to define community area radius, so our results using a simple geographic radius do not account for the ease of access to hospital care as a determinant of catchment area and of inclusivity. While driving distance may give a more accurate estimate for some hospitals, particularly those on shore lines or with other geographic barriers, Boscoe et al. showed that travel time, driving distance and straight-line distances were overall highly correlated.[32]

Conclusion

We present here a novel method of quantifying socioeconomic and racial inclusion in US hospitals that provides insights into important patterns of health care utilization. Inclusivity, the degree to which a hospital's patients' characteristics by race and socioeconomic status reflect its community area, varied widely. We found that inclusivity varied most in metro areas and among major teaching hospitals, revealing patterns of local segregation. Safety net hospitals' inclusivity scores show they serve an important social function that is evident in both urban and rural settings.

The metric of inclusivity holds promise for allowing identification of other specific associations, but more research is required on its broader utility. It represents an objective, transparent method of measuring one specific aspect of inequities in health care and provides an opportunity to pursue new avenues of research on health care disparities. Inclusivity assessed as a time series may allow an assessment of historical trends as well as providing a tool for the measurement of the effects of policy actions designed to mitigate structural racism and economic class disparities in US hospital care.

Acknowledgements: The authors thank Paula Smith and Valérie Gopinath for assistance in data collection and preparation and Carissa Fu for assistance with manuscript preparation.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

References

1. Riley P, Hayes SL, Ryan J. Closing the equity gap in health care for Black Americans. To the Point, The Commonwealth Fund. Published 2016. Accessed May 11, 2021. https://www.commonwealthfund.org/blog/2016/closing-equity-gap-health-care-black-americans

2. Artiga S, Orgera K, Damico A. Changes in Health Coverage by Race and Ethnicity since the ACA, 2010-2018. KFF. Published March 5, 2020. Accessed May 11, 2021. https://www.kff.org/racial-equity-and-health-policy/issue-brief/changes-in-health-coverage-by-race-and-ethnicity-since-the-aca-2010-2018/

3. Geiger HJ. Racial and ethnic disparities in diagnosis and treatment: a review of the evidence and a consideration of causes. Published online 2003. https://www.ncbi.nlm.nih.gov/books/NBK220337/

4. Cronin KA, Howlader N, Stevens JL, Trimble EL, Harlan LC, Warren JL. Racial Disparities in the Receipt of Guideline Care and Cancer Deaths for Women with Ovarian Cancer. *Cancer Epidemiol Biomarkers Prev*. 2019;28(3):539-545. doi:10.1158/1055-9965.Epi-18-0285

5. Burton BN, Munir NA, Labastide AS, Sanchez RA, Gabriel RA. An Update on Racial Disparities With 30-Day Outcomes After Coronary Artery Bypass Graft Under the Affordable Care Act. *J Cardiothorac Vasc Anesth*. 2019;33(7):1890-1898. doi:10.1053/j.jvca.2018.10.021

6. Ali I, Vattigunta S, Jang JM, et al. Racial Disparities are Present in the Timing of Radiographic Assessment and Surgical Treatment of Hip Fractures. *Clinical Orthopaedics and Related Research*®. 2020;478(3):455-461. doi:10.1097/corr.0000000000001091

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

7. Sanei-Moghaddam A, Kang C, Edwards RP, et al. Racial and Socioeconomic Disparities in Hysterectomy Route for Benign Conditions. *Journal of Racial and Ethnic Health Disparities*. 2018;5(4):758-765. doi:10.1007/s40615-017-0420-7

8. Krieger N, Waterman PD, Spasojevic J, Li W, Maduro G, Van Wye G. Public Health Monitoring of Privilege and Deprivation With the Index of Concentration at the Extremes. *Am J Public Health*. 2016;106(2):256-263. doi:10.2105/ajph.2015.302955

9. Duncan OD, Duncan B. A Methodological Analysis of Segregation Indexes. *American Sociological Review*. 1955;20(2):210-217. doi:10.2307/2088328

10. James DR, Taeuber KE. Measures of Segregation. *Sociological Methodology*. 1985;15:1-32. doi:10.2307/270845

11. Tikkanen RS, Woolhandler S, Himmelstein DU, et al. Hospital Payer and Racial/ethnic Mix at Private Academic Medical Centers in Boston and New York City. *International Journal of Health Services*. 2017;47(3):460-476. doi:10.1177/0020731416689549

12. Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. *Lancet*. 2017;389(10077):1453-1463. doi:10.1016/s0140-6736(17)30569-x

13. Case A, Deaton A. Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. *Proceedings of the National Academy of Sciences*. 2015;112(49):15078-15083.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

14. US Department of Health and Human Services. Hospital Compare Federal Dataset. HealthData.gov. Accessed February 10, 2021. https://healthdata.gov/dataset/hospital-general-information/resource/6cb04f57-b05b-4e74-8a0f-e7eb04fdf59f

15. Eicheldinger C, Bonito A. More accurate racial and ethnic codes for Medicare administrative data. *Health care financing review*. 2008;29(3):27.

16. US Census Bureau. American Community Survey 5-Year Data (2009-2019). The United States Census Bureau. Accessed May 11, 2021. https://www.census.gov/data/developers/data-sets/acs-5year.html

17. Karney CFF. Algorithms for geodesics. *J Geod*. 2013;87:43-55. doi:10.1007/s00190-012-0578-z

18. Das K, Krzywinski M, Altman N. Quantile regression. *Nat Methods*. 2019;16(6):451-452. doi:10.1038/s41592-019-0406-y

19. Koenker R. Quantile Regression [R package quantreg version 5.85]. Published February 24, 2021. Accessed May 19, 2021. https://CRAN.R-project.org/package=quantreg

20. National Bureau of Economic Research. Dartmouth Atlas Data - Geography and Coding Trend Files -- Crosswalks and Boundary Files. NBER. Published 2013. Accessed May 19, 2021. https://www2.nber.org/data/dartmouth-atlas-geography.html

21. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*. 2014;67(1):1-48.

22. Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. *JOSS*. 2019;4(43):1686. doi:10.21105/joss.01686

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

23. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Published online 2016. https://ggplot2.tidyverse.org

24. Givens ML, Gennuso KP, Pollock EA, Johnson SL. Deconstructing Inequities — Transparent Values in Measurement and Analytic Choices. *New England Journal of Medicine*. 2021;384(19):1861-1865. doi:10.1056/NEJMms2035717

25. Garnick DW, Luft HS, Robinson JC, Tetreault J. Appropriate measures of hospital market areas. *Health Serv Res*. 1987;22(1):69-89.

26. Phibbs CS, Robinson JC. A variable-radius measure of local hospital market structure. *Health Serv Res*. 1993;28(3):313-324.

27. Gresenz CR, Rogowski J, Escarce JJ. Updated Variable-Radius Measures of Hospital Competition. *Health Serv Res*. 2004;39(2):417-430. doi:10.1111/j.1475-6773.2004.00235.x

28. Lopes L, Muñana C, 2020. KFF/The Undefeated Survey on Race and Health. KFF. Published October 13, 2020. Accessed May 15, 2021. https://www.kff.org/racial-equity-and-health-policy/report/kff-the-undefeated-survey-on-race-and-health/

29. Decker SL, Doshi JA, Knaup AE, Polsky D. Health Service Use among the Previously Uninsured: Is Subsidized Health Insurance Enough? *Health Econ*. 2012;21(10). doi:10.1002/hec.1780

30. Brunt CS. Supplemental Insurance and Racial Health Disparities under Medicare Part B. *Health Serv Res*. 2017;52(6):2197-2218. doi:10.1111/1475-6773.12599

31. Alsan M, Garrick O, Graziani GC. *Does Diversity Matter for Health? Experimental Evidence from Oakland*. National Bureau of Economic Research; 2018. doi:10.3386/w24787

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

32. Boscoe FP, Henry KA, Zdeb MS. A Nationwide Comparison of Driving Distance Versus Straight-Line Distance to Hospitals. *Prof Geogr*. 2012;64(2). doi:10.1080/00330124.2011.583586

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Tables

Table 1. The characteristics of hospitals and patients in our sample.

Hospital characteristic	**N = 3,548**
CBSA area type	
Metro	2,207 (62%)
Micro	605 (17%)
Rural	736 (21%)
Ownership type	
For-profit	529 (15%)
Non-profit	3,019 (85%)
Teaching class	
Major	225 (6.3%)
Minor	975 (27%)
Non-teaching	2,348 (66%)
Hospital type	
Acute Care Hospitals	2,675 (75%)
Critical Access Hospitals	873 (25%)
Safety net status	
Non-safety net	2,840 (80%)
Safety net	708 (20%)
Bed size	
Extra small	1,028 (29%)
Small	541 (15%)
Medium	796 (22%)
Large	747 (21%)

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Extra large	436 (12%)
Patient characteristics	**N = 4,230,080**
Age (Mean, SD)	77.7 (8.2)
Race*	
Non-Hispanic white	3,474,579 (82.1%)
Black or African-American	365,132 (8.6%)
Asian/Pacific Islander	84,267 (2.0%)
Hispanic	212,282 (5.0%)
American Indian or Alaska Native	19,562 (0.5%)
Other	27,827 (0.7%)
Unknown	46,431 (1.1%)
Sex	
Female	2,349,447 (55.5%)

*Race reported using Research Triangle Institute codes from MedPar

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Table 2. Community area radius and the radius (miles) of the 50th percentile perimeter by the core based statistical area (CBSA) of the hospitals and the hospital size.

	Community area radius				**Radius with 50% of patients**			
	Median (IQR)	**Mean (SD)**	**Range**	**p-value[1]**	**Median (IQR)**	**Mean (SD)**	**Range**	**p-value[1]**
Overall	24.7 (17.3, 37.2)	30.7 (21.9)	0.8, 276.4		5.9 (2.8, 9.6)	7.3 (6.7)	0.1, 66.8	
CBSA Type								
Metro	24.1 (16.5, 38.8)	31.3 (24.2)	0.8, 276.4	<0.001	6.3 (4.0, 9.6)	7.9 (6.7)	0.1, 66.8	<0.001
Micro	29.1 (21.7, 41.6)	34.0 (19.1)	6.4, 139.3		5.2 (1.6, 9.9)	6.8 (6.5)	0.1, 44.3	
Rural	23.4 (17.3, 31.1)	26.2 (15.5)	2.6, 244.1		3.9 (0.9, 9.3)	5.9 (6.5)	0.1, 61.3	
Hospital bed size								
Extra small	21.5 (16.2, 28.6)	24.3 (14.4)	2.6, 244.1	<0.001	3.1 (0.9, 8.2)	5.2 (5.8)	0.1, 61.3	<0.001
Small	25.1 (18.6, 33.8)	28.3 (14.9)	0.8, 112.9		5.5 (2.1, 9.0)	6.5 (5.5)	0.1, 34.1	
Medium	25.7 (17.8, 38.3)	30.6 (20.1)	2.1, 164.1		5.7 (3.6, 9.1)	7.0 (5.2)	0.1, 39.4	
Large	27.4 (16.9, 45.4)	35.1 (28.5)	1.8, 276.4		6.4 (4.3, 9.8)	8.2 (6.3)	0.4, 50.2	
Extra large	34.8 (21.6, 55.1)	41.3 (27.9)	3.7, 182.8		9.6 (6.6, 15.0)	12.4 (9.5)	1.2, 66.8	

[1]Kruskal-Wallis rank sum test

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

FIGURES

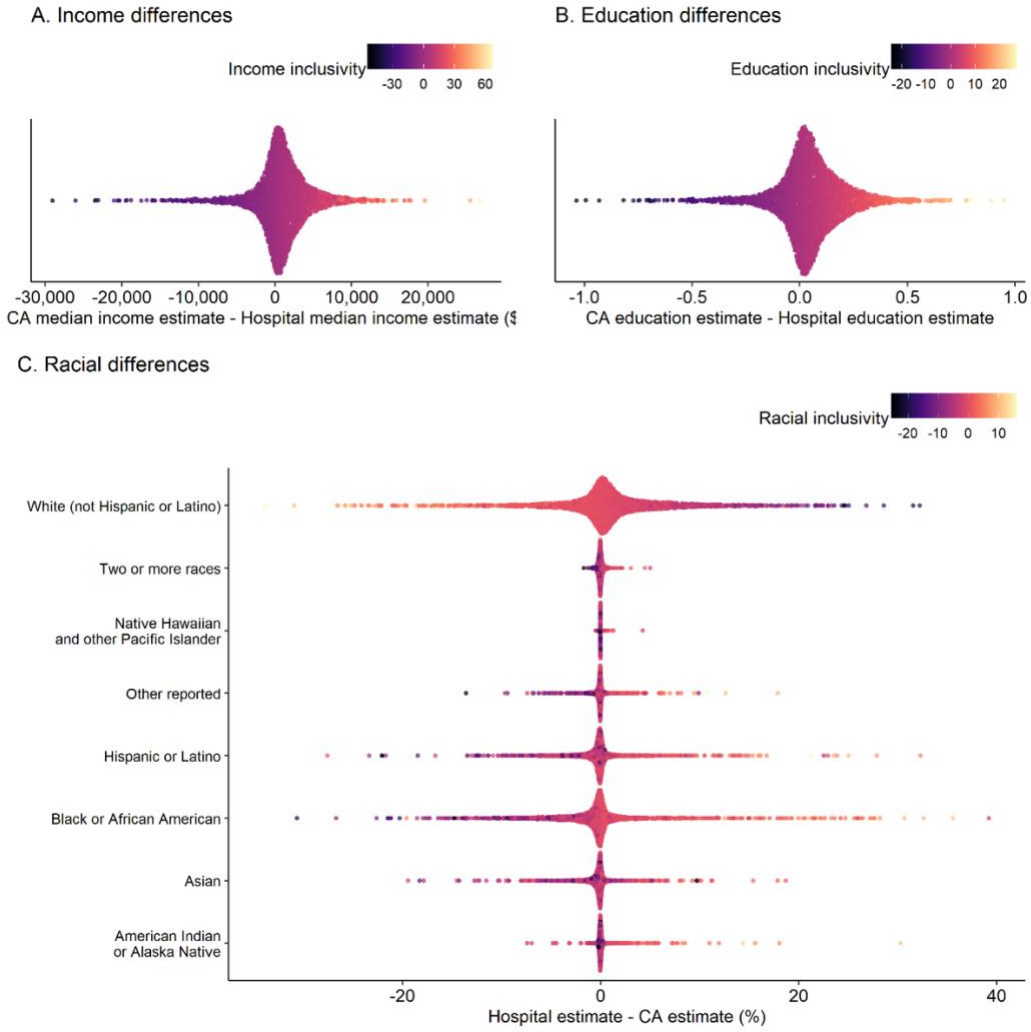

Figure 1.

Individual hospital results of the differences between the hospital demographic estimates and the community area (CA) estimates for A. median income, B. education value and C. percentage of each racialized group. Points are shown for every hospital (N = 3,548), with the relative density of these points shown on the y-axis. Hospital points are colored based on each inclusivity score metric.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

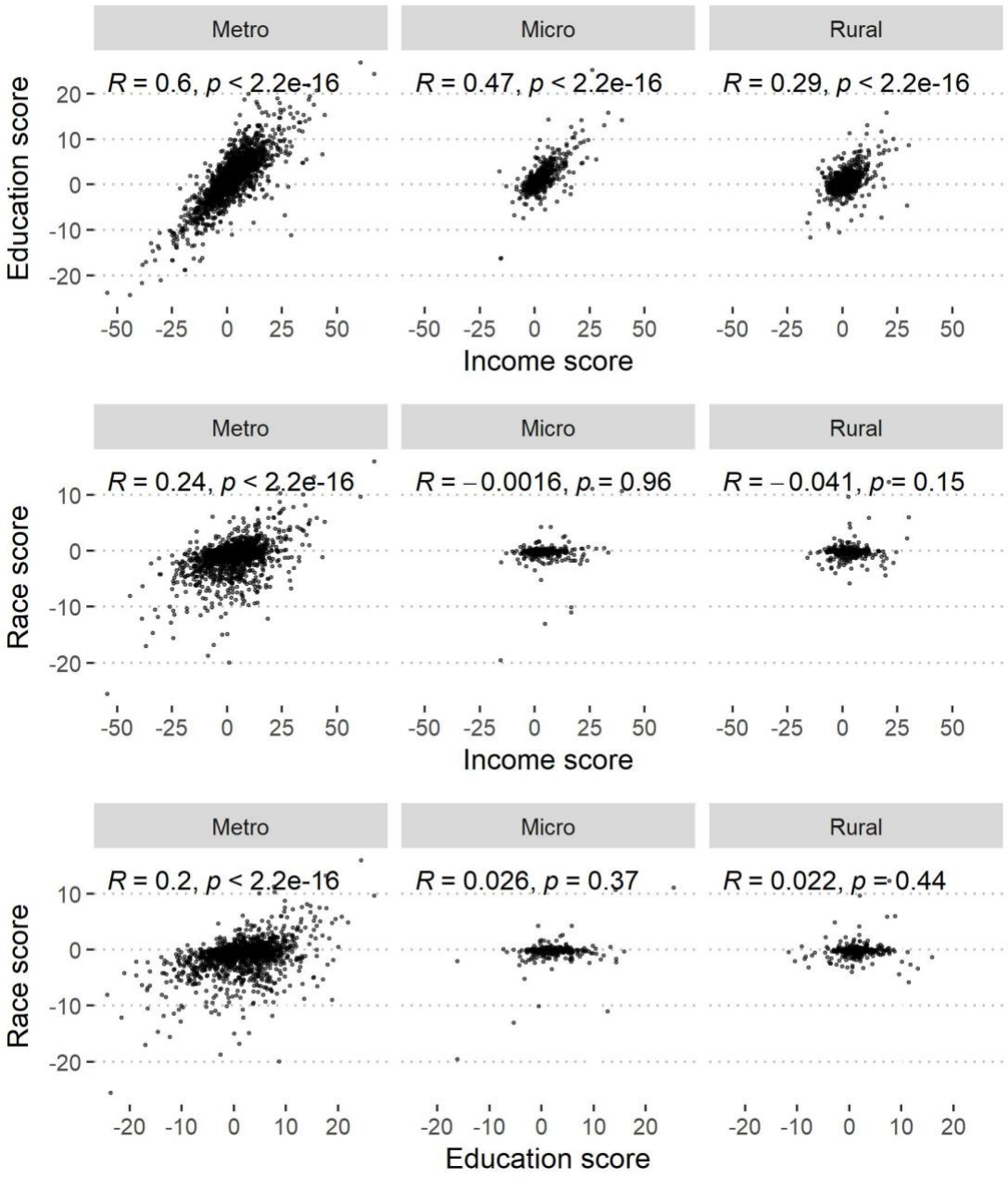

Figure 2. Comparison of the three inclusivity scores across hospitals: income, education and race, along with the histograms of each score. The Kendall rank correlation coefficient and p-value for the statistical test that this correlation is not zero between each variable pair is shown.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

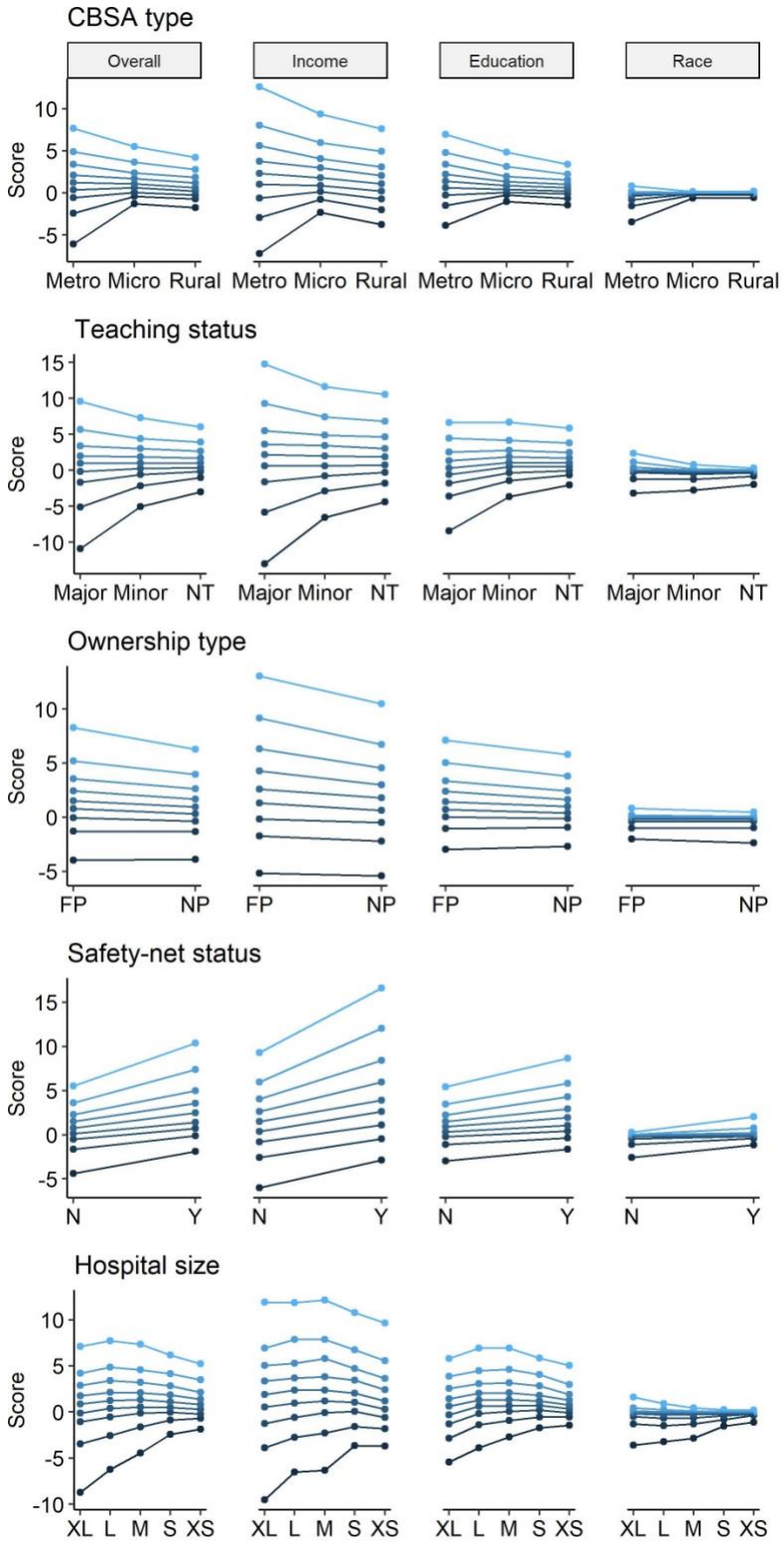

Figure 3. The overall, income, education and race inclusivity scores by hospital characteristic. Points show the 10th to 90th percentile scores for each hospital characteristic level, with lines drawn between the respective percentiles for each hospital characteristic level.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.24.21257551; this version posted May 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

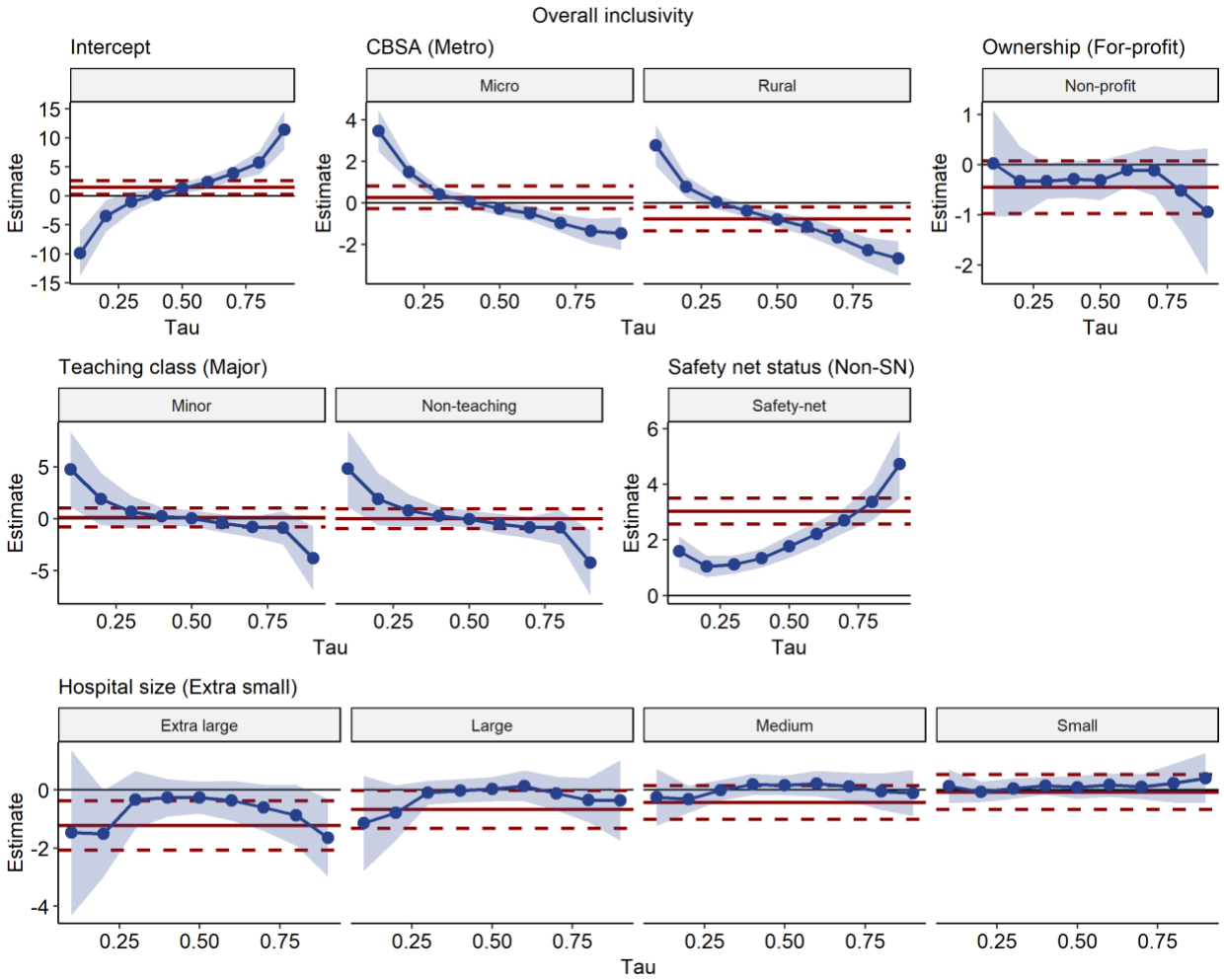

Figure 4. Quantile regression results for overall inclusivity scores.