JAMA Internal Medicine | Original Investigation | HEALTH EQUITY

Attributing Racial Differences in Care to Health Plan Performance or Selection

Jacob Wallace, PhD; Chima D. Ndumele, PhD; Anthony Lollo, PhD; Danil Agafiev Macambira, BSc; Matthew Lavallee, BA; Beniamino Green, MSc; Kate A. Duchowny, PhD, MPH; J. Michael McWilliams, MD, PhD

IMPORTANCE There is increased interest in public reporting of, and linking financial incentives to, the performance of organizations on health equity metrics, but variation across organizations could reflect differences in performance or selection bias.

OBJECTIVE To assess whether differences across health plans in sex- and age-adjusted racial disparities are associated with performance or selection bias.

DESIGN, SETTING, AND PARTICIPANTS This cross-sectional study leveraged a natural experiment, wherein a southern US state randomly assigned much of its Medicaid population to 1 of 5 plans after shifting to managed care in 2012. Enrollee-level administrative claims and enrollment data from 2011 to 2015 were obtained for self-identified Black and White enrollees. The analyses were limited to Black and White Medicaid enrollees because they accounted for the largest percentages of the population and could be compared with greater statistical power than other groups. Data were analyzed from June 2021 to September 2024.

EXPOSURES Plan enrollment via self-selection (observational population) vs random assignment (randomized population).

MAIN OUTCOMES AND MEASURES Annual counts of primary care visits, low-acuity emergency department visits, prescription drug fills, and total spending. For observational and randomized populations, models of each outcome were fit as a function of plan indicators, indicators for race, interactions between plan indicators and race, and age and sex. Models estimated the magnitude of racial differences within each plan and tested whether this magnitude varied across plans.

RESULTS Of 118 101 enrollees (mean [SD] age, 9.3 [7.5] years; 53.0% female; 61.4% non-Hispanic Black; and 38.6% non-Hispanic White), 70.2% were included in the randomized population, and 29.8% were included in the observational population. Within-plan differences in primary care visits, low-acuity emergency department visits, prescription drug use, and total spending between Black and White enrollees were large but did not vary substantially and were not statistically significantly different across plans in the randomized population, suggesting minimal effects of plans on racial differences in these measures. In contrast, in the observational population, racial differences varied substantially across plans (standard deviations 2-3 times greater than in the randomized population); this variation was statistically significant after adjustment for multiple testing, except for emergency department visits. Greater between-plan variation in racial differences in the observational population was only partially explained by sampling error. Stratifying by race did not bring observational estimates of plan effects meaningfully closer to randomized estimates.

CONCLUSIONS AND RELEVANCE This cross-sectional study showed that selection bias may mischaracterize plans' relative performance on measures of health care disparities. It is critical to address disparities in Medicaid, but adjusting plan payments based on disparity measures may have unintended consequences.

Editor's Note

Supplemental content

Author Affiliations: Yale School of Public Health, New Haven, Connecticut (Wallace, Ndumele, Lollo); Yale University, New Haven, Connecticut (Agafiev Macambira, Lavallee, Green); University of Michigan, Ann Arbor (Duchowny); Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts (McWilliams).

Corresponding Author: Jacob Wallace, PhD, Yale School of Public Health, 60 College St, New Haven, CT 06510 (jacob.wallace@yale.edu).

JAMA Intern Med. doi:10.1001/jamainternmed.2024.5451
Published online November 25, 2024.

Studies of nearly every sector of the health care system show poorer access and outcomes for historically disadvantaged groups.[1-7] The Medicaid program is the largest insurer of low-income populations in the US,[8] including disproportionately high shares of individuals who identify as belonging to racial and/or ethnic minority groups.[9] In recent years, states and researchers have started to document health disparities in Medicaid[10-14] and consider the available policy options to address them.[15,16]

One policy direction is to build on pay-for-performance or public reporting initiatives to include health disparity measures. As of December 2023, 11 state Medicaid programs had tied financial incentives to health equity, and this coincides with widespread calls for public reporting on quality and utilization measures stratified by race and ethnicity to inform the choices made by members of historically disadvantaged groups.[17-20] Payment reforms intended to promote health equity either (1) direct additional resources to organizations that disproportionately serve historically disadvantaged populations[21] or (2) measure and reward organizations based on measures of health care disparities or stratified measures.

As with standard pay-for-performance arrangements and public reporting programs, accurate performance profiling is vital to initiatives that report on, or tie financial incentives to, health care disparity measures. There is now evidence that traditional pay-for-performance can lead to unmerited financial transfers when there is selection bias (ie, when patient populations vary in ways not accounted for by the systems of risk adjustment).[22-29] These unintended redistributions may exacerbate health care disparities by directing resources away from organizations serving more disadvantaged populations. If such selection bias is severe, initiatives that directly target resources to organizations serving historically disadvantaged populations may be preferable to those that reward organizations based on health disparity measures.

This study leverages a natural experiment in which a state Medicaid program randomly assigned a subset of enrollees to different managed care plans (randomized population), while others chose among the same set of plans (observational population). We quantified within-plan differences in care for Black and White Medicaid enrollees and the extent to which these differences varied across plans—both in the observational population (the scenario typically available for judging plan performance) and in the randomized population (where confounding due to nonrandom sorting of enrollees to plans was removed).

Key Points

Question Are differences across health plans in racial disparities in health care utilization due to differences in plan performance or selection bias?

Findings In this cross-sectional study of 118 101 enrollees in a state Medicaid program, there were large differences in utilization between Black and White enrollees within each plan. Variations in racial differences were statistically significant across plans among enrollees who selected a plan but not among those who were randomly assigned to a plan.

Meaning Selection bias may cause across-plan comparisons of within-plan racial disparities to mischaracterize plan performance, and large within-plan disparities suggest it remains critical to address inequitable care.

Methods

Population

We obtained enrollee-level administrative claims and enrollment data from 2011 to 2015 for a southern US state whose Medicaid program transitioned in 2012 from traditional fee for service to contracting with 5 Medicaid managed care (MMC) plans. Enrollees were randomized to 1 of the 5 plans if they did not select a plan within 30 days of being notified of the transition to MMC (see the eMethods in Supplement 1 for additional details on the transition, autoassignment process, and randomization scheme). Using self-identified race from enrollment data (eTable 1 in Supplement 1), we limited the analyses to Black and White Medicaid enrollees, who accounted for 52.5% and 33.1% of the population, respectively, and could be compared with greater statistical power than other groups.

The institutional review board at Yale University deemed the study exempt from review, and informed consent was waived owing to use of deidentified data. The study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines.

Study Design

We examined across-plan variation in within-plan differences in health care use between Black and White enrollees (racial differences). These comparisons were performed separately in the observational population of 35 187 enrollees who selected a plan and the randomized population of 82 914 enrollees assigned to a plan. In supplementary analyses, we similarly examined differences between Hispanic and non-Hispanic White enrollees.

Results from the randomized population quantify the impact of MMC plans on racial differences in health care utilization. In this population, where random assignment eliminates selection bias, differences between plans could be due to different physician networks, utilization management, or quality improvement efforts.[30-32]

We then assessed whether results from the observational population yielded the same conclusion as the results from the randomized population about the effects of plans on racial differences or an erroneous conclusion due to residual confounding from nonrandom sorting.[29]

Study Variables

Plan Exposure

From administrative data, we determined each enrollee's plan in each study year and whether the plan was randomly assigned or actively chosen. For assigned enrollees, we used the assigned plan in an intention-to-treat framework because using the plan selected by those who did not comply with their assigned plan could introduce selection bias. A high proportion

(86%) of autoassigned enrollees remained in their assigned plans (eFigure 1 in Supplement 1). In sensitivity analyses, we used instrumental variables to rescale intention-to-treat estimates based on plan-specific rates of noncompliance (eMethods in Supplement 1).

Primary and Secondary Outcomes

As primary outcomes, we constructed 4 annual measures of health care utilization for each enrollee: (1) number of visits with a primary care clinician, (2) number of prescriptions filled, (3) number of low-acuity emergency department (ED) visits, and (4) total health care spending (payments to clinicians and hospitals) for all medical services and prescription drugs. We considered total spending a summary utilization measure, as prices varied minimally across plans.[33] Low-acuity ED visits were considered visits to the ED that could have been appropriately managed within 24 hours in a primary care or other ambulatory setting.[34] We focused on these 4 measures of health care use, as they should reflect plan variation both in enrollee characteristics and plan attributes that could affect plan performance on a range of measures derived from use of outpatient care, acute care, or prescription drugs. In supplementary analyses, we examined total ED visits, guideline-concordant primary care use stratified by age, potentially high-value drug use, and quality measures related to care of acute and chronic conditions.

Enrollee Characteristics

We obtained enrollee age, sex, self-identified race (Black or White), and Medicaid eligibility category from administrative data. For each study year, we used concurrent claims to construct 141 condition indicators derived from the Health and Human Services Hierarchical Condition Categories model used in the Affordable Care Act marketplaces. As a summary measure of enrollee risk, we report predicted spending as a function of all observed characteristics (eMethods in Supplement 1).

Statistical Analysis

First, we examined the balance of baseline enrollee characteristics across plans in the observational and randomized populations, overall and by race. For each baseline characteristic we performed an F-test of the joint significance of plan differences, adjusting for multiple testing using the Benjamini-Hochberg method.[35]

Second, within each population, we used linear regression to estimate plan effects on racial differences in utilization. In the observational population (ie, those who made active plan choices), we used the plan of enrollment in the model. In the randomized population, we used the assigned plan. Specifically, we fit a model of each outcome as a function of plan indicators, interactions between plan indicators and race, and enrollee covariates. The interaction terms yielded the estimates of interest—the difference between Black and White enrollees in utilization within each plan. We performed an F-test of the joint significance of these interaction terms to test whether plan variation in racial differences was statistically significant, adjusting for multiple hypothesis testing using the Benjamini-Hochberg correction (eMethods in Supplement 1).[35] Models included indicators for enrollees' established physician organization, as randomization was conditional on that factor (eMethods in Supplement 1).

Following accepted frameworks for measuring health care disparities,[18,19] we included only enrollee age (using indicator variables for each 5-year age bin) and sex as covariates in the primary analysis, as we were interested in racial differences in care mediated by any factor affected by structural disadvantage. In addition, adjustment for risk scores derived from diagnoses in claims (recorded only when individuals used care) could introduce bias because of racial differences in access to care or plan variation in diagnosis coding intensity.

Even without adjusting for additional characteristics, one might expect variation across plans in racial differences in care to be similar in the observational and randomized populations if nonrandom sorting in the observational population followed a similar pattern for Black and White enrollees. A pay-for-equity approach might make this assumption, negating the need for risk adjustment when assessing and comparing disparities that, conceptually, should not be risk adjusted.[18,19] However, this assumption may not hold for 2 reasons. First, Black and White enrollees may differ in their nonrandom sorting to plans. Second, the pattern of nonrandom sorting may be similar—meaning that both Black and White enrollees with a given characteristic (eg, very low incomes or high illness burden) may disproportionately choose the same plan—but differences in Black and White enrollees' care use may vary by such enrollee characteristics. Either of these mechanisms would contribute to plan variation in racial disparities in the observational population. We explored these mechanisms through additional adjustments for enrollee characteristics—including geography, eligibility categories, and Health and Human Services Hierarchical Condition Categories indicators—and for interactions between race and enrollee characteristics (eMethods in Supplement 1).

All statistical analyses were performed using Stata, version 14 (StataCorp). The level of significance was $P < .05$, and tests were 2-sided. Data were analyzed from June 2021 to September 2024.

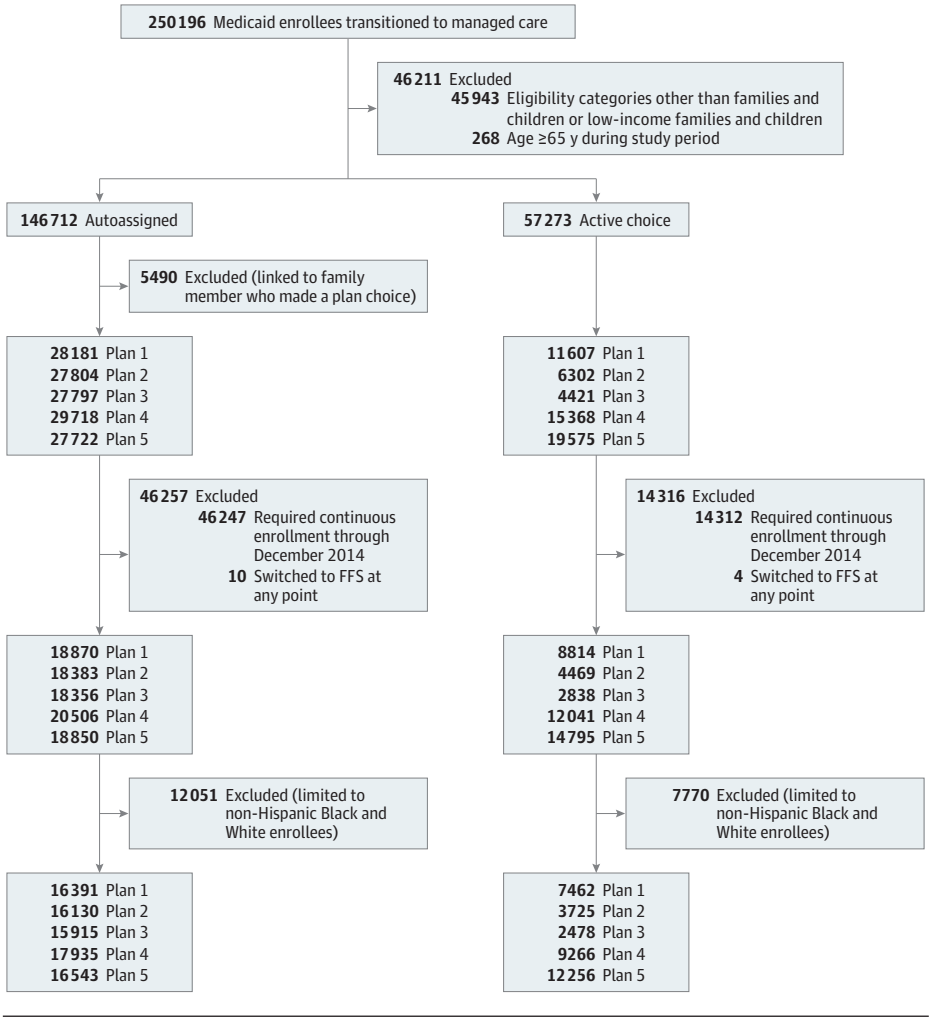
Secondary Analyses

In secondary analyses, we assessed whether stratification by race improved the accuracy of plan effects estimated in the observational population for Black enrollees (eMethods in Supplement 1). Specifically, we determined whether estimates of plan effects for Black enrollees in the randomized population (the criterion standard) were better approximated by estimates based on Black enrollees in the observational population than by estimates based on the full, pooled population of Black and White enrollees in the observational population. In this comparison, we normalized differences in mean utilization between the Black population and full population by expressing estimated plan effects in relative terms, as a percentage of the population-specific mean.

Sensitivity Analyses

We performed several sensitivity analyses. First, we assessed whether differences in plan variation between the random-

Figure 1. Sample Restrictions



This figure indicates all of the sample exclusions enforced on the 250 196 Medicaid enrollees who were transitioned to Medicaid managed care on February 1, 2012, to arrive at the final study population for both the randomized (autoassigned) and observational (active choice) populations. FFS indicates fee for service.

ized and observational populations were due to smaller sample sizes in the observational population. Second, we assessed whether spending results were sensitive to alternate transformations of the dependent variable (eFigure 2 in Supplement 1).[36] Third, we assessed whether results were sensitive to weighting the randomized population to match the characteristics of the observational population to limit confounding from heterogeneity in estimated plan effects across the 2 (different) populations.

Results

Study Population

After exclusions, the primary study population comprised 118 101 enrollees (**Figure 1**). The mean (SD) age was 9.3 (7.5) years, 6.0% were adults, 53.0% were female, 61.4% self-identified as non-Hispanic Black, and 38.6% self-identified as non-Hispanic White. There were 82 914 enrollees (70.2%) in the randomized population and 35 187 (29.8%) in the observational population. Enrollees in the randomized population were more likely to identify as Black, were moderately older, and had lower levels of predicted spending (**Table 1**).

Differences in Enrollee Characteristics Between Plans

In the observational population, enrollees' age, sex, race, health conditions (eg, pregnancy, cardiovascular conditions), and predicted spending varied substantially across the 5 plans, consistent with nonrandom sorting. By comparison, enrollee characteristics varied minimally across the 5 assigned plans in the randomized population (Table 1). Enrollee characteristics were similarly balanced in the randomized population and imbalanced in the observational population when examined separately for Black and White enrollees (eTables 2 and 3 in Supplement 1). The patterns of nonrandom sorting to plans were not statistically significantly different for Black and White enrollees based on observable characteristics (eTable 4 in Supplement 1).

Differences Between Plans in Racial Differences in Care

In the observational population, within-plan racial differences in utilization were large (**Figure 2**). Relative to White enrollees,

Table 1. Differences Between Plans in Enrollee Characteristics in the Observational and Randomized Populations

Characteristic	Population mean, %	No. (%)					SD of plan means, %	*P* value from F-test[a]	Benjamini-Hochberg–corrected *P* values
		Plan 1	Plan 2	Plan 3	Plan 4	Plan 5			
Observational population									
Total No.	NA	7462	3725	2478	9266	12 256	NA	NA	NA
Age, y									
≤5	34.2	2626 (36.4)	1203 (31.6)	735 (33.7)	3224 (34.5)	4237 (33.5)	1.7	<.001	<.001
6-17	60.7	4424 (59.3)	2267 (62.5)	1285 (55.3)	5828 (61.1)	7568 (61.9)	2.9	<.001	<.001
18-64	5.1	412 (4.3)	255 (6.0)	458 (11.0)	214 (4.4)	451 (4.6)	2.9	<.001	<.001
Sex									
Female	52.7	3999 (53.2)	2025 (53.9)	1457 (55.4)	4657 (51.4)	6401 (52.5)	1.5	.007	.01
Male	47.3	3463 (46.8)	1700 (46.1)	1021 (44.6)	4609 (48.6)	5855 (47.5)	1.5	.007	.01
Race									
Non-Hispanic Black	53.0	4520 (57.0)	2095 (56.0)	1558 (59.5)	4156 (49.5)	6319 (51.0)	4.2	<.001	<.001
Non-Hispanic White	47.0	2942 (43.0)	1630 (44.0)	920 (40.5)	5110 (50.5)	5937 (49.0)	4.2	<.001	<.001
Health condition[b]									
Asthma	7.8	561 (7.7)	294 (7.9)	193 (7.3)	741 (7.8)	950 (8.0)	0.2	.88	.88
Serious mental illness	3.4	265 (3.7)	99 (3.2)	70 (3.4)	444 (3.5)	334 (3.3)	0.2	.43	.53
Diabetes	0.9	73 (0.9)	29 (0.7)	45 (1.2)	57 (0.8)	103 (0.9)	0.2	.56	.62
Pregnancy	1.1	91 (1.0)	66 (1.6)	85 (2.0)	47 (0.9)	107 (1.1)	0.5	.007	.01
Cardiovascular conditions	1.4	93 (1.2)	64 (1.8)	23 (1.0)	122 (1.3)	186 (1.5)	0.3	.01	.02
Predicted annual spending based on enrollee characteristics, $[c]	1914	1874	1936	1897	1948	1909	29.8	.01	.02
Randomized population									
Total No.	NA	16 391	16 130	15 915	17 935	16 543	NA	NA	NA
Age, y									
≤5	29.7	4836 (29.5)	4658 (30.7)	4463 (29.5)	5693 (29.1)	4977 (29.7)	0.6	.76	.83
6-17	64.0	10 523 (64.2)	10 161 (62.9)	10 220 (64.1)	11 567 (64.8)	10 559 (63.8)	0.7	.63	.83
18-64	6.3	1032 (6.3)	1311 (6.5)	1232 (6.4)	675 (6.1)	1007 (6.5)	0.2	.71	.83
Sex									
Female	53.1	8761 (53.2)	8631 (52.7)	8534 (53.2)	9257 (52.9)	8862 (53.7)	0.4	.36	.83
Male	46.9	7630 (46.8)	7499 (47.3)	7381 (46.8)	8678 (47.1)	7681 (46.3)	.04	.36	.83
Race									
Non-Hispanic Black	64.9	11 173 (64.7)	10 943 (66.4)	10 077 (65.6)	10 595 (63.4)	11 039 (64.7)	1.2	.007	.08
Non-Hispanic White	35.1	5218 (35.3)	5187 (33.6)	5838 (34.4)	7340 (36.6)	5504 (35.3)	1.2	.007	.08

(continued)

Table 1. Differences Between Plans in Enrollee Characteristics in the Observational and Randomized Populations (continued)

Characteristic	Population mean, %	No. (%) Plan 1	Plan 2	Plan 3	Plan 4	Plan 5	SD of plan means, %	P value from F-test[a]	Benjamini-Hochberg-corrected P values
Health condition[b]									
Asthma	6.3	1041 (6.1)	1094 (6.4)	983 (6.3)	1110 (6.5)	990 (6.1)	0.2	.60	.83
Serious mental illness	2.8	435 (2.7)	486 (3.0)	462 (2.8)	562 (2.7)	389 (2.8)	0.1	.60	.83
Diabetes	0.6	117 (0.7)	100 (0.7)	129 (0.6)	74 (0.6)	102 (0.7)	0.1	.49	.83
Pregnancy	1.3	232 (1.5)	263 (1.2)	223 (1.3)	157 (1.2)	212 (1.4)	0.1	.21	.83
Cardiovascular conditions	1.3	196 (1.2)	215 (1.4)	217 (1.3)	222 (1.2)	200 (1.2)	0.1	.27	.83
Predicted annual spending based on enrollee characteristics, $[c]	1511	1514	1509	1515	1506	1511	3.7	.97	.97

Abbreviation: NA, not applicable.

[a] This column reports P values from F-tests of joint significance of the difference in plan means in the regression model.

[b] Health conditions were assigned based on the Health and Human Services Hierarchical Condition Categories model using all diagnoses in 2011, the calendar year prior to assignment. The conditions correspond to the following Hierarchical Condition Categories: asthma (160, 161.1, and 161.2), serious mental illness (87.1, 87.2, 88, 90, 102, and 103), diabetes (20 and 21), pregnancy (203, 204, 205, 207, 208, 209, 210, 211, and 212), and cardiovascular conditions (125, 126, 127, 128, 129, 130, 131, 132, 135, 137, 138, 139, and 142).

[c] Predicted annual spending is based on a spending outcome that is winsorized at $50 000 annually, or the 99.93 percentile of annual spending in this population. Predicted annual spending uses a cross-validated LASSO (least absolute shrinkage and selection operator) regression evaluated at the enrollee level, where the outcome is average annual spending after assignment and the predictors are an enrollee's baseline characteristics (5-year age bins, sex, prior physician, county of residence, eligibility category, and indicators for all diagnosis codes listed and prescription drugs used in the year prior to assignment).

Black enrollees had lower primary care use, prescription drug use, and spending, but higher use of the ED for low-acuity visits. The magnitude of these differences varied substantially and statistically significantly across plans (Figure 2 and **Table 2**). For primary care visits, for example, the White enrollee mean (SD) was 552.1 (510.5) visits per 100 enrollees per year, and the difference between Black and White enrollees within plans ranged from 178 fewer visits per 100 enrollees per year for Black enrollees in plan 4 to 112 fewer visits per 100 enrollees per year for Black enrollees in plan 3 (range, 66 visits; $P = .01$ after adjustment for multiple testing). Racial differences in prescription drug utilization and total spending also differed significantly between plans in the observational population. Plan variation in racial differences in low-acuity ED use did not reach statistical significance but was greater as a proportion of the population mean than variation in the other measures, and plan variation in racial differences in total ED visits was statistically significant (eTable 5 in Supplement 1).

In the randomized population, within-plan racial differences in care use between Black and White enrollees also were large, but variation in racial differences was not statistically significant across plans (Table 2 and Figure 2). Sampling error only partially explained the greater plan variation in racial differences in the smaller observational population, though the contribution of sampling error differed across outcomes (eTable 6 in Supplement 1).

Secondary and Sensitivity Analyses

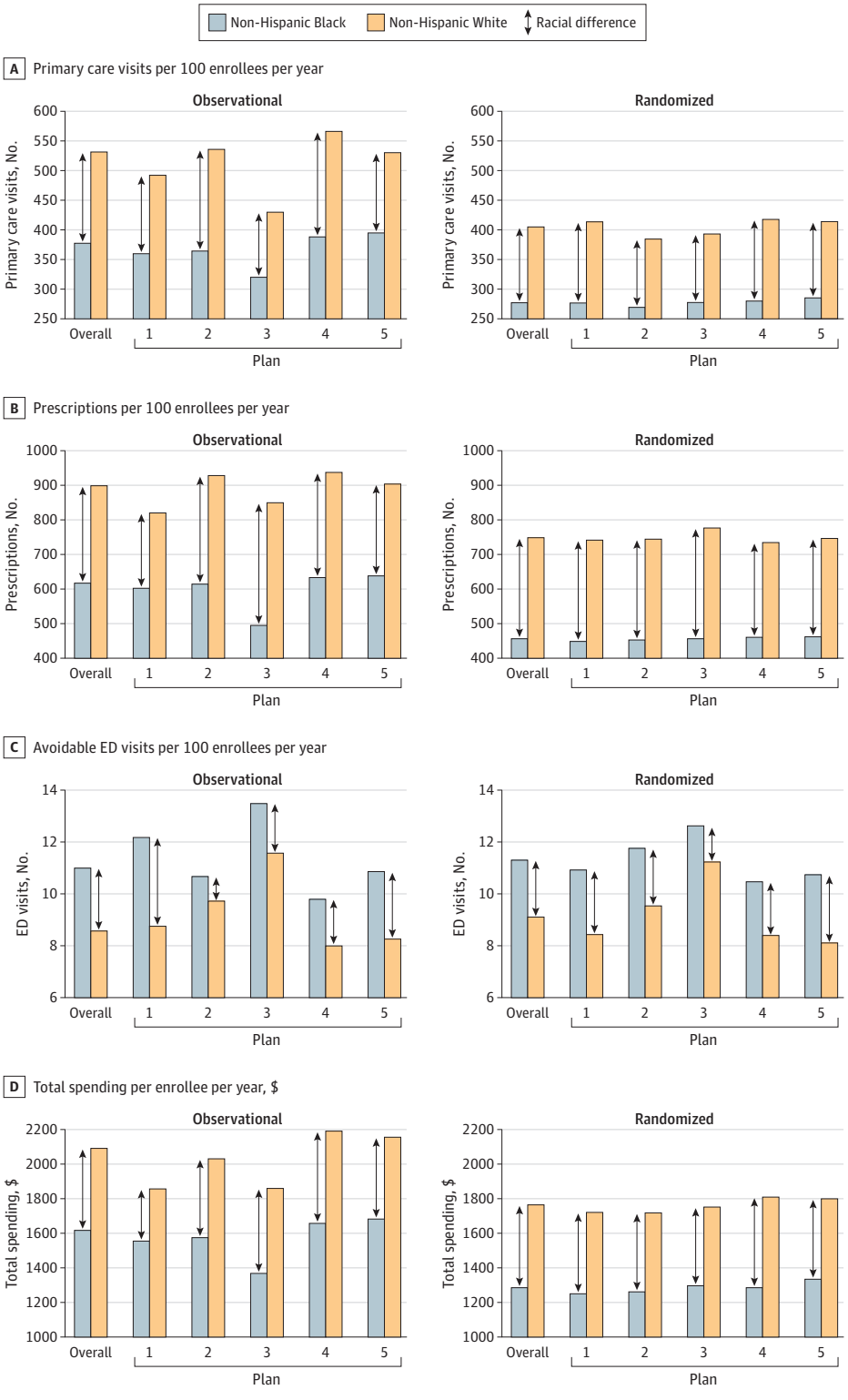
Stratified analyses revealed that estimates of plan performance in both the Black and White observational populations were similarly biased by nonrandom sorting (eFigures 3-7 and eTables 7-10 in Supplement 1). Moreover, stratifying by race did not consistently or meaningfully improve the accuracy of observational estimates (**Figure 3**). Compared with observational estimates of plan performance based on the pooled population of Black and White enrollees, the stratified observational estimates for Black enrollees were not, on average, closer to estimates based on the randomized population of Black enrollees; this was the case for 12 of the 20 plan-outcome estimates (Figure 3).

Sensitivity analyses supported the conclusions of the main analyses (eTables 11-16 in Supplement 1). In the observational population, adjustment for geography, eligibility category, and condition indicators reduced plan variation somewhat (eTables 17 and 18 in Supplement 1), as did interactions between race and enrollee characteristics (eTable 19 in Supplement 1). Results comparing Hispanic and non-Hispanic White enrollees were qualitatively similar to the main findings (eFigure 8 and eTables 20 and 21 in Supplement 1).

Discussion

In this study of 1 state's MMC program, we observed large racial differences in utilization between Black and White enrollees within each plan. There was statistically significant variation in racial differences in care use across plans in a population of enrollees who chose plans (the observational population),

Figure 2. Mean Health Care Utilization and Spending by Plan and Race for Enrollees in the Observational and Randomized Populations



Averages for each plan are calculated among all non-Hispanic Black and non-Hispanic White enrollees who made an active choice and are adjusted for 5-year age bins, sex, and enrollee's prior physician (the unit of randomization). Although the observational population was not randomized, data were adjusted for prior physician in both populations for consistency. ED indicates emergency department.

but plan variation was limited and not statistically significant in a population of enrollees randomized to the same plans. Although based on 5 plans in a single state, the findings illustrate how assessments of health care disparities at a plan level may be subject to selection bias when based on observational data typically used for public reporting and pay-for-

Table 2. Variation in Racial Differences Across Plans in Health Care Utilization and Spending[a]

Outcome measure	Unadjusted non-Hispanic White enrollees, mean No.	Plan specific gap between non-Hispanic Black and non-Hispanic White enrollees (95% CI)[b]					Difference between plans in disparities		
		Plan 1	Plan 2	Plan 3	Plan 4	Plan 5	SD (range)	*P* value from F-test for difference in plan[c]	Benjamini-Hochberg-corrected *P* values
Observational population									
Primary care visits per 100 enrollees/y	552.1	−131.3 (−156.3 to −106.4)	−171.5 (−211.9 to −131.1)	−111.9 (−151.9 to −72.0)	−178.1 (−201.3 to −154.9)	−135.5 (−152.8 to −118.1)	28.1 (66.2)	.005	.01
Prescription drugs per 100 enrollees/y	942.6	−218.3 (−275.7 to −160.8)	−314.3 (−431.8 to −196.8)	−353.1 (−504.7 to −201.5)	−304.8 (−354.3 to −255.2)	−266.5 (−314.9 to −218.0)	51.2 (134.8)	.003	.01
Low-acuity ED visits per 100 enrollees/y	8.5	3.5 (1.8 to 5.1)	1.0 (−1.6 to 3.5)	1.9 (−1.4 to 5.1)	1.8 (0.5 to 3.1)	2.6 (1.0 to 4.2)	1.0 (2.5)	.05	.05
Total spending per enrollee/y, $[d]	2134	−302 (−426 to −178)	−458 (−673 to −242)	−495 (−785 to −205)	−533 (−665 to −402)	−472 (−559 to −384)	88.7 (232)	.02	.02
Randomized population									
Primary care visits per 100 enrollees/y	429.8	−136.9 (−158.6 to −115.2)	−115.8 (−135.6 to −95.9)	−116.8 (−135.3 to −98.3)	−137.8 (−167.2 to −108.4)	−127.7 (−141.2 to −114.1)	10.6 (22.0)	.14	.56
Prescription drugs per 100 enrollees/y	794.3	−293.3 (−334.6 to −252.0)	−290.6 (−332.3 to −248.8)	−319.5 (−366.7 to −272.4)	−273.4 (−307.8 to −239.0)	−284.1 (−327.7 to −240.5)	17.1 (46.1)	.55	.66
Low-acuity ED visits per 100 enrollees/y	9.0	2.5 (1.3 to 3.7)	2.2 (1.1 to 3.3)	1.4 (0.0 to 2.8)	2.1 (1.3 to 2.8)	2.7 (1.4 to 3.9)	0.5 (1.3)	.49	.66
Total spending per enrollee/y, $[d]	1798	−473 (−554 to −392)	−452 (−548 to −357)	−455 (−570 to −340)	−525 (−607 to −442)	−457 (−592 to −322)	30.3 (72)	.66	.66

Abbreviation: ED, emergency department.

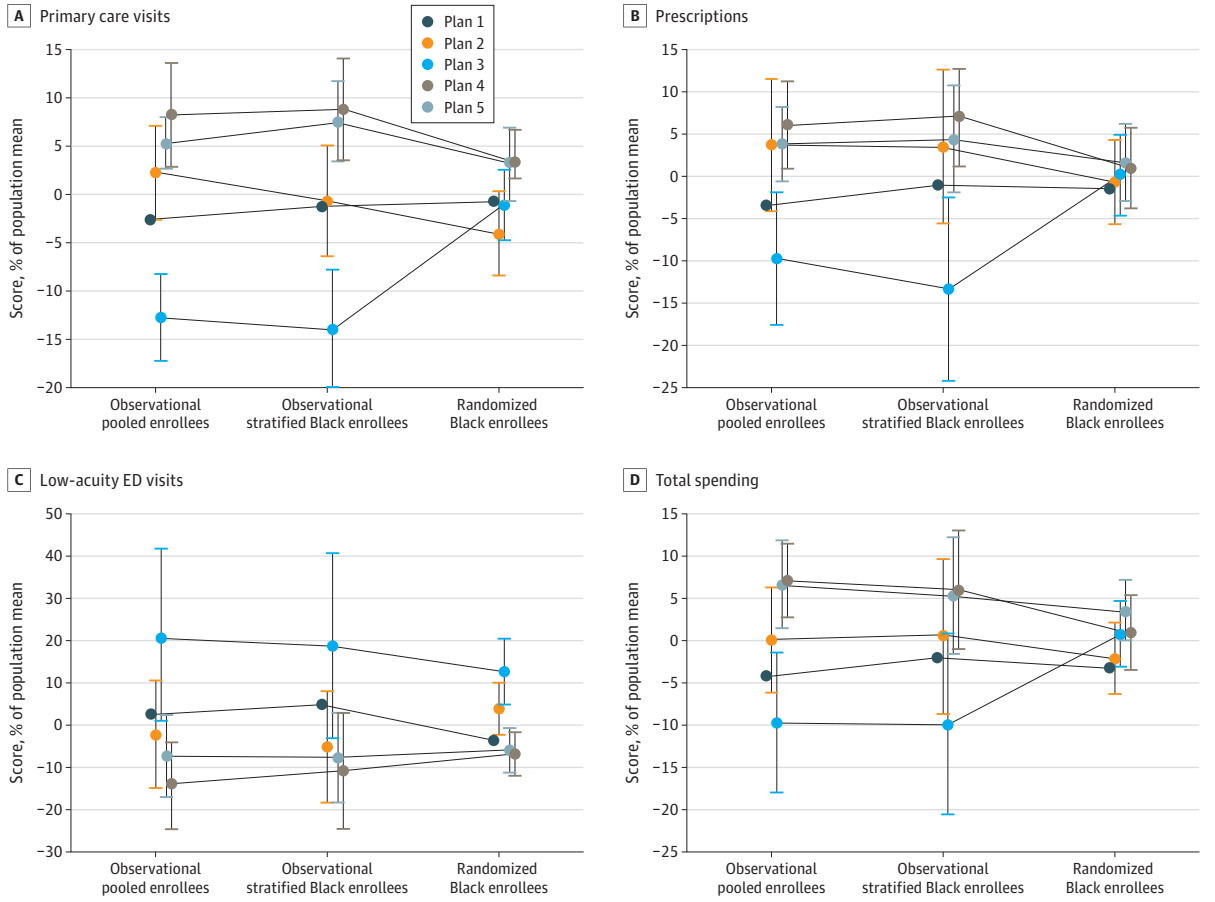
[a] For both populations, plan-specific differences were derived from regressions that control for an enrollee's age (using 5-year age bins), sex, and the unit of randomization (see the Methods section and eMethods in Supplement 1 for additional details).

[b] Gap is calculated as plan effects estimated for non-Hispanic Black enrollees minus plan effects estimated for non-Hispanic White enrollees.

[c] This column reports *P* values from F-tests of the joint significance of the mean plan-specific differences in the regression model.

[d] Total annual spending is winsorized at $50 000 annually, or the 99.93 percentile of annual spending within the population.

Figure 3. Assessing Bias in Stratified Reporting Among Non-Hispanic Black Enrollees



For all populations, plan-specific utilization and spending measures are derived from regressions that control for an enrollee's age (using 5-year age bins), sex, and enrollees' established physician organization, within which randomization to plans was conducted. A plan's score is equal to that plan's difference from average plan performance within the population. The observational pooled sample includes plan scores estimated by pooling data for Black and White enrollees who made active plan choices. The observational stratified sample includes plan scores estimated using only data for Black enrollees who made active plan choices. The randomized sample includes plan scores estimated using only data for Black enrollees who were randomly assigned to health plans. ED indicates emergency department. Error bars represent 95% CIs.

performance arrangements. Specifically, the present findings are consistent with Black and White enrollees differing in their nonrandom sorting to plans based on unobservable characteristics (sorting to plans on observable characteristics did not differ much between Black and White enrollees) and with differences in care use between Black and White enrollees varying across enrollee characteristics.

The substantively different conclusion that would be reached by analysis of the observational vs randomized population should caution payers, policymakers, and researchers against drawing inferences about plan effects on health equity from available data, as plan variation in racial differences in care may reflect selection bias rather than true differences in plan performance that impact health disparities.[15] These results imply that attaching financial incentives to such observational performance profiles may result in payment reallocations unrelated to plans' influence on health care disparities, potentially compromising the effectiveness of initiatives crafted around such comparisons.

In stratified analyses, we found evidence of selection bias in observational estimates of plan effects on care for both Black and White enrollees. Although stratified estimates differed somewhat from the overall population estimates of plan performance, they were not consistently and meaningfully closer to the true effects of plans on utilization, as estimated in the randomized population. These findings suggest that stratified reporting of measured plan performance by race may not better inform plan selections by Black enrollees.

In both the randomized and observational populations, within-plan racial differences in care use were large. While racial differences do not always imply a disparity (eg, if due to overuse among White enrollees),[18] there are reasons to believe a disparity may exist here. First, it is well documented that structural and interpersonal racism decrease access for racial and ethnic minority groups,[37-46] and these barriers continue to exist within the Medicaid setting.[14] Second, historical concerns about access to primary care in Medicaid suggest that the lower utilization of primary care among Black enroll-

ees (and their increased use of the ED for avoidable reasons) reflects underservice due to long-standing barriers rather than overuse by White enrollees.

The large size of the racial differences in utilization, and minimal to modest variation in those differences across plans in the randomized population, suggest that MMC plans may have limited means or incentive to meaningfully address disparities without addressing the root cause of structural racism.[47] Notably, the impact of plans on racial differences was, at most, modest despite substantial differences in physician networks, utilization management strategies, and quality-improvement activities.[33] Given the potential drawbacks of pay-for-equity approaches underscored by the present findings, alternative approaches to advance health equity are likely needed. For example, capitated payments to plans may be increased for socially disadvantaged groups, independent of plan performance, to strengthen incentives for plans to attract those groups with enhanced benefits, access, and care.[48] However, structural disadvantages, such as clinician shortages in historically marginalized communities and other barriers (ie, limited transportation), are likely to persist without more substantial increases in Medicaid payment rates or other investments.

Limitations

This study has several limitations. First, the study is based on a single Medicaid program with a young population (only 6.0% were adults, accounting for 12.7% of spending) and a limited set of utilization measures. Hence, it is unclear whether the selection bias we observed would generalize to an adult population. Second, since the data come from a southern US state,[49,50] we acknowledge the unique historical context that may have shaped the results observed in this study[51,52] and limits their generalizability to other states in the US. Third, different patterns of plan variation in disparities in the randomized and observational populations may reflect heterogeneity in estimated plan effects in the 2 populations rather than revealing that differences between plans in the observational population are associated with selection. However, results were similar after reweighting the populations to balance their characteristics.

Conclusions

In this cross-sectional study, we found substantial variation across MMC plans in racial differences in health care that was largely an artifact of selection bias. These findings highlight the difficulty of measuring a plan's impact on health equity while underscoring that health care disparities remain large in Medicaid. Although these findings suggest that initiatives to adjust payments to MMC plans based on health equity measures may have unintended consequences, they also demonstrate a pressing need for well-designed policies to measure and address inequities in Medicaid.

ARTICLE INFORMATION

Accepted for Publication: August 20, 2024.

Published Online: November 25, 2024. doi:10.1001/jamainternmed.2024.5451

Author Contributions: Prof Wallace and Dr Lollo had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Wallace, Ndumele, McWilliams.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: Wallace, Lollo, McWilliams.
Critical review of the manuscript for important intellectual content: All authors.
Statistical analysis: Lollo, Agafiev Macambira, Lavallee, Green, McWilliams.
Obtained funding: Wallace.
Administrative, technical, or material support: Ndumele, Duchowny.
Supervision: Wallace.

Conflict of Interest Disclosures: Prof Wallace reported that his spouse is a counsel at Manatt, Phelps & Phillips, a Medicaid-focused consultancy. Dr McWilliams reported personal fees from the Center for Medicare and Medicaid Innovation, RTI, Analysis Group, the MITRE Corporation, America's Physician Groups, Phillips and Cohen, and Action Now Initiative; grants from the National Institute on Aging and the Commonwealth Fund; and serving as a member of the board of directors for the Institute for Accountable Care. No other disclosures were reported.

Funding/Support: This study was supported by the Commonwealth Fund. Dr McWilliams was supported by a grant from Arnold Ventures, and Dr Duchowny was supported by a grant from the National Institute on Aging (K99AG066846).

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The content of this article is solely the responsibility of the authors and does not reflect the views of any organization with which they are affiliated.

Data Sharing Statement: See Supplement 2.

REFERENCES

1. Cook NL, Ayanian JZ, Orav EJ, Hicks LS. Differences in specialist consultations for cardiovascular disease by race, ethnicity, gender, insurance status, and site of primary care. *Circulation*. 2009;119(18):2463-2470. doi:10.1161/CIRCULATIONAHA.108.825133

2. Cook NL, Orav EJ, Liang CL, Guadagnoli E, Hicks LS. Racial and gender disparities in implantable cardioverter-defibrillator placement: are they due to overuse or underuse? *Med Care Res Rev*. 2011;68(2):226-246. doi:10.1177/1077558710379421

3. Epstein AM, Ayanian JZ, Keogh JH, et al. Racial disparities in access to renal transplantation—clinically appropriate or due to underuse or overuse? *N Engl J Med*. 2000;343(21):1537-1544. doi:10.1056/NEJM200011233432106

4. Joshi S, Gaynor JJ, Bayers S, et al. Disparities among Blacks, Hispanics, and Whites in time from starting dialysis to kidney transplant waitlisting. *Transplantation*. 2013;95(2):309-318. doi:10.1097/TP.0b013e31827191d4

5. Ferdows NB, Aranda MP, Baldwin JA, Baghban Ferdows S, Ahluwalia JS, Kumar A. Assessment of racial disparities in mortality rates among older adults living in US rural vs urban counties from 1968 to 2016. *JAMA Netw Open*. 2020;3(8):e2012241. doi:10.1001/jamanetworkopen.2020.12241

6. Ayanian JZ, Landon BE, Newhouse JP, Zaslavsky AM. Racial and ethnic disparities among enrollees in Medicare Advantage plans. *N Engl J Med*. 2014;371(24):2288-2297. doi:10.1056/NEJMsa1407273

7. Caraballo C, Massey DS, Ndumele CD, et al. Excess mortality and years of potential life lost among the Black population in the US, 1999-2020. *JAMA*. 2023;329(19):1662-1670. doi:10.1001/jama.2023.7022

8. Donohue JM, Cole ES, James CV, Jarlenski M, Michener JD, Roberts ET. The US Medicaid program: coverage, financing, reforms, and implications for health equity. *JAMA*. 2022;328(11):1085-1099. doi:10.1001/jama.2022.14791

9. Distribution of the nonelderly with Medicaid by race/ethnicity. Kaiser Family Foundation. Accessed September 16, 2024. https://www.kff.org/medicaid/state-indicator/medicaid-distribution-nonelderly-by-raceethnicity/

10. McCue MJ, Bailit MH. Assessing the financial health of Medicaid managed care plans and the quality of patient care they provide. The Commonwealth Fund. June 15, 2011. Accessed September 16, 2024. https://www.commonwealthfund.org/publications/issue-briefs/2011/jun/assessing-financial-health-medicaid-managed-care-plans-and

11. Gordon SH, Gadbois EA, Shield RR, Vivier PM, Ndumele CD, Trivedi AN. Qualitative perspectives of primary care providers who treat Medicaid managed care patients. *BMC Health Serv Res*. 2018; 18(1):728. doi:10.1186/s12913-018-3516-9

12. Kuziemko I, Meckel K, Rossin-Slater M. Does managed care widen infant health disparities? evidence from Texas Medicaid. *Am Econ J Econ Policy*. 2018;10(3):255-283. doi:10.1257/pol.20150262

13. Cook BL. Effect of Medicaid Managed Care on racial disparities in health care access. *Health Serv Res*. 2007;42(1 pt 1):124-145. doi:10.1111/j.1475-6773.2006.00611.x

14. Wallace J, Lollo A, Duchowny KA, Lavallee M, Ndumele CD. Disparities in health care spending and utilization among Black and White Medicaid enrollees. *JAMA Health Forum*. 2022;3(6):e221398. doi:10.1001/jamahealthforum.2022.1398

15. State Health and Value Strategies. Medicaid managed care contract language: health disparities and health equity. Princeton University. August 2021. Accessed November 19, 2021. https://www.shvs.org/wp-content/uploads/2021/02/SHVS-MCO-Contract-Language-Health-Equity-and-Disparities_August-2021.pdf

16. Machledt D. Addressing health equity in Medicaid managed care. National Health Law Program. May 2021. Accessed November 19, 2021. https://healthlaw.org/wp-content/uploads/2021/05/NHeLP_AddressingHealthEquityMedicaidManagedCare06032021.pdf

17. Hinton E, Williams E, Raphael J, et al. Amid unwinding of pandemic-era policies, Medicaid programs continue to focus on delivery systems, benefits, and reimbursement rates: results from an annual Medicaid budget survey for state fiscal years 2023 and 2024. Kaiser Family Foundation. November 14, 2023. Accessed September 21, 2024. https://www.kff.org/report-section/50-state-medicaid-budget-survey-fy-2023-2024-delivery-systems

18. Smedley BD, Stith AY, Nelson AR. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. National Academy Press; 2003.

19. *Accounting for Social Risk Factors in Medicare Payment*. National Academies Press; 2017.

20. Innovation Center strategy refresh. Center for Medicare and Medicaid Innovation. Accessed November 20, 2022. https://www.cms.gov/priorities/innovation/strategic-direction-whitepaper

21. Fact sheet: health equity enhanced payment initiative. North Carolina Medicaid. March 2021. Accessed July 24, 2024. https://medicaid.ncdhhs.gov/ncmt-provider-factsheet-health-equity-payments-20210331/download

22. Joynt KE, Jha AK. A path forward on Medicare readmissions. *N Engl J Med*. 2013;368(13):1175-1177. doi:10.1056/NEJMp1300122

23. Barnett ML, Hsu J, McWilliams JM. Patient characteristics and differences in hospital readmission rates. *JAMA Intern Med*. 2015;175(11): 1803-1812. doi:10.1001/jamainternmed.2015.4660

24. DuGoff E, Bishop S, Rawal P. Hospital readmission reduction program reignites debate over risk adjusting quality measures. HealthAffairs. August 14, 2014. Accessed September 23, 2024. https://www.healthaffairs.org/content/forefront/hospital-readmission-reduction-program-reignites-debate-over-risk-adjusting-quality

25. Gu Q, Koenig L, Faerberg J, Steinberg CR, Vaz C, Wheatley MP. The Medicare Hospital Readmissions Reduction Program: potential unintended consequences for hospitals serving vulnerable populations. *Health Serv Res*. 2014;49(3):818-837. doi:10.1111/1475-6773.12150

26. Ryan AM. Will value-based purchasing increase disparities in care? *N Engl J Med*. 2013;369(26): 2472-2474. doi:10.1056/NEJMp1312654

27. Lipstein SH, Dunagan WC. The risks of not adjusting performance measures for sociodemographic factors. *Ann Intern Med*. 2014; 161(8):594-596. doi:10.7326/M14-1601

28. Chen LM, Epstein AM, Orav EJ, Filice CE, Samson LW, Joynt Maddox KE. Association of practice-level social and medical risk with performance in the Medicare Physician Value-Based Payment Modifier Program. *JAMA*. 2017;318(5):453-461. doi:10.1001/jama.2017.9643

29. Wallace J, McWilliams JM, Lollo A, Eaton J, Ndumele CD. Residual confounding in health plan performance assessments: evidence from randomization in Medicaid. *Ann Intern Med*. 2022; 175(3):314-324. doi:10.7326/M21-0881

30. Wallace J. What does a provider network do? evidence from random assignment in Medicaid managed care. *Am Econ J Econ Policy*. 2023;15(1): 473-509. doi:10.1257/pol.20210162

31. Geruso M, Layton TJ, Wallace J. What difference does a health plan make? evidence from random plan assignment in Medicaid. *Am Econ J Appl Econ*. 2023;15(3):341-379. doi:10.1257/app.20210843

32. Swaminathan S, Ndumele CD, Gordon SH, Lee Y, Trivedi AN. Association of Medicaid-focused or commercial Medicaid managed care plan type with outpatient and acute care. *JAMA Intern Med*. 2020;180(12):1672-1679. doi:10.1001/jamainternmed.2020.5408

33. Macambira DA, Geruso M, Lollo A, Ndumele CD, Wallace J. *The Private Provision of Public Services: Evidence from Random Assignment in Medicaid*. National Bureau of Economic Research; 2022. doi:10.3386/w30390

34. Statewide collaborative quality improvement project: reducing avoidable emergency room visits. California Department of Health Care Services. June 2012. Accessed September 17, 2024. https://www.dhcs.ca.gov/dataandstats/reports/Documents/MMCD_Qual_Rpts/EQRO_QIPs/CA2011-12_QIP_Coll_ER_Remeasure_Report.pdf

35. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B*. 1995;57(1):289-300. doi:10.1111/j.2517-6161.1995.tb02031.x

36. Buntin MB, Zaslavsky AM. Too much ado about two-part models and transformation? comparing methods of modeling Medicare expenditures. *J Health Econ*. 2004;23(3):525-542. doi:10.1016/j.jhealeco.2003.10.005

37. Probst JC, Laditka SB, Wang JY, Johnson AO. Effects of residence and race on burden of travel for care: cross sectional analysis of the 2001 US National Household Travel Survey. *BMC Health Serv Res*. 2007;7(1):40. doi:10.1186/1472-6963-7-40

38. 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

39. Bulatao RA, Anderson NB, eds. *Understanding Racial and Ethnic Differences in Health in Late Life: a Research Agenda*. National Academies Press; 2004.

40. Boulware LE, Cooper LA, Ratner LE, LaVeist TA, Powe NR. Race and trust in the health care system. *Public Health Rep*. 2003;118(4):358-365. doi:10.1016/S0033-3549(04)50262-5

41. Johnson RL, Roter D, Powe NR, Cooper LA. Patient race/ethnicity and quality of patient-physician communication during medical visits. *Am J Public Health*. 2004;94(12):2084-2090. doi:10.2105/AJPH.94.12.2084

42. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447-453. doi:10.1126/science.aax2342

43. Cooper LA, Roter DL, Johnson RL, Ford DE, Steinwachs DM, Powe NR. Patient-centered communication, ratings of care, and concordance of patient and physician race. *Ann Intern Med*. 2003;139(11):907-915. doi:10.7326/0003-4819-139-11-200312020-00009

44. Gaskin DJ, Spencer CS, Richard P, Anderson GF, Powe NR, Laveist TA. Do hospitals provide lower-quality care to minorities than to whites? *Health Aff (Millwood)*. 2008;27(2):518-527. doi:10.1377/hlthaff.27.2.518

45. Gaskin DJ, Price A, Brandon DT, Laveist TA. Segregation and disparities in health services use. *Med Care Res Rev*. 2009;66(5):578-589. doi:10.1177/1077558709336445

46. Hua CL, Bardo AR, Brown JS. Mistrust in physicians does not explain black-white disparities in primary care and emergency department utilization: the importance of socialization during the Jim Crow era. *J Natl Med Assoc*. 2018;110(6): 540-546. doi:10.1016/j.jnma.2018.01.006

47. Moskowitz D, Guthrie B, Bindman AB. The role of data in health care disparities in Medicaid managed care. *Medicare Medicaid Res Rev*. 2012;2 (4):mmrr.002.04.a02. doi:10.5600/mmrr.002.04.A02

48. McWilliams JM, Weinreb G, Ding L, Ndumele CD, Wallace J. Risk adjustment and promoting health equity in population-based payment: concepts and evidence. *Health Aff (Millwood)*. 2023;42(1):105-114. doi:10.1377/hlthaff.2022.00916

49. Chandra A, Skinner JS. Geography and racial health disparities. National Bureau of Economic Research working paper 9513. February 2003. Accessed September 17, 2024. https://www.nber.org/papers/w9513

50. Wallace J, Jiang K, Goldsmith-Pinkham P, Song Z. Changes in racial and ethnic disparities in access to care and health among US adults at age 65 years. *JAMA Intern Med*. 2021;181(9):1207-1215. doi:10.1001/jamainternmed.2021.3922

51. Krieger N, Chen JT, Coull B, Waterman PD, Beckfield J. The unique impact of abolition of Jim Crow laws on reducing inequities in infant death rates and implications for choice of comparison groups in analyzing societal determinants of health. *Am J Public Health*. 2013; 103(12):2234-2244. doi:10.2105/AJPH.2013.301350

52. Krieger N, Jahn JL, Waterman PD. Jim Crow and estrogen-receptor-negative breast cancer: US-born black and white non-Hispanic women, 1992-2012. *Cancer Causes Control*. 2017;28 (1):49-59. doi:10.1007/s10552-016-0834-2