Cureus
Part of SPRINGER NATURE

Open Access Original Article

Racial Disparity in Primary Cesarean Delivery Persists Within Physicians' Practices: A Multilevel Analysis of Physician Characteristics and Implicit Association Test Results

Kristina Roloff [1], Chioma Camille Okekpe [1], Robert Stowe [1], Guillermo J. Valenzuela [1]

1. Department of Women's Health, Arrowhead Regional Medical Center, Colton, USA

Corresponding author: Kristina Roloff, kristyroloff@gmail.com

Review began 10/19/2026
Review ended 10/26/2026
Published 10/29/2026

© Copyright 2026
Roloff et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.118804

Abstract

Objective

The objective of this study was to measure the Black-non-Black gap in primary cesarean delivery at a safety-net hospital and to test whether it varies by the delivering physician's demographics, training, or Implicit Association Test (IAT) result.

Methods

This retrospective cohort study included term, singleton, vertex births without a prior cesarean delivery, at any parity, at a single California safety-net teaching hospital from November 2016 to October 2020. Each birth was linked to the attending physician of record. We fit a multilevel logistic regression model with patients nested within physicians, adjusting for maternal age, BMI, gestational age, parity, and year. We tested physician gender, race, years in practice, training, and race-IAT result as modifiers of the patient-race effect. A physician fixed-effects model estimated the gap within physicians. Nulliparous, term, singleton, vertex (NTSV) births were analyzed as a secondary analysis.

Results

Of 7,249 singleton deliveries, 4,706 births attended by 21 physicians were analyzed. Primary cesarean delivery occurred in 98/478 (20.5%) Black patients and 502/4,228 (11.9%) non-Black patients, and the gap was present in every parity group. The adjusted OR (aOR) was 1.89 (95% CI, 1.47-2.44). In a physician fixed-effects model, it was 1.90 (95% CI, 1.47-2.44) before and 1.88 (95% CI, 1.45-2.42) after accounting for the physician. The disparity did not vary across physicians (random-slope variance estimated at zero), and no physician characteristic modified it (all interaction $p \geqslant 0.41$). Ten of 22 physicians submitted IAT results. The gap was not larger among physicians with a pro-White preference (interaction OR, 0.69; 95% CI, 0.36-1.31); it was present among those with no preference, with primary cesarean delivery occurring in 52/224 (23.2%) Black patients and 236/1,992 (11.8%) non-Black patients. In NTSV births, the aOR was 1.86 (95% CI, 1.24-2.79).

Conclusions

A twofold Black-non-Black gap in primary cesarean delivery was present across parity groups and persisted within individual physicians' practices. We found no evidence that it depended on physician demographics or IAT results, although the study could detect only moderate-to-large effects. Our findings point toward patient-, labor management-, and system-level drivers rather than which physician attends the birth.

Categories: Obstetrics/Gynecology
Keywords: cesarean delivery, health equity, implicit association test, implicit bias, maternal health, ntsv, physician characteristics, primary cesarean delivery, provider demographics, racial disparities

Introduction

Every primary cesarean delivery carries immediate surgical morbidity and shapes the course of subsequent pregnancies; therefore, preventing the first cesarean is central to reducing maternal morbidity [1,2]. Most quality-improvement efforts focus on nulliparous, term, singleton, vertex (NTSV) births, for which cesarean delivery is most discretionary [3,4]. However, primary cesarean deliveries in parous patients without a prior cesarean carry the same downstream consequences and are excluded from that measure.

Black patients have higher NTSV and primary cesarean delivery rates than White patients. This disparity persists after adjustment for clinical risk factors and payer [5-9]. It has also narrowed less than overall cesarean delivery rates under statewide quality-improvement efforts [4,7]. Differences in measured comorbidities, labor management, and induction explain only part of the disparity [8-10].

How to cite this article
Roloff K, Okekpe C, Stowe R, et al. (October 29, 2026) Racial Disparity in Primary Cesarean Delivery Persists Within Physicians' Practices: A Multilevel Analysis of Physician Characteristics and Implicit Association Test Results. Cureus 18(10): e118804. DOI 10.7759/cureus.118804

Clinician implicit bias is one proposed mechanism. Implicit bias refers to automatic associations that can shape judgment without conscious intent [11-13]. Cesarean delivery decisions during labor often rely on subjective assessments, such as labor progress or fetal heart rate interpretation, in which such associations could plausibly operate [10,12]. Pro-White implicit preferences measured by the race Implicit Association Test (IAT) are common among physicians. Some vignette studies have linked these preferences to differences in recommended care [14,15]. Whether IAT scores predict actual clinical behavior remains contested [16,17]. To our knowledge, no study has linked obstetricians' IAT results to the cesarean delivery decisions they actually made.

If physician attitudes or characteristics drove the disparity, its magnitude should differ between physicians. It should also decrease when Black and non-Black patients are compared within the same physician's practice. If the disparity reflects factors associated with the patient or the health care system, it should be similar across physicians and persist within individual physicians' practices.

We therefore examined the magnitude of the Black-non-Black gap in primary cesarean delivery overall and across parity groups; whether it persisted within individual physicians' practices and varied between physicians; and whether it was modified by the attending physician's gender, race, years in practice, training, or race-IAT result.

Materials And Methods

Design and setting

This retrospective cohort study used the labor and delivery electronic health record (IntelliSpace, Philips, Amsterdam, the Netherlands) of a county safety-net teaching hospital in Southern California. The hospital has approximately 2,500 births per year and no private attending physicians. Obstetrics and gynecology (OB/GYN) generalists and maternal-fetal medicine (MFM) subspecialists staff the obstetric service in 24-hour in-house shifts. Family medicine physicians, some with obstetric fellowship training, attend births on a separate teaching service. Most patients are insured by Medicaid or are medically indigent. The study included all available delivery records from November 9, 2016, to October 31, 2020. The institutional review board approved the study with a waiver of consent.

Cohort

The primary analysis included all term, singleton, vertex births without a prior cesarean, regardless of parity, so that every primary cesarean delivery contributed to the analysis. Each primary cesarean delivery adds immediate surgical morbidity and commits most patients to repeat cesarean delivery in subsequent pregnancies [1,2]. We also analyzed the nulliparous subgroup (NTSV), the established quality measure [3,4], as a prespecified secondary analysis.

The source file contained 7,469 delivery records. After 220 infant records from multiple gestations were removed, identified by the same mother delivering on the same date, 7,249 records remained. We then excluded, in order, terminations or records with no recorded mode of delivery (n = 8), residual multiple gestations documented in free text (n = 1), births before 37 0/7 weeks or with missing gestational age (n = 927), nonvertex presentations documented in the cesarean indication or labor complication fields (n = 88), records indicating any prior cesarean identified from repeat-cesarean and vaginal birth after cesarean fields or free text (n = 1,312), and records missing patient race or an identifiable attending physician (n = 207).

Recorded parity in this extract usually counted the current birth: gravida 1, para 1 was recorded approximately seven times as often as gravida 1, para 0. Parity was therefore grouped into three categories: unambiguously nulliparous (gravidity 1 or recorded parity 0), recorded para 1 with gravidity ⩾2 (which may represent a nulliparous patient or a patient with parity 1), and recorded para ⩾2. The NTSV subgroup comprised the unambiguously nulliparous category.

Several limitations of the source data affected variable definitions. BMI was not captured for deliveries in 2020 because of a change in the delivery database. Labor characteristics and cesarean indications were incompletely documented, and most indications were recorded only as references to the operative report; therefore, indication-specific analyses were not possible. Parity was recorded inconsistently in the electronic health record, with many first births coded as parity 1 at delivery. Parity was therefore categorized as described above, and an alternative definition was tested in a sensitivity analysis. Deliveries from July 2017 to June 2018 were underrepresented because of database corruption during that period.

Outcome and exposure

The outcome was cesarean delivery; operative and spontaneous vaginal births were combined. Race and ethnicity were self-reported at admission in a single combined field. Consistent with the study's primary question, the exposure was Black vs non-Black race, with the non-Black category comprising Hispanic, White, Asian, Pacific Islander, Native American, and other patients. Four-level race (White, Hispanic, Black, and other) was used for descriptive analyses and in a Black-vs-White sensitivity analysis.

Covariates

Models were adjusted for maternal age (continuous), BMI at delivery (<30, 30-34.9, 35-39.9 and ⩾40 kg/m², with a missing category), gestational age in completed weeks, parity category (omitted in the NTSV analysis), and calendar year.

BMI was unavailable for one-quarter of births because it was not captured after a change in the hospital's electronic record system; therefore, a complete-case BMI analysis was conducted as a sensitivity check. Type of labor (spontaneous, elective induction, or indicated induction) can lie on the pathway between race and cesarean delivery. It was therefore added in a separate model rather than treated as a confounder. Birthweight was not adjusted for because it is measured after delivery.

Physician characteristics and IAT

Each birth was attributed to the attending physician of record. Physicians self-reported their gender, race and ethnicity, years in practice since training (0-5 vs >5), and training (OB/GYN, family medicine with or without an obstetric fellowship, or MFM). Physician age (<45 vs ⩾45 years) was nearly collinear with years in practice and was not modeled.

Between November 2020 and April 2021, all attending physicians were contacted by email, with personal follow-up, and asked to complete the race-IAT (Project Implicit) and voluntarily share their results. Results were recorded in the test's categories: strong, moderate, or slight preference for White over Black people; no preference; or preference for Black over White people. IAT results were analyzed in three ways: as a three-level group (no preference as the reference, any pro-White preference, or did not submit); as a binary indicator among submitters (any pro-White preference vs no preference); and as an ordinal score among submitters (0 = no preference to 3 = strong pro-White preference).

Statistical analysis

Patient characteristics were compared across the four race groups using one-way ANOVA for means, the Kruskal-Wallis test for variables summarized as medians, and Pearson's χ^2 test for categorical variables. Each test was applied once per variable across all its categories. Crude cesarean delivery rates were compared using Pearson's χ^2 test. The primary analysis was a two-level logistic regression model with patients nested within physicians and a random intercept for physician. An empty model (M0) estimated between-physician variation; successive models added Black race (M1), the covariates described above (M2, the primary model), and type of labor (M3).

Between-physician variation was summarized using the latent-scale intraclass correlation coefficient and the median OR [18,19]. To test whether the disparity differed across physicians, M4 added a physician-level random slope for Black race. A race × parity interaction tested whether the gap differed by parity.

Each physician characteristic and IAT variable was tested individually as a cross-level modifier. Each model added the characteristic and its interaction with Black race to M2. To estimate the within-physician gap directly, we fitted a conventional logistic regression model with physician fixed effects, restricted to the 17 physicians with ⩾20 births, and compared the Black-race estimate with and without physician fixed effects.

Sensitivity analyses repeated M2, restricting the sample to Black and White patients; using complete-case BMI data; using generalized estimating equations (GEE) with an exchangeable correlation structure and robust standard errors; and adding patients with a prior cesarean delivery, excluding scheduled repeat cesarean deliveries, and adjusting for prior cesarean delivery.

The NTSV subgroup was analyzed using the same models, without the parity term. Its cesarean delivery rate was also compared with the Healthy People 2030 target of 23.6% using a one-sided binomial test with a normal approximation and exact Clopper-Pearson CIs [20].

At the physician level, we calculated each physician's crude Black-minus-non-Black cesarean delivery rate difference and correlated it with IAT score using Spearman's rank correlation among physicians with ⩾10 Black patients. Mixed models were estimated using GENLINMIXED in IBM SPSS Statistics for Windows, version 31.0 (released 2025; IBM Corp., Armonk, NY, USA). Fixed effects in the mixed models were tested using t tests with Satterthwaite degrees of freedom; fixed-effects logistic regression and GEE estimates were tested using Wald χ^2 tests. A two-sided $p < 0.05$ ($\alpha = 0.05$) was considered statistically significant. Interaction tests were treated as exploratory, and their CIs are reported in full.

Results

Cohort

Of 7,249 singleton delivery records, 4,706 term, singleton, vertex births without a prior cesarean delivery were included (Figure *1*). The cohort was predominantly Hispanic, reflecting the hospital's catchment area;

approximately one in 10 patients self-identified as Black, and nearly one-third of births were nulliparous and formed the NTSV subgroup (Table *1*). Compared with patients in the other groups, Black patients were more often recorded as para 1, more often had a BMI of 40 kg/m² or higher, were more often electively induced, and had smaller infants. Use of neuraxial analgesia and first-stage duration were similar across groups, whereas second-stage duration was shortest among Black patients. Apart from BMI and induction, these differences would not be expected to increase cesarean delivery risk, and the labor-course measures do not suggest longer or more difficult labor in Black patients.

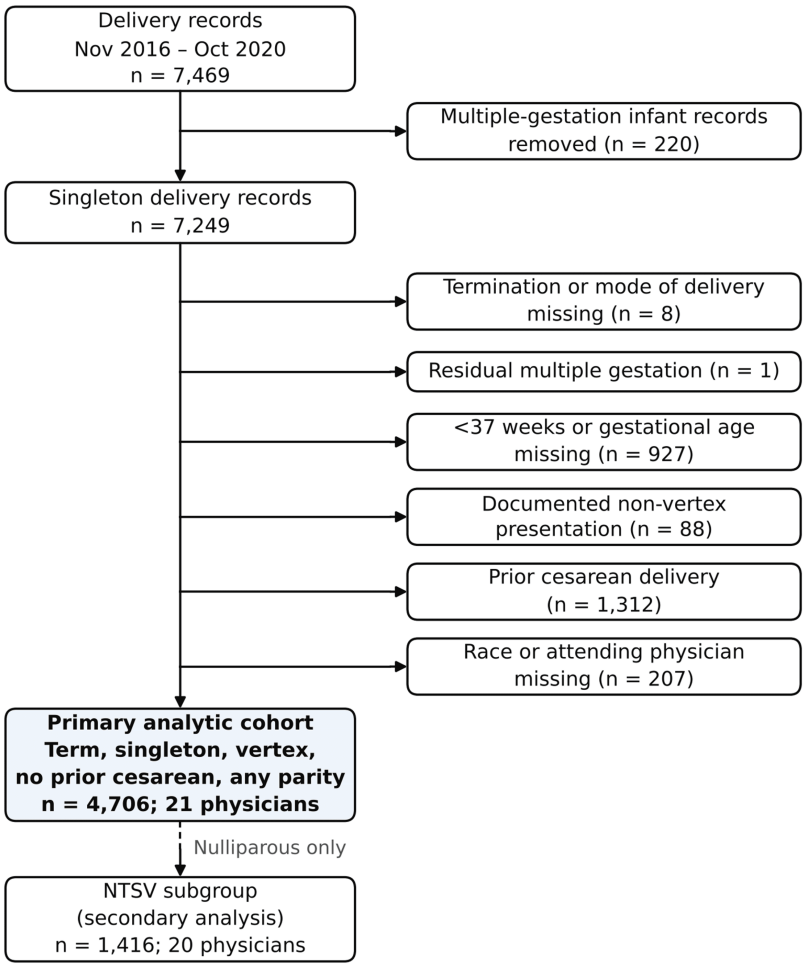


FIGURE 1: Cohort selection.

Exclusions were applied in the order shown. The primary analytic cohort comprised term, singleton, vertex births without a prior cesarean delivery, regardless of parity. The NTSV subgroup (nulliparous patients within the primary cohort) was analyzed as a secondary analysis.

NTSV: nulliparous, term, singleton, vertex

Characteristic	White (n = 368)	Hispanic (n = 3,706)	Black (n = 478)	Other (n = 154)	Test statistic	p-value
Age, years, mean ± SD	27.1 ± 5.6	26.8 ± 5.9	26.6 ± 5.2	28.6 ± 5.7	F(3, 4702) = 5.223	0.001
Parity, n (%)	-	-	-	-	$\chi^2(6) = 27.993$	<0.001
Nulliparous	109 (29.6)	1,120 (30.2)	136 (28.5)	51 (33.1)	-	-
Recorded para 1	42 (11.4)	359 (9.7)	80 (16.7)	24 (15.6)	-	-
Para ≥2	217 (59.0)	2,227 (60.1)	262 (54.8)	79 (51.3)	-	-
BMI, kg/m², mean ± SD (n with BMI)	32.3 ± 7.6 (253)	33.6 ± 7.3 (2,793)	33.9 ± 8.0 (356)	32.3 ± 8.4 (116)	F(3, 3514) = 3.327	0.019
BMI group, kg/m², n (%)	-	-	-	-	$\chi^2(12) = 29.452$	0.003
<30	110 (29.9)	998 (26.9)	128 (26.8)	56 (36.4)	-	-
30-34.9	70 (19.0)	810 (21.9)	102 (21.3)	23 (14.9)	-	-
35-39.9	43 (11.7)	521 (14.1)	54 (11.3)	15 (9.7)	-	-
≥40	30 (8.2)	464 (12.5)	72 (15.1)	22 (14.3)	-	-
Missing	115 (31.3)	913 (24.6)	122 (25.5)	38 (24.7)	-	-
Gestational age, weeks, mean ± SD	39.0 ± 1.1	38.9 ± 1.0	38.8 ± 1.0	38.9 ± 1.0	F(3, 4702) = 1.076	0.358
Type of labor, n (%)	-	-	-	-	$\chi^2(12) = 44.593$	<0.001
Spontaneous or augmented	137 (37.2)	1,681 (45.4)	166 (34.7)	73 (47.4)	-	-
Elective induction	71 (19.3)	523 (14.1)	102 (21.3)	16 (10.4)	-	-
Indicated induction	91 (24.7)	832 (22.5)	122 (25.5)	40 (26.0)	-	-
Prelabor cesarean	2 (0.5)	7 (0.2)	3 (0.6)	1 (0.6)	-	-
Missing	67 (18.2)	663 (17.9)	85 (17.8)	24 (15.6)	-	-
Neuraxial labor analgesia, n/N (%)[a]	188/357 (52.7)	1,851/3,641 (50.8)	250/466 (53.6)	78/151 (51.7)	$\chi^2(3) = 1.612$	0.657
First stage of labor, hours, median[b]	4	3.8	3.9	4.7	H(3) = 2.470	0.481
Second stage of labor, minutes, median[b]	18	17	15	24.5	H(3) = 11.315	0.01
Birthweight, g, median	3,300	3,310	3,170	3,250	H(3) = 41.744	<0.001
Cesarean delivery, n (%)	55 (14.9)	420 (11.3)	98 (20.5)	27 (17.5)	$\chi^2(3) = 37.273$	<0.001

TABLE 1: Characteristics of the primary cohort by patient race (n = 4,706).

Percentages use the column totals in the header (N = 368, 3,706, 478, and 154) as the denominators, except where n/N is shown. Each variable was tested once across all its categories; the test statistic and p-value appear in the variable's bold header row, and the indented rows beneath it represent its categories (including missing categories in the test). Tests included one-way ANOVA, F(degrees of freedom), for means; the Kruskal-Wallis test, H(df), for medians; and Pearson's χ^2(df) test for categorical variables. Other: Asian, Pacific Islander, Native American, and other.

[a] Denominator N = births with labor and recorded analgesia (prelabor cesareans excluded). For type of labor, 3 of 20 cells (15%) had expected counts <5, all in the prelabor cesarean row.

[b] Labor stage durations were recorded for 1,962 (first stage) and 3,544 (second stage) births; birthweight for 4,597.

The births were attended by 21 physicians (Table 2). OB/GYN generalists attended more than half of all births, and births to Black patients were distributed across physician gender, training, and IAT groups in broadly similar proportions to births overall. Thus, Black patients were not concentrated among particular groups of physicians.

Characteristic	Physicians, n (%) (N = 21)	Births, n (%) (N = 4,706)	Births to Black patients, n (%) (N = 478)
Female	12 (57.1)	2,106 (44.8)	223 (46.7)
Male	9 (42.9)	2,600 (55.2)	255 (53.3)
White	8 (38.1)	2,704 (57.5)	285 (59.6)
Hispanic	6 (28.6)	987 (21.0)	91 (19.0)
Black	3 (14.3)	430 (9.1)	42 (8.8)
Asian/South Asian	4 (19.0)	585 (12.4)	60 (12.6)
0-5 years in practice	8 (38.1)	2,029 (43.1)	217 (45.4)
>5 years in practice	13 (61.9)	2,677 (56.9)	261 (54.6)
OB/GYN	8 (38.1)	2,642 (56.1)	285 (59.6)
Family medicine ± OB fellowship	6 (28.6)	485 (10.3)	32 (6.7)
MFM	7 (33.3)	1,579 (33.6)	161 (33.7)
IAT: no preference	6 (28.6)	2,216 (47.1)	224 (46.9)
IAT: pro-White preference	4 (19.0)	1,159 (24.6)	132 (27.6)
IAT: not submitted	11 (52.4)	1,331 (28.3)	122 (25.5)

TABLE 2: Physicians and births in the primary cohort by physician characteristic.

Values are n (%); percentages use the column totals in the header as denominators (N = 21 physicians, 4,706 births, and 478 births to Black patients). One additional physician (22 in total) attended no births in the primary cohort. Among the 21 physicians, IAT results were not submitted by 11: nine did not respond, one declined, and one was deceased.

IAT: Implicit Association Test; MFM: maternal-fetal medicine; OB/GYN: obstetrics and gynecology

Cesarean rates

The overall primary cesarean delivery rate was low (600/4,706; 12.7%), but the rate was highest among Black patients, lowest among Hispanic patients, and intermediate among White and other patients (Table *1*). Black patients had nearly twice the crude odds of cesarean delivery as non-Black patients, and this excess was observed among nulliparous, recorded para 1, and para ≥2 patients alike (Table *3*). Baseline cesarean delivery rates among non-Black patients decreased steeply with parity, yet the relative gap was similar across all three groups and did not change over the four study years (race × year interaction OR, 1.07; 95% CI, 0.88-1.30).

Model	n	OR (95% CI)	Test statistic	p-value
Primary cohort	-	-	-	-
Crude (2 × 2)	4,706	1.91 (1.50-2.44)	$\chi^2(1) = 28.744$	<0.001
Crude, nulliparous: Black 43/136 (31.6%) vs non-Black 238/1,280 (18.6%)	1,416	2.02 (1.37-2.98)	$\chi^2(1) = 13.110$	<0.001
Crude, recorded para 1: Black 23/80 (28.8%) vs non-Black 69/425 (16.2%)	505	2.08 (1.20-3.60)	$\chi^2(1) = 7.077$	0.008
Crude, para ≥2: Black 32/262 (12.2%) vs non-Black 195/2,523 (7.7%)	2,785	1.66 (1.12-2.47)	$\chi^2(1) = 6.377$	0.012
M1: random intercept, race only	4,706	1.90 (1.49-2.42)	t = 5.194	<0.001
M2: + age, BMI, gestational age, parity, year (primary)	4,706	1.89 (1.47-2.44)	t = 4.948	<0.001
M3: M2 + type of labor	4,706	1.92 (1.47-2.50)	t = 4.818	<0.001
M4: M2 + random slope for race	4,706	1.92 (1.49-2.47)	t = 5.030	<0.001
Fixed effects: M2 covariates, no physician terms	4,686	1.90 (1.47-2.44)	Wald $\chi^2(1) = 24.618$	<0.001
Fixed effects: + physician indicators (within-physician)	4,686	1.88 (1.45-2.42)	Wald $\chi^2(1) = 23.422$	<0.001
GEE, exchangeable, robust SE	4,706	1.91 (1.49-2.44)	Wald $\chi^2(1) = 26.217$	<0.001
Sensitivity: complete-case BMI	3,518	1.93 (1.44-2.57)	t = 4.440	<0.001
Sensitivity: Black vs White only	846	1.43 (0.97-2.11)	t = 1.829	0.068
Sensitivity: + prior-cesarean patients, scheduled repeats excluded[a]	5,288	1.78 (1.42-2.24)	t = 4.960	<0.001
Secondary: NTSV subgroup	-	-	-	-
Crude (2 × 2)	1,416	2.02 (1.37-2.98)	$\chi^2(1) = 13.110$	<0.001
M2 (without parity term)	1,416	1.86 (1.24-2.79)	t = 2.984	0.003

TABLE 3: Odds of cesarean delivery for Black vs non-Black patients.

ORs compare Black with non-Black patients; n denotes the number of births included in each model. Tests included Pearson's χ^2(df) for crude 2 × 2 comparisons; t tests of the Black-race fixed effect for multilevel models (M1-M4 and sensitivity models), using Satterthwaite degrees of freedom; and Wald χ^2(df) tests for conventional fixed-effects logistic regression and GEE. For M0 (empty model), physician variance was 0.047 (latent intraclass correlation, 0.014; median OR, 1.23). In M4, the random-slope variance was estimated at the boundary of zero. Race × parity interaction ORs were as follows: recorded para 1, 1.11 (95% CI, 0.56-2.21; t = 0.292, p = 0.770); para ≥2, 0.95 (95% CI, 0.54-1.68; t = -0.164, p = 0.869). Fixed-effects models were restricted to the 17 physicians with ≥20 births.

[a] Additionally adjusted for prior cesarean delivery.

GEE: generalized estimating equations

Multilevel models

Physicians differed little in their overall use of cesarean delivery, and the racial gap was not explained by measured patient characteristics (Table *3*). The OR for Black patients was essentially unchanged before and after adjustment for maternal age, BMI, gestational age, parity, and year. Older age and a BMI of 40 kg/m² or higher were associated with higher odds of cesarean delivery, whereas higher parity was associated with lower odds. Black patients were more likely to undergo induction even after adjustment (adjusted OR (aOR), 1.54; 95% CI, 1.24-1.92), and both elective and indicated induction were associated with cesarean delivery compared with spontaneous labor. Nevertheless, adding type of labor to the model left the racial gap unchanged.

Three analyses examined whether the gap was concentrated among particular physicians. When the race effect was allowed to vary by physician, the between-physician variance was essentially zero, and model fit did not improve. When Black and non-Black patients treated by the same physician were compared directly using physician fixed effects, the estimate was virtually unchanged. At the individual physician level, 12 of the 13 physicians who attended at least 10 Black patients had a higher crude cesarean delivery rate among Black than among non-Black patients, and the differences clustered around a common value rather than being driven by a few outliers (Figure *2*). Together, these analyses suggest that the gap was shared across the service rather than concentrated in individual practices.

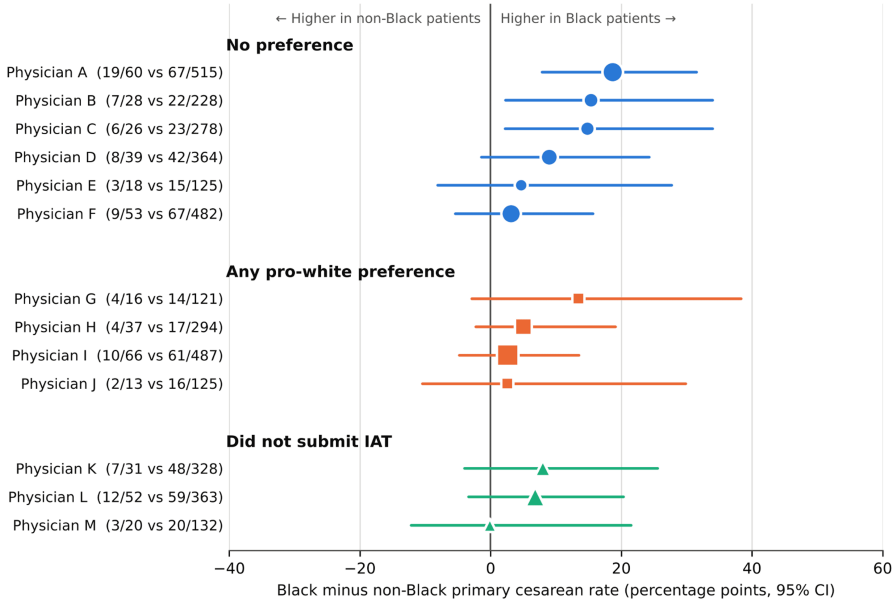


FIGURE 2: Black-minus-non-Black primary cesarean delivery rate by attending physician, grouped by race-IAT result.

Each row represents one physician who attended at least 10 Black patients in the primary cohort (n = 13). Row labels indicate cesarean deliveries/births among Black vs non-Black patients. Points show the crude difference in percentage points, with point size proportional to the number of Black patients; bars show Newcombe 95% CIs. Twelve of 13 physicians had a higher cesarean delivery rate among Black patients (median difference, +6.8 percentage points).

IAT: Implicit Association Test

Physician characteristics

Physician characteristics had little association with overall cesarean delivery rates and did not modify the racial gap (Table *4*). Among non-Black patients, MFM physicians had modestly lower cesarean delivery rates than OB/GYN generalists, while female physicians and family medicine physicians showed nonsignificant trends in the same direction. No characteristic, including physician race, modified the Black-non-Black difference, and every interaction CI included 1. Racial concordance did not narrow the gap, although only 42 Black patients were delivered by one of the three Black physicians.

Physician characteristic	Cohort (n)	aOR in non-Black patients (95% CI)	Interaction with Black race, OR (95% CI)	Test statistic (t)	p-value
Female (vs male)	Primary (4,706)	0.79 (0.57-1.08)	0.96 (0.57-1.60)	-0.167	0.867
Female (vs male)	NTSV (1,416)	0.61 (0.39-0.97)	1.16 (0.51-2.66)	0.354	0.723
>5 years in practice (vs 0-5)	Primary (4,706)	1.17 (0.84-1.63)	1.06 (0.63-1.77)	0.219	0.826
>5 years in practice (vs 0-5)	NTSV (1,416)	1.52 (0.94-2.45)	0.84 (0.36-1.91)	-0.426	0.67
MFM (vs OB/GYN)	Primary (4,706)	0.71 (0.51-0.996)	1.26 (0.73-2.17)	0.817	0.414
MFM (vs OB/GYN)	NTSV (1,416)	0.58 (0.34-0.98)	1.10 (0.46-2.62)	0.206	0.837
Family medicine (vs OB/GYN)	Primary (4,706)	0.67 (0.44-1.04)	0.77 (0.26-2.28)	-0.470	0.639
Family medicine (vs OB/GYN)	NTSV (1,416)	0.78 (0.42-1.46)	0.27 (0.03-2.59)	-1.136	0.256
Black physician (vs non-Black)	Primary (4,706)	0.90 (0.49-1.66)	0.97 (0.39-2.37)	-0.075	0.94

TABLE 4: Physician characteristics: association with cesarean delivery and modification of the racial gap.

Each characteristic was tested in a separate multilevel logistic regression model: M2 plus the characteristic and its interaction with Black race. The third column shows the effect of the characteristic among non-Black patients in that model. The test statistic and p-value refer to the interaction term (t test of the fixed effect, using Satterthwaite degrees of freedom). n denotes the number of births in the model. The racial concordance model was not run in the NTSV subgroup because too few Black patients had a Black physician to permit estimation.

aOR: adjusted OR; MFM: maternal-fetal medicine; NTSV: nulliparous, term, singleton, vertex; OB/GYN: obstetrics and gynecology

IAT

Fewer than half of the physicians shared their IAT results, but those who did attended most of the births in the cohort (Table *2*). Of the 10 submitters, four showed a pro-White preference, and six showed no preference; none showed a pro-Black preference. Of the 12 non-submitters, 10 did not respond, one declined, and one had died. Overall cesarean rates were similar across IAT groups after adjustment. The racial gap was present in every IAT group and was largest among physicians with no preference; however, it did not reach statistical significance in the smaller pro-White group (Table *5*, panel A). In adjusted models, neither a pro-White preference nor non-submission was associated with a larger gap. Among submitters, both the binary and ordinal IAT measures pointed in the direction opposite to the hypothesis (Table *5*, panels B and C). The physician-level correlation between IAT score and each physician's crude racial gap was also negative and nonsignificant (Table *5*, panel E).

IAT group or comparison	Black: CS n/N (%)	Non-Black: CS n/N (%)	OR (95% CI)	Test statistic	p-value
A. Crude Black vs non-Black comparison, primary cohort	-	-	-	-	-
No preference (six physicians)	52/224 (23.2)	236/1,992 (11.8)	2.25 (1.60-3.16)	$\chi^2(1) = 23.009$	<0.001
Any pro-White preference (four physicians)	20/132 (15.2)	108/1,027 (10.5)	1.52 (0.91-2.55)	$\chi^2(1) = 2.558$	0.11
Did not submit (11 physicians)	26/122 (21.3)	158/1,209 (13.1)	1.80 (1.13-2.87)	$\chi^2(1) = 6.320$	0.012
B. Adjusted interaction with Black race, primary cohort (n = 4,706)	-	-	-	-	-
Pro-White preference vs no preference	-	-	0.69 (0.36-1.31)	t = -1.143	0.253
Did not submit vs no preference	-	-	0.66 (0.36-1.21)	t = -1.344	0.179
Submitted vs did not submit	-	-	1.34 (0.76-2.37)	t = 1.005	0.315
C. Adjusted interaction, submitters only (10 physicians, 3,375 births)	-	-	-	-	-
Any pro-White preference vs none	-	-	0.69 (0.36-1.30)	t = -1.153	0.249
Per step of ordinal IAT score (0-3)	-	-	0.79 (0.53-1.17)	t = -1.184	0.236
D. Adjusted interaction, NTSV subgroup (n = 1,416)	-	-	-	-	-
Pro-White preference vs no preference	-	-	0.64 (0.23-1.82)	t = -0.839	0.401
Did not submit vs no preference	-	-	0.79 (0.30-2.10)	t = -0.473	0.636
E. Physician level (10 submitters with ≥10 Black patients)	-	-	-	-	-
IAT score vs crude Black-minus-non-Black gap	-	-	-	Spearman ρ = -0.548	0.101

TABLE 5: Race-IAT result of the attending physician and primary cesarean delivery.

Panel A: n/N = cesarean deliveries/births to patients of the specified race attended by physicians in the IAT group; percentages use N as the denominator; crude ORs with Pearson's χ^2 test (df = 1). Panels B-D: interaction ORs from multilevel logistic regression models (M2 plus the IAT variable and its interaction with Black race), tested using t tests of the fixed effect (Satterthwaite degrees of freedom); an OR below 1 indicates a smaller Black-non-Black gap than in the reference group. Panel E: Spearman's rank correlation. A dash (-) indicates a cell intentionally left blank.

CS: cesarean delivery; IAT: Implicit Association Test; NTSV: nulliparous, term, singleton, vertex

Secondary analysis: NTSV subgroup

The NTSV subgroup reproduced the primary findings (Table *3*). The NTSV cesarean rate among Black patients exceeded the Healthy People 2030 target of 23.6% (one-sided p = 0.018), whereas the rate among non-Black patients was below the target, and the aOR was similar to that in the primary cohort. In this subgroup, female and MFM physicians had lower cesarean rates among non-Black patients, but no physician characteristic or IAT group modified the racial gap (Table *4*; Table *5*, panel D).

Sensitivity analyses

The disparity estimate was robust to complete-case BMI analysis, the use of GEE, and the addition of patients with a prior cesarean who did not have a scheduled repeat cesarean (Table 3). The one exception was the restriction to Black and White patients, which yielded a smaller estimate whose CI just included 1. This reflects the choice of comparison group: White patients had a higher cesarean rate than the Hispanic patients who made up most of the non-Black group, and the White group was small.

Discussion

Principal findings

In this safety-net cohort, the overall primary cesarean rate was low: 600/4,706 (12.7%). Black patients nonetheless had approximately twice the adjusted odds of primary cesarean delivery as non-Black patients, and the gap was present at every parity level. It was essentially unchanged when Black and non-Black patients were compared within the same physician's practice (aOR 1.88), and its magnitude did not vary between physicians. Physician gender, race, years in practice, training, and race-IAT result were not associated with the magnitude of the gap. The gap was clearly present among physicians whose IAT showed no racial preference. The NTSV subgroup, the standard quality measure, yielded a similar result (aOR 1.86).

Comparison with previous studies

The magnitude of the disparity is consistent with earlier reports. Among low-risk nulliparous patients, Black patients had higher adjusted odds of cesarean delivery than White patients at an academic medical center [5], in an integrated health care system [9], and in a multicenter U.S. cohort [8]. In population-based data, the excess persisted after adjustment for clinical and sociodemographic factors [6], and racial gaps in obstetric quality measures have persisted over time [7], including under statewide quality-improvement efforts that lowered overall NTSV cesarean rates [4]. Our cohort adds two features to this literature. Nearly all patients were insured by Medicaid or were medically indigent, which limits payer as an explanation, and the overall cesarean rate was low. A relative disparity of similar magnitude in a low-cesarean unit suggests that lowering overall rates does not, by itself, close the gap.

Previous studies have mostly compared hospitals or populations. They could not separate a disparity arising within each physician's practice from one arising from differences in which physicians attend Black patients. The multilevel and fixed-effects design used here addresses that question directly.

Interpretation

The within-physician design is the main contribution of this analysis. If Black patients were disproportionately delivered by physicians with high cesarean rates, or if certain physicians performed cesareans more readily for Black patients, adding physician fixed effects should reduce the disparity, and the random-slope variance should be large. Neither occurred. The between-physician variance in the race effect was estimated at zero, and the gap behaved as a feature of the patients and the care system rather than of individual physicians.

Between-physician variation in overall cesarean use was also small (median OR, 1.23). Across U.S. hospitals, cesarean rates vary as much as tenfold, and wide variation persists even among low-risk women [21]. That pattern points to unit-level practice as a major determinant of cesarean use. In our unit, where in-house attending physicians share protocols and shifts, little variation remained between physicians, making unit-level processes a more plausible source of the gap than individual practice style.

Extending the analysis from NTSV to all primary cesareans strengthens this conclusion. The number of Black patients more than tripled, from 136 to 478, narrowing the CIs around every physician-level test. The gap appeared among parous as well as nulliparous patients, with no evidence that it differed by parity. Baseline cesarean rates among non-Black patients ranged from 195/2,523 (7.7%) among those with para ≥2 to 238/1,280 (18.6%) among nulliparous patients. A disparity of similar relative magnitude across such different baseline rates points to a shared mechanism rather than one specific to first labors.

This does not rule out bias. Implicit attitudes may be widespread enough to operate at the level of the whole unit, and labor and delivery decisions are made by teams. Residents, nurses, and successive attending physicians across shift changes all shape a labor that ends in cesarean delivery. Patient experience points in the same direction: in a national survey, women of color reported mistreatment by care providers during pregnancy and childbirth more often than White women [22]. A unit-wide or structural effect would produce the pattern we observed, and it cannot be distinguished from unmeasured patient risk with these data.

Our IAT findings fit the broader evidence on implicit bias in medicine. Systematic reviews have found pro-White implicit preferences in most samples of health professionals, at levels similar to those in the general population [11,13]. Their association with clinical decisions is inconsistent. In vignette studies, pro-White IAT scores predicted fewer thrombolysis recommendations for Black patients among internal medicine and emergency medicine residents [14] and less opioid prescribing for postoperative pain in one of four pediatric scenarios [15]. By contrast, implicit racial bias was not associated with 27 vignette-based treatment decisions

among acute care surgical clinicians [23], and most studies in a systematic review of physicians found no relationship between implicit bias and clinical decisions [12]. The most consistent associations have been with communication: clinicians with stronger pro-White preferences were more verbally dominant with Black patients and received lower ratings of interpersonal care from them [24]. A meta-analysis found that IAT scores poorly predict discriminatory behavior [16], and the test's developers caution against using it to classify individuals [17].

Our results are consistent with that literature. The gap was not larger among physicians with a pro-White IAT preference, and every point estimate was in the opposite direction, including the physician-level correlation, although none reached statistical significance. A cesarean delivery during labor is a discrete treatment decision, often shaped by a team over several shifts; this is the kind of outcome that the IAT has predicted least well. A single post hoc measurement linked to 10 physicians cannot exclude a meaningful effect of implicit bias. It does show that the disparity did not depend on the IAT results available to us.

Patient and process factors

Black patients more often had a BMI ⩾40 kg/m², one of the strongest covariates in the model. Yet adjustment for BMI and the other covariates barely changed the disparity (OR 1.90 unadjusted vs 1.89 adjusted). This agrees with multicenter and health-system studies in which measured clinical risk factors explained only part of the gap [8,9].

Black patients were more likely to undergo induction, even after adjustment (aOR 1.54), and induction was associated with cesarean delivery compared with spontaneous labor. Induction itself need not increase cesarean risk: in a large randomized trial of low-risk nulliparous patients, induction at 39 weeks lowered the cesarean rate compared with expectant management [25]. Our comparison is with spontaneous labor, a different reference group, and adjustment for type of labor did not reduce the gap. This is consistent with previous work showing that labor-management strategies do not explain racial differences [10]. How inductions are indicated, timed, and conducted may still differ by race, and a standardized induction protocol has been associated with smaller racial disparities in obstetric outcomes [26].

We could not examine indication-specific disparities. Others have reported that the excess among Black patients is concentrated in more subjective indications, such as labor arrest and nonreassuring fetal status [5,9,27]. This is the most important gap in our data.

The smaller Black-vs-White estimate (aOR 1.43) reflects our population. Hispanic patients, who accounted for 3,706/4,706 (78.8%) of the cohort, had the lowest cesarean rate, and White patients had an intermediate rate. Disparity estimates depend on the comparison group, and we report both.

Strengths and limitations

This study has several strengths. The primary cohort included every primary cesarean among term, singleton, vertex births, matching the outcome that drives future repeat cesareans, and the NTSV quality measure was reported alongside it. The population was relatively homogeneous in terms of payer, which limits confounding by insurance. The multilevel design separated variation between physicians from variation within physicians, and the fixed-effects analysis compared Black and non-Black patients cared for by the same physician. To our knowledge, this is the first study to link obstetricians' IAT results to the delivery outcomes of their own patients.

The study also has important limitations. With 21 physicians and 10 IAT submitters, it could detect only moderate-to-large effect modification; therefore, the null interaction findings are imprecise rather than evidence of no effect. Fewer than half of the physicians submitted IAT results, and the IAT was completed after the births to which the results were linked. Births were attributed to the attending physician of record. This may misclassify decisions made by residents or earlier attending physicians and would bias physician-level associations toward the null.

Because this was an observational study, residual confounding cannot be excluded, and the findings describe associations rather than causal effects. The small number of physicians who submitted IAT results limited the power to detect effect modification, so a contribution of implicit bias cannot be ruled out. The persistence of the gap within physicians is consistent with system-level contributors, but the study was not designed to identify specific mechanisms. Finally, the non-Black reference group was predominantly Hispanic, and nonvertex presentation was identified partly from free-text documentation. Results may differ in populations with other demographic compositions, and replication will require locally appropriate methods for case identification.

Data completeness was limited by the source system. Parity was recorded inconsistently, so it was modeled in three categories, one of which combines nulliparous and para 1 patients. Because the gap was similar across all three categories, residual confounding by parity is unlikely to explain it. Nonvertex presentation could be identified only when documented in free text, and most cesarean indications were recorded as "see operative report," so some breech cesareans likely remain in the cohort. Comorbidities and cesarean

indications were unavailable. BMI was unavailable for one-quarter of births after an electronic health record system change. Type of labor was missing more often for cesarean births, limiting the model that adjusted for it.

Race and ethnicity were recorded in a single combined field, preventing separate analysis of Hispanic ethnicity among Black patients. Finally, this was a single teaching hospital with in-house attending physicians and a predominantly Hispanic, Medicaid-insured population, so the findings may not generalize to private or community practice.

Conclusions

Racial disparities in primary cesarean delivery may reflect features of the obstetric unit and its systems of care, not only the individual clinicians who work within it. Assessing individual physicians' implicit attitudes may have value for reflection and education, but it is unlikely, on its own, to reveal where inequity arises or how to correct it. Units committed to birth equity can begin with measures within their own control: standardized labor and induction protocols, routine review of cesarean indications and outcomes stratified by race, and unit-wide training in equitable care.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Kristina Roloff, Chioma Camille Okekpe, Guillermo J. Valenzuela

Acquisition, analysis, or interpretation of data: Kristina Roloff, Robert Stowe, Guillermo J. Valenzuela

Drafting of the manuscript: Kristina Roloff

Critical review of the manuscript for important intellectual content: Chioma Camille Okekpe, Robert Stowe, Guillermo J. Valenzuela

Disclosures

Human subjects: Informed consent for treatment and open access publication was obtained or waived by all participants in this study. Arrowhead Regional Medical Center Institutional Review Board issued approval 21-12. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Spong CY, Berghella V, Wenstrom KD, Mercer BM, Saade GR: Preventing the first cesarean delivery: summary of a joint Eunice Kennedy Shriver National Institute of Child Health and Human Development, Society for Maternal-Fetal Medicine, and American College of Obstetricians and Gynecologists Workshop. Obstet Gynecol. 2012, 120:1181-93. 10.1097/aog.0b013e3182704880
2. Miller ES, Hahn K, Grobman WA: Consequences of a primary elective cesarean delivery across the reproductive life. Obstet Gynecol. 2013, 121:789-97. 10.1097/AOG.0b013e3182878b43
3. Main EK, Moore D, Farrell B, et al.: Is there a useful cesarean birth measure? Assessment of the nulliparous term singleton vertex cesarean birth rate as a tool for obstetric quality improvement. Am J Obstet Gynecol. 2006, 194:1644-51; discussion 1651-2. 10.1016/j.ajog.2006.03.013
4. Rosenstein MG, Chang SC, Sakowski C, et al.: Hospital quality improvement interventions, statewide policy initiatives, and rates of cesarean delivery for nulliparous, term, singleton, vertex births in California. JAMA. 2021, 325:1631-9. 10.1001/jama.2021.3816
5. Edmonds JK, Yehezkel R, Liao X, Moore Simas TA: Racial and ethnic differences in primary, unscheduled cesarean deliveries among low-risk primiparous women at an academic medical center: a retrospective cohort study. BMC Pregnancy Childbirth. 2013, 13:168. 10.1186/1471-2393-13-168
6. Huesch M, Doctor JN: Factors associated with increased cesarean risk among African American women: evidence from California, 2010. Am J Public Health. 2015, 105:956-62. 10.2105/AJPH.2014.302381
7. Janevic T, Egorova NN, Zeitlin J, Balbierz A, Hebert PL, Howell EA: Examining trends in obstetric quality measures for monitoring health care disparities. Med Care. 2018, 56:470-6. 10.1097/MLR.0000000000000919
8. Debbink MP, Ugwu LG, Grobman WA, et al.: Racial and ethnic inequities in cesarean birth and maternal morbidity in a low-risk, nulliparous cohort. Obstet Gynecol. 2022, 139:73-82. 10.1097/AOG.0000000000004620
9. Okwandu IC, Anderson M, Postlethwaite D, Shirazi A, Torrente S: Racial and ethnic disparities in cesarean

delivery and indications among nulliparous, term, singleton, vertex women. J Racial Ethn Health Disparities. 2022, 9:1161-71. 10.1007/s40615-021-01057-w

10. Yee LM, Costantine MM, Rice MM, et al.: Racial and ethnic differences in utilization of labor management strategies intended to reduce cesarean delivery rates. Obstet Gynecol. 2017, 130:1285-94. 10.1097/AOG.0000000000002343
11. Hall WJ, Chapman MV, Lee KM, et al.: Implicit racial/ethnic bias among health care professionals and its influence on health care outcomes: a systematic review. Am J Public Health. 2015, 105:e60-76. 10.2105/AJPH.2015.302903
12. Dehon E, Weiss N, Jones J, Faulconer W, Hinton E, Sterling S: A systematic review of the impact of physician implicit racial bias on clinical decision making. Acad Emerg Med. 2017, 24:895-904. 10.1111/acem.13214
13. FitzGerald C, Hurst S: Implicit bias in healthcare professionals: a systematic review . BMC Med Ethics. 2017, 18:19. 10.1186/s12910-017-0179-8
14. Green AR, Carney DR, Pallin DJ, Ngo LH, Raymond KL, Iezzoni LI, Banaji MR: Implicit bias among physicians and its prediction of thrombolysis decisions for black and white patients. J Gen Intern Med. 2007, 22:1231-8. 10.1007/s11606-007-0258-5
15. Sabin JA, Greenwald AG: The influence of implicit bias on treatment recommendations for 4 common pediatric conditions: pain, urinary tract infection, attention deficit hyperactivity disorder, and asthma. Am J Public Health. 2012, 102:988-95. 10.2105/AJPH.2011.300621
16. Oswald FL, Mitchell G, Blanton H, Jaccard J, Tetlock PE: Predicting ethnic and racial discrimination: a meta-analysis of IAT criterion studies. J Pers Soc Psychol. 2013, 105:171-92. 10.1037/a0032734
17. Greenwald AG, Brendl M, Cai H, et al.: Best research practices for using the Implicit Association Test . Behav Res Methods. 2022, 54:1161-80. 10.3758/s13428-021-01624-3
18. Merlo J, Chaix B, Ohlsson H, et al.: A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. J Epidemiol Community Health. 2006, 60:290-7. 10.1136/jech.2004.029454
19. Austin PC, Merlo J: Intermediate and advanced topics in multilevel logistic regression analysis . Stat Med. 2017, 36:3257-77. 10.1002/sim.7336
20. Reduce cesarean births among low-risk women with no prior births — MICH-06 . Accessed: October 13, 2026: https://odphp.health.gov/healthypeople/objectives-and-data/browse-objectives/pregnancy-and-childbirth/reduce-cesarean....
21. Kozhimannil KB, Law MR, Virnig BA: Cesarean delivery rates vary tenfold among US hospitals; reducing variation may address quality and cost issues. Health Aff (Millwood). 2013, 32:527-35. 10.1377/hlthaff.2012.1030
22. Vedam S, Stoll K, Taiwo TK, et al.: The Giving Voice to Mothers study: inequity and mistreatment during pregnancy and childbirth in the United States. Reprod Health. 2019, 16:77. 10.1186/s12978-019-0729-2
23. Haider AH, Schneider EB, Sriram N, et al.: Unconscious race and social class bias among acute care surgical clinicians and clinical treatment decisions. JAMA Surg. 2015, 150:457-64. 10.1001/jamasurg.2014.4038
24. Cooper LA, Roter DL, Carson KA, Beach MC, Sabin JA, Greenwald AG, Inui TS: The associations of clinicians' implicit attitudes about race with medical visit communication and patient ratings of interpersonal care. Am J Public Health. 2012, 102:979-87. 10.2105/AJPH.2011.300558
25. Grobman WA, Rice MM, Reddy UM, et al.: Labor induction versus expectant management in low-risk nulliparous women. N Engl J Med. 2018, 379:513-23. 10.1056/NEJMoa1800566
26. Hamm RF, Srinivas SK, Levine LD: A standardized labor induction protocol: impact on racial disparities in obstetrical outcomes. Am J Obstet Gynecol MFM. 2020, 2:100148. 10.1016/j.ajogmf.2020.100148
27. Williams A, Little SE, Bryant AS, Smith NA: Mode of delivery and unplanned cesarean: differences in rates and indication by race, ethnicity, and sociodemographic characteristics. Am J Perinatol. 2024, 41:834-41. 10.1055/a-1785-8843