BMJ Open

Explaining racial disparities in surgical survival: a tapered match analysis of patient and hospital factors

Karen B Lasater ,[1,2] Paul R Rosenbaum,[2,3] Linda H Aiken,[1,2] J Margo Brooks-Carthon,[1,2] Rachel R Kelz,[2,4] Joseph G Reiter,[5] Jeffrey H Silber,[2,5,6] Matthew D McHugh [1,2]

To cite: Lasater KB, Rosenbaum PR, Aiken LH, *et al*. Explaining racial disparities in surgical survival: a tapered match analysis of patient and hospital factors. *BMJ Open* 2023;**13**:e066813. doi:10.1136/bmjopen-2022-066813

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-066813).

Received 20 July 2022
Accepted 26 April 2023

For numbered affiliations see end of article.

Correspondence to
Dr Karen B Lasater;
karenbl@nursing.upenn.edu

ABSTRACT

Objectives Evaluate whether hospital factors, including nurse resources, explain racial differences in Medicare black and white patient surgical outcomes and whether disparities changed over time.

Design Retrospective tapered-match.

Setting 571 hospitals at two time points (Early Era 2003–2005; Recent Era 2013–2015).

Participants 6752 black patients and three sets of 6752 white controls selected from 107 001 potential controls (Early Era). 4964 black patients and three sets of 4964 white controls selected from 74 108 potential controls (Recent Era).

Interventions Black patients were matched to white controls on demographics (age, sex, state and year of procedure), procedure (demographics variables plus 136 International Classification of Diseases (ICD)-9 principal procedure codes) and presentation (demographics and procedure variables plus 34 comorbidities, a mortality risk score, a propensity score for being black, emergency admission, transfer status, predicted procedure time).

Outcomes 30-day and 1-year mortality.

Results Before matching, black patients had more comorbidities, higher risk of mortality despite being younger and underwent procedures at different percentages than white patients. Whites in the demographics match had lower mortality at 30 days (5.6% vs 6.7% Early Era; 5.4% vs 5.7% Recent Era) and 1-year (15.5% vs 21.5% Early Era; 12.3% vs 15.9% Recent Era). Black–white 1-year mortality differences were equivalent after matching patients with respect to presentation, procedure and demographic factors. Black–white 30-day mortality differences were equivalent after matching on procedure and demographic factors. Racial disparities in outcomes remained unchanged between the two time periods spanning 10 years. All patients in hospitals with better nurse resources had lower odds of 30-day (OR 0.60, 95% CI 0.46 to 0.78, p<0.010) and 1-year mortality (OR 0.77, 95% CI 0.65 to 0.92, p<0.010) even after accounting for other hospital factors.

Conclusions Survival disparities among black and white patients are largely explained by differences in demographic, procedure and presentation factors. Better nurse resources (eg, staffing, work environment) were associated with lower mortality for all patients.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ Tapered multivariate matching approach allows for sequentially matching black patients to different sets of white patients to understand which patient and hospital-level factors contribute to the observed outcomes disparity.
⇒ Measures of hospital nurse resources are derived directly from staff nurses.
⇒ Patient outcomes include 30-day and 1-year mortality and 30-day readmission.
⇒ Comorbidities used to match black and white patients may be fallible markers of clinical severity and frailty.

INTRODUCTION

Major National Academy of Medicine reports,[1 2] document the existence of racial disparities in hospital outcomes. Worse outcomes among black patients have been attributed to differences in illness severity,[3 4] disparities in treatment[5] and variation in hospital quality.[6 7] Each of these factors is a function of structural racism arising out of long-standing discriminatory systems, policies and institutions across sociopolitical domains including education, housing, criminal justice and healthcare.[8] Although systematic differences in hospitals where patients receive care may contribute to disparities,[9–12] little evidence specifies exactly which hospital factors are associated with worse disparities.

We focus on a modifiable aspect of hospitals—nurse resources. An evaluation of the role of nurse resources is warranted since they vary widely across hospitals[13 14] and a large literature shows that patients in hospitals where nurses care for fewer patients at a time, have a skill mix rich in registered nurses (RNs), high proportions of bachelors-educated nurses (BSNs) and a favourable nurse work environment, experience better outcomes including lower mortality.[14–17] Evidence suggests the survival benefits

conferred by better nurse resources accrue to all patients; however, they may be particularly beneficial for black patients.[18–21] Our motivation was to understand whether variation in hospital nurse resources differentially impact survival outcomes of black and white patients following surgery, whether improving these resources hold promise as an interventional target for reducing racial disparities and improving outcomes; and whether racial disparities in surgical outcomes have improved or worsened over time.

METHODS

Design and data sources

This is a retrospective multivariate tapered matching study that uses secondary data of patients and hospitals at two cross-sections in time: 2003–2005 (ie, Early Era) and 2013–2015 (ie, Recent Era). Data about patients were obtained from Centers for Medicare and Medicaid Services. Data about hospitals were obtained from the American Hospital Association Annual Survey which provided information on hospital size, the Healthcare Cost Report Information System data set which provided information on hospital teaching status and the RN4CAST-US survey which provided information about hospital nurse resources. Time periods for the Early and Recent Era were selected based on the availability of the RN4CAST-US survey data.

Patient population

The patient sample included non-Hispanic black and non-Hispanic white Medicare fee-for-service beneficiaries, who were 65.5 years or older and who were admitted to one of the study hospitals for general surgery (online appendix table 1) either between 1 January 2004 and 30 September 2006 or 1 January 2013 and 30 September 2015. Using race to characterise patients should not be interpreted as race representing innate biological differences. Race is a social construct; it reflects differences in experiences and exposure to systematic discrimination that produces observable harm and differences in health outcomes. Patient data included Research Identifiable Files: inpatient, outpatient, carrier (physician Part B), hospice and the master beneficiary summary file. Patients were excluded if there was missing data on age or sex, had an invalid date of death, or were enrolled in an Health Maintenance Organization (HMO) or lacked Part B coverage in the 6 months prior to their index hospitalisation.

For patients with multiple admissions, the index hospitalisation was defined by randomly selecting one admission. A 180-day look-back from the index admission was performed across all patient files to identify comorbidities. A 30-day mortality risk model to estimate each patient's probability of death at the time of admission was constructed using a 10% random sample of data that did not overlap with the analytical sample (online appendix tables 2A,B). Propensity scores to be a black individual were estimated using the covariates controlled in each match (online appendix tables 3–4). Other characteristics included age, sex, transfer-in status, emergent admission and 34 comorbidities. This is a retrospective study of patient claims data and thus there was no participation consent for patients.

Hospital sample

The RN4CAST-US is a large panel survey of RNs, conducted at two points in time (ie, 2005–2006; 2015–2016) in four large US states: California, Florida, New Jersey and Pennsylvania. Both surveys employed the same methodology—a modified Dillman *et al* approach[22] to randomly sample actively licenced RNs from state licensure lists.[23] Nurses consented to participation in the RN4CAST-US by completing the survey. Nurses reported the name of their employer, demographics, and details about resources in their hospital, including patient-to-nurse staffing ratios, nurse skill mix and the quality of the work environment. Our focus was adult, general, acute care hospitals in the four states.

Averages among RNs in the same hospitals were used to create aggregated hospital-level measures of nurse resources, consistent with prior research[15] and is a validated method of using multiple informants to generate organisational measures.[23] Our hospital-level measure of staffing, that is, patient-to-nurse ratios, is derived by taking the average number of patients per direct-care RN on medical-surgical units within the same hospital. Skill mix is the proportion of RNs to all nursing personnel (ie, RNs, licenced practice nurses, unlicenced assistive personnel). Nurse education is the hospital proportion of RNs holding a BSN or higher. Nurse work environment is derived from the National Quality Forum-endorsed 31-item Practice Environment Scale of the Nursing Work Index, comprised of five subscales: Nurse Participation in Hospital Affairs; Nursing Foundations for Quality of Care; Nurse Manager Ability, Leadership and Support of Nurses; Staffing and Resource Adequacy; Collegial Nurse–Physician Relations.[24]

Hospital nurse resources are presented as a three-category variable characterised by terciles of hospitals according to their percentile ranking which ranged from 0% (poorest nurse resources) to 100% (best nurse resources) based on a coherence rank score.[25] This approach gives equal weight to the four nurse resources in computing the coherence rank score, as we have done in prior studies,[15 26] since we had no a priori hypothesis that one resource would be more important to patient outcomes than another. The score describes how each hospital compared with others based on the four resources.[15 26] Hospitals present in both eras were ranked twice, once in each era. Ranks were formed by comparing hospitals two at a time—which of the two hospitals is better?—and then aggregating the pairwise comparisons. If hospital *i* had better nurse resources on all measures than hospital *j*, it received 1 point; if hospital *i* had worse nurse resources than hospital *j*, it lost 1 point,

or received—1 point; and if hospital *i* was better on some measures and worse on others, it received 0 points. The rank for hospital *i* is its total points, that is, the number of hospitals that were worse than hospital *i* minus the number that were better than hospital *i*.

Outcomes

Thirty-day and 1-year mortality (defined as a death within 30 days and 1 year of admission, respectively). Thirty-day readmission (or death) outcomes are reported in the online appendix tables 5–6). Mortality and readmission outcomes were 'all-cause' and determined by data reported in the Centers for Medicare & Medicaid Services (CMS) patient files.

Statistical analysis

Matching methodology

The tapered multivariate matching approach[3 27–30] sequentially matches the same black patients to different sets of white patients, controlling for consecutively more variables to understand the contribution of various factors to the outcomes disparity.[27] The goal is to understand the extent of and factors driving the racial disparities in outcomes between black and white patients. By incrementally matching white patients to black patients on additional variables, we can directly observe how the matched white cohort changes with respect to their outcomes. Our tapered matching procedure includes three tapers (or sets of matches). First, the demographics taper included variables for age, sex, state and year of procedure. Second, the procedure taper included all the variables from the demographics taper and added International Classification of Diseases (ICD)-9 principal procedure codes. Third, the presentation taper included all the variables from the procedure and demographics tapers and added patient risk factors related to health status at the time of surgery, including 34 comorbidities, a mortality risk score, emergency admission, transfer status and predicted procedure time. Patients were exactly matched within era and state (with New Jersey and Pennsylvania combined), for 136 ICD-9 procedure codes and mortality risk quintile (online appendix tables 3–4). Fine balance and distance minimisation techniques were used to make matched groups as similar as possible.

Statistical methods

Comparisons within pairs used McNemar's test and conditional logit regression. We compared the black–white difference in the Early and Late Eras to test whether the disparity changed over time. These analyses used Gart's test[31] to compare disparities in the Early Era to disparities in the Recent Era.[32] Conditional logit regression models were performed at the presentation match (ie, using the white patient cohort that was similar to the black patients with respect to demographic, procedure and presentation variables), and using data from both eras combined to test nurse resources, race and combinations of their interactions, accounting for structural hospital characteristics (ie, size, teaching status, technology capabilities, general surgery volume). Hospital size was defined as large (>250 beds), medium (101–250 beds) or small (≤100 beds). Teaching status was defined by the medical resident to beds (RB) ratio (non-teaching: 0 RB; minor teaching: >0 RB and ≤0.25 RB; major teaching: >0.25 RB). A high technology hospital was defined as having the capability to perform major organ transplantation and/or open-heart surgery. General surgery volume was defined as a continuous measure of the number of general surgical cases per 100 patients in each hospital during the study period.

Patient and public involvement

Patients were not directly involved in the development of the research question or outcome measures, the study design or the recruitment and conduct of the study.

RESULTS

Quality of patient matches

The matches are shown in table 1 (Recent Era) and online appendix table 7 (Early Era). Table 1 describes 4964 black patients and three sets of 4964 white controls—selected from a population of 74 108 white patients. In each taper, white controls become more like the black patients. Matched variables (ie, left of the zigzag line) were similar: the standardised differences in means never exceeded 0.11 SDs. Unmatched variables (ie, right of the zigzag line) show the disparity prior to matching. Comparisons in the demographic match reveal differences in the types of procedures black and white patients receive. For example, black patients underwent a laparoscopic cholecystectomy less (18.9%) than white patients (21.3%, p<0.01). Black patients had more comorbidities, and in some cases were much more likely to have a chronic condition such as diabetes (51.3% vs 32.8%), despite being 1.5 years younger on average. The demographics match removed age, sex, state and year of procedure differences, the procedure match included demographics match variables and removed differences in procedures and the presentation match included all demographics and procedure variables and further matched on variables reflecting health status by selecting white controls that had similar mortality risk and comorbidity burden as black patients. The cohort of white patients in the presentation match are different than the 'unmatched' white patients, in that the white patients in the presentation match have a substantially higher burden of comorbidities that are more comparable to the burden of comorbidities observed in the black population.

We made no attempt to match on measures of socioeconomic status (SES), including dual-eligibility and neighbourhood-level socioeconomic variables (ie, median household income, percentage of high school graduates, percentage of college graduates) because SES variables are highly correlated with race in the USA. Black patients were nearly four times more likely to be dual-eligible compared

Table 1 Quality of matches for selected§ variables, Recent Era (2013–2015)

Variable	Black patients (n=4964)	Tapered matches: Presentation + procedure + demographics (n=4964)	Tapered matches: Procedure + demographics (n=4964)	Tapered matches: Demographics (n=4964)	White patients (unmatched) (n=74 108)
State (%)					
California	24.3	24.3	24.3	24.3	**27.8‡**
Florida	34.1	34.1	34.1	34.1	35.4
New Jersey/Pennsylvania	41.6	41.6	41.6	41.6	**36.8‡**
Year of procedure (%)					
2013	23.1	23.2	23.1	23.1	22.9
2014	43.7	43.7	43.7	43.7	44.6
2015	33.2	33.1	33.2	33.2	32.6
Age at procedure	75.5	**75.0†**	75.4	75.5	**77.0‡**
% Male	39.3	39.3	39.3	39.3	**44.7‡**
Procedures (%)					
Laparoscopic cholecystectomy (5123)	18.9	18.9	18.9	**21.3†**	**21.6‡**
Open right hemicolectomy (4573)	7.6	7.6	7.6	**6.4***	**6.6†**
Partial resection of small intestine (4562)	7.0	7.0	7.0	**5.7***	**5.6‡**
Laparoscopic right hemicolectomy (1733)	4.8	4.8	4.8	4.3	4.4
Open cholecystectomy (5122)	3.1	3.1	3.1	3.2	3.3
Selected comorbidities (%)					
Hypertension	93.2	93.3	**84.9‡**	**84.7‡**	**85.1‡**
Diabetes	51.3	51.1	**33.7‡**	**32.6‡**	**32.8‡**
Congestive heart failure	26.1	25.9	**18.0‡**	**18.2‡**	**19.4‡**
Renal dialysis	42.2	41.7	**26.9‡**	**26.1‡**	**28.4‡**
Renal failure	14.0	6.5	**5.5‡**	**5.7‡**	**4.1‡**
Paraplegia	6.1	**4.5**	**2.1‡**	**2.1‡**	**2.1‡**
Mortality Risk Score (prob)	0.069	0.067	**0.055‡**	**0.050‡**	**0.056‡**
Emergency admission (%)	56.9	58.4	**50.2‡**	**50.2‡**	**50.5‡**
Transfer status (%)	1.1	1.0	0.9	0.9	**0.8***
Anaesthesia time (minutes)	155	**150‡**	**150‡**	**152‡**	**151‡**
Dual-eligible (%)	37.4	**14.8‡**	**11.7‡**	**10.6‡**	**10.4‡**
Neighbourhood median household income (US$)	24 267	**32 070‡**	**32 970‡**	**32 843‡**	**32 755‡**
Neighbourhood high school graduate (%)	83.2	**88.8‡**	**89.3‡**	**89.2‡**	**89.2‡**
Neighbourhood college graduate (%)	32.8	**39.9‡**	**40.9‡**	**40.9‡**	**40.9‡**

The zigzag diagonal line indicates which variables are controlled in each match: variables to the right of the line are not controlled. The table shows only a few of the variables,—in particular, a few of the surgical procedures—that were controlled in each match. Bolded numbers represent significant differences.

*<0.005

†<0.01

‡<0.001.

§The complete balance tables with all variables are available in online appendix table 4 for Recent Era (2013–2015) patient matches. Dual-eligible is a beneficiary of both Medicare and Medicaid. Measures of patient socioeconomic status were obtained through the American Community Survey and are based on neighbourhood-level characteristics: median household income, percentage of high school graduates and percentage of college graduates.

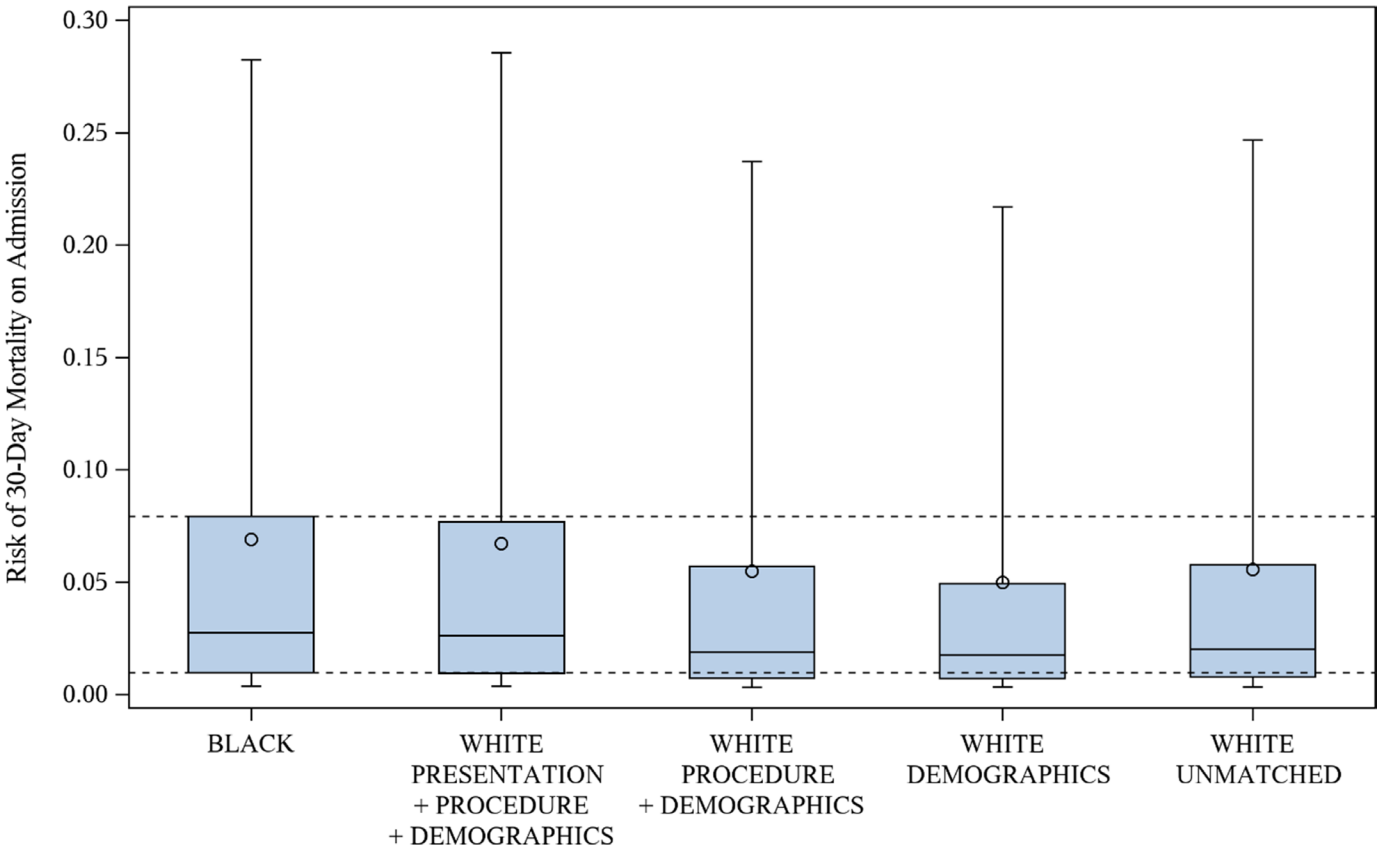

Figure 1 Distribution of Mortality Risk Score for the black study population, the total white study population and three matched white populations, Recent Era (2013–2015). The tails of each box plot represent the lower 5% and upper 95% of the distribution. The mortality risk estimates presented here are based on risk at the time of admission. Early Era results look similar and are presented in appendix figure 1. Summary: Until matched for surgical procedure and comorbid conditions in the 'White Presentation' match, black patients had a combination of surgical procedures and comorbid conditions that placed them at elevated risk of death compared to white controls.

with unmatched whites, and more likely to live in neighbourhoods with markers of lower SES. After matching on demographic, procedure and presentation variables, white controls looked more like black patients with respect to SES indicators, however large and important differences remained (eg, 37.4% black patients were dual-eligible vs 14.8% of white controls, p<0.001).

Figure 1 demonstrates differences in black and white patients' estimated mortality risk on admission prior to matching (ie, white unmatched) and at each taper of the match. The largest disparity in estimated mortality risk is observed in the demographics match—likely because this match requires patients to be the same on age and sex, which selects for white controls who were 1.5 years younger than the typical white patient and fewer men. As we move through the tapers, the racial disparity in estimated mortality risk narrows. The result of the matching process is a white control group that is profoundly different than the initial white population. Online appendix figure 1 presents comparisons in the Early Era with similar findings.

OUTCOME RESULTS

Mortality outcomes for black patients and the three sets of white controls are reported in table 2. In the Early and Recent Eras, after matching white controls with similar demographics as the black cohort (ie, demographics match), we observe higher 1-year mortality among black patients. One-year mortality differences narrow after matching on procedure but remain significantly higher among black patients. After selecting white controls that presented as sick as black patients (ie, presentation match), 1-year mortality differences become statistically insignificant. Thirty-day mortality differences diminished after matching on procedure. The bottom most panel of table 2 reports whether the black–white difference changed over time, defined by the black–white difference in the Recent Era minus the black–white difference in the Early Era. Survival disparities did not change significantly over the two eras separated by 10 years. Survival curves of black patients and white controls are presented in figure 2. In the Early and Recent Eras, white controls at the presentation match had the lower probability of survival in the time period most proximal to hospital admission; however, at 1 year from hospitalisation black patients had lower survival odds. The mortality in white control groups changed significantly as more covariates were controlled in all cases, except the move from the demographics control group to the demographics +

Table 2 Mortality outcomes for black study population and three matched white populations: Early Era (2003–2005), Recent Era (2013–2015) and the difference between the eras to evaluate whether the black–white difference is different in the two eras

			Tapered matches of white controls		
		Black patients	Presentation + procedure + demographics	Procedure + demographics	Demographics
Early Era (2003–2005)	1-year mortality	21.45%	20.51%	17.54%***	15.52%***
	30-day mortality	6.71%	7.81%**	6.47%	5.60%**
Recent Era (2013–2015)	1-year mortality	15.87%	16.16%	12.99%***	12.29%***
	30-day mortality	5.70%	7.88%***	5.74%	5.42%
Difference in difference (Recent—Early)	1-year mortality	--	−1.23%	−1.03%	−2.35%
	30-day mortality	--	−1.08%	−0.28%	−0.83%

Black–white difference between eras is defined by the black–white difference in Recent Era minus the black–white difference in Early Era. Significance tests for binary variables used McNemar test (*<0.05, **<0.01, ***<0.001). For the difference in difference across eras, Gart's test for binary outcomes was used (+<0.05, ++<0.01, +++<0.001). The symbols were marked in the later era if the difference in difference was significant.

procedure control group in the Recent Era, where the difference in mortality at 30 days and 1 year was not significant (online appendix table 8).

In tables 1 and 2, the statistics (ie, comparisons of differences in variables between black and white patients) are crude in the sense that we do not employ regression modelling for adjustments. With each tapered match of our multivariate tapered matching procedure we are selecting a new cohort of white patients who more closely resemble the black patients on the variables of that match. Thus, there is no formal adjustment procedure occurring since these are the observed characteristics of the white and black cohorts.

Conditional logit models further analyse black–white patient pairs (table 3). These models attempt to tease apart race, nurse resources, their interaction and other hospital attributes. Each model has a parameter for each matched pair, representing the many covariates that were made similar in that pair. If a covariate is matched, it is already in the model via these pair effects. Aside from the many pair effects, race and hospital-level variables are the only variables in the model. All hospital-level

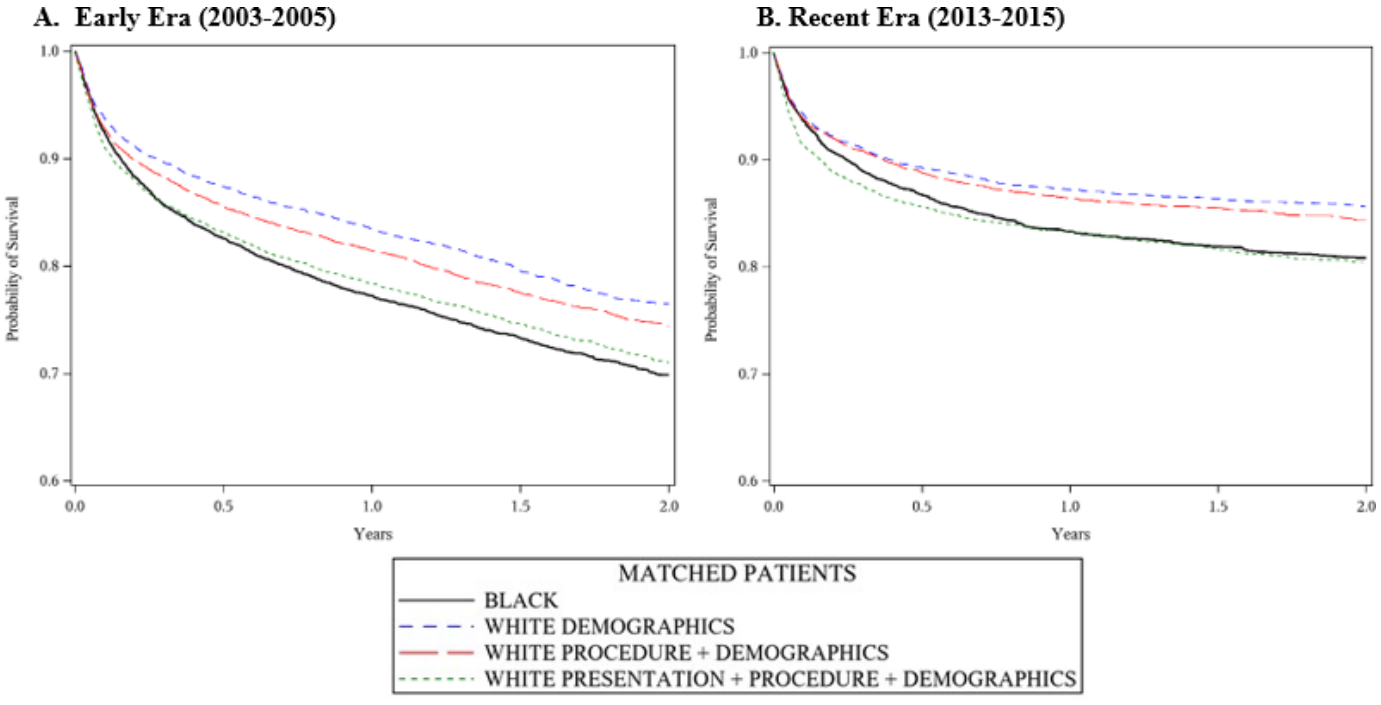

Figure 2 Kaplan-Meier lot for survival for black study population and three matched white populations. The substantially higher mortality among black patients is most evident over a longer span of time, is not concentrated in the brief period around surgery and reflects a greater burden of comorbid conditions and a more frequent need for higher risk procedures. Black and white patients had lower mortality in the Recent Era (2013–2015), but there is no clear indication that the black–white disparity has diminished.

Table 3 Effect of race and hospital nurse resources on 30-day and 1-year mortality odds, after matching patients on demographics, procedure and presentation variables

	30-day mortality				1-year mortality			
	Model 1a	Model 2a	Model 3a	Model 4a	Model 1b	Model 2b	Model 3b	Model 4b
Variables in the model	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Black (vs white)	0.77*** (0.69 to 0.85)	0.79*** (0.71 to 0.88)	0.75*** (0.64 to 0.87)	0.75*** (0.64 to 0.88)	1.03 (0.96 to 1.11)	1.05 (0.98 to 1.13)	1.05 (0.95 to 1.17)	1.05 (0.94 to 1.16)
Nurse resources (high vs low)	—	0.58*** (0.46 to 0.74)	0.59*** (0.46 to 0.74)	0.60** (0.46 to 0.78)	—	0.75*** (0.64 to 0.88)	0.75*** (0.64 to 0.88)	0.77** (0.65 to 0.92)
Nurse resources (middle vs low)	—	0.83 (0.68 to 1.00)	0.82* (0.68 to 1.00)	0.83 (0.68 to 1.01)	—	0.91 (0.80 to 1.03)	0.91 (0.80 to 1.03)	0.91 (0.80 to 1.04)
Black*nurse resources (high vs low)	—	—	0.91 (0.78 to 1.07)	0.92 (0.78 to 1.08)	—	—	1.01 (0.90 to 1.13)	1.02 (0.91 to 1.13)
Black*nurse resources (middle vs low)	—	—	0.95 (0.80 to 1.13)	0.95 (0.80 to 1.13)	—	—	0.99 (0.88 to 1.11)	0.99 (0.88 to 1.11)
Major teaching hospital	—	—	—	0.93 (0.72 to 1.21)	—	—	—	0.97 (0.82 to 1.15)
Minor teaching hospital	—	—	—	0.97 (0.80 to 1.18)	—	—	—	1.03 (0.90 to 1.17)
Large size (>250 beds)	—	—	—	0.98 (0.81 to 1.19)	—	—	—	0.97 (0.85 to 1.10)
High technology hospital	—	—	—	1.08 (0.90 to 1.29)	—	—	—	1.07 (0.94 to 1.21)
General surgery volume, per 100 patients	—	—	—	0.99 (0.96 to 1.01)	—	—	—	0.99 (0.97 to 1.00)
Test for improvement in fit with greater model complexity								
X^2	—	20.03	1.25	1.80	—	12.40	0.13	4.94
Df	—	2	2	5	—	2	2	5
P value	—	<0.0001	0.5350	0.8756	—	0.0020	0.9363	0.4227

Conditional logit models show the effects of race and hospital nurse resources for pairs of black and white patients who have been closely matched on demographic characteristics (age, sex, state and year of procedure), procedure (ICD-9 principal procedure code) and presentation (34 comorbidities, mortality risk score, propensity score for being black, emergency admission indicator, transfer status indicator, predicted procedure time). Data from both eras are combined in this analysis. Nurse resources represent a three-category variable characterised by terciles of hospitals according to their percentile ranking. The general surgery volume variable represents the effect of a 100-patient increase in general surgery volume on patient odds of 30-day mortality. Black*nurse resources indicate an interaction of the variables Black and Nurse Resoures. Summary: High levels of nursing resources were associated with substantially lower mortality for both black and white patients, with no indication of interaction.
*p<0.05; **p<0.01; ***p<0.001.
ICD-9, International Classification of Diseases-9.

variables (except volume) are two-category variables, and the coefficient is the OR comparing the two categories. The volume variable is in units of 100 patients on the logit scale, such that an OR 0.9, for example, would be comparing two hospitals, one with 100 more general surgery patients than the other. Model 1a is like the table 2 presentation match in which black patients have lower odds of 30-day mortality (OR 0.77, 95% CI 0.69 to 0.85, p<0.001). In Model 2a, high nurse resources are associated with substantially lower mortality (OR 0.58, 95% CI 0.46 to 0.74, p<0.001), and this pattern appears to be the same or nearly so for black and white patients. As in table 2, 1-year mortality outcomes are not significantly different among black and white patients who were matched on demographic, procedure and presentation characteristics (Models 1b–4b). High nurse resources are strongly associated with lower 1-year mortality (Model 2b), apparently in the same way for blacks and whites (Model 3c), persisting even after adjusting for hospital-level characteristics (Model 4b). Findings were similar for 30-day readmission (online appendix table 6).

The simplest model that fits well includes race and nurse resources (Models 2a and 2b). The addition of interactions between race and nurse resources or additional hospital attributes did not improve the model. This is evident in the test-statistics reported in the bottom of table 3 which describe the improvement in fit for each model compared with the prior model. P values>0.05 mean we fail to reject the simpler model in favour of the more complex model.

DISCUSSION

Study results reveal outcomes disparities are largely explained by significant differences in clinical presentation between black and white patients. Among black and white patients matched for demographics (ie, age, sex, state and year of procedure), we found significantly higher 30-day and 1-year mortality among black patients. This is consistent with prior evidence of racial outcomes disparities in surgical patients.[2 33 34] Black patients in our sample had a heavier burden of comorbidity and mortality risk than white patients. Despite being younger, black patients had more comorbidities, more emergency admissions and higher mortality risk on admission. Black patients also underwent procedures at different percentages. Only after closely matching patients to account for these differences did the mortality advantage for white controls disappear.

Our research is not the first to find higher mortality among white patients after accounting for racial differences in clinical presentation.[3 4 18 35–37] Cumulative effects of centuries of systematic discrimination in virtually all domains of life (eg, education, housing, criminal justice, policy benefits, job opportunities, pay, political power and access to high quality healthcare) underlie observable clinical presentation differences. Thus, system-level reforms across these domains are necessary to begin to undo the harms generating differences in health status and survival outcomes.

Our second major finding is that surgical disparities—at least for general surgeries—have not narrowed overtime. This is in contrast to what Mehtsun and colleagues found[38]—though that analysis focused on eight procedures and included orthopaedic and vascular surgeries. In our study, we found that while mortality and readmissions were lower in the Recent Era (2013–2015) for both black patients and white controls, the differences between the two groups remained unchanged overtime.

Our third major finding is that differences in hospitals are a significant contributor to variation in outcomes for all surgical patients, both black and white. Specifically, receiving care in hospitals with better nurse resources was associated with lower odds of death, even after accounting for other hospital factors (ie, teaching status, technology capability, size and surgery volume). Being in a hospital with high nurse resources predicted a much larger reduction in mortality than did race. High nurse resources predicted lower mortality for both black and white patients, to the same or similar degree. Some research has shown that nurse resource deficiencies result in even worse outcomes for black patients,[18–21] but perhaps this difference is a function of our use of a composite measure which simultaneously evaluates all four aspects of nurse resources versus isolating the effect of a single resource; other investigations focused mainly on nurse staffing.

That our results suggest that better nurse resources, as opposed to other hospital factors are associated with higher survival outcomes, is important. Whereas the other hospital factors we measured here are difficult to modify, nurse resources are modifiable through actions of hospital administrators or policy intervention. Hospital administrators can make it their strategic priority to staff greater numbers of nurses, including higher proportions of BSN-prepared nurses and a richer skill mix of RNs, as well as improve their nurse work environments via management reforms and evidence-based interventions like the American Nurses Credentialling Center Magnet Program.[39 40] At the policy-level, states can follow the example of California—the first and only state to legislate hospitals hire enough nurses to safely care for patients. The result of this policy has improved nurse staffing ratios and made more even the staffing variability across the state.[41 42] Recent studies show wide variation in the average nurse staffing ratios within states,[13 43] ranging from 3.3 to 9.7 patients-per-nurse on medical-surgical units.[13] If other states followed California's example by enacting minimum safe nurse staffing policies, it would raise the floor on hospital nurse staffing while making more even the variability across hospitals.

Limitations

Despite carefully matching on demographic, procedure and presentation differences, we are unable to account for possible within-hospital differences experienced by black and white patients, for example, the possibility

of selection bias wherein surgeons may be less likely to operate on black patients compared with similarly ill white patients.[5 44] Thus, our analysis of surgical patients may include somewhat healthier black patients than their matched white controls. Comorbidities used for matching patients may be fallible markers of clinical severity and frailty or have within-category variation leading to residual differences in presentation despite careful and comprehensive matching. Next, although we use the white population as the reference group, it should not be interpreted that the white population's outcomes are the ideal referent or the best that could be achieved in terms of outcomes for black patients. Studies using other referent groups (eg, not-low-SES white[45 46]) would be useful, as would research within the black population alone to understand possible strengths that could be leveraged to improve outcomes that may be unique to the population. Finally, our tapered-matched design makes transparent the comparisons between black and white patients and shows that the black–white survival disparity is largely explained by differences in demographic, procedure and presentation factors. It is possible; however, that unmeasured confounders may be important to further investigate health disparities after discharge, which we did not do in this study but could be relevant to survival outcomes over a year following surgery.

CONCLUSIONS

In summary, there is a large racial disparity in mortality among Medicare patients undergoing general surgery. Black and white patients present differently even when undergoing the same procedure. Despite being younger, black patients are more likely to have higher comorbidity burden and greater risk of mortality. We found racial outcomes disparities following surgery have not improved over the decade, but organisational and policy reform have the potential to improve outcomes for black and white patients alike. Even after accounting for demographic, procedure and presentation differences, better nurse resources—a modifiable feature of hospitals—were significantly associated with improved survival for both black and white patients.

Author affiliations
[1]Center for Health Outcomes and Policy Research, University of Pennsylvania School of Nursing, Philadelphia, Pennsylvania, USA
[2]The Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, Pennsylvania, USA
[3]The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania, USA
[4]Department of Surgery, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA
[5]Center for Outcomes Research, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA
[6]Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Twitter Karen B Lasater @k_lasater, Linda H Aiken @LindaAiken_Penn and Matthew D McHugh @matthewdmchugh

Contributors All authors meet the criteria recommended by the International Committee of Medical Journal Editors (ICMJE). PRR, LHA, JMB-C, RRK, JHS and MDM contributed to the original idea and design of the study. KBL, LHA, JMB-C and MDM contributed to the collection of nurse survey data. JGR conducted the data analysis. All authors contributed to the interpretation of the data and preparation of the submitted manuscript. All authors approved the submitted manuscript.

Funding This research was funded by grants from the National Institute on Minority Health and Health Disparities (R01 MD011679, JHS and MDM), the National Institute of Nursing Research, NIH (R01 NR014855, LHA) and National Institute on Aging, NIH (R01 AG041099, MDM).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study was approved by the Children's Hospital of Philadelphia Institutional Review Board (19-016296).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. No data are available.

ORCID iDs
Karen B Lasater http://orcid.org/0000-0002-5834-1954
Matthew D McHugh http://orcid.org/0000-0002-1263-0697

REFERENCES

1 Richardson WBD, Bisgard J, *et al*. *Crossing the quality chasm: a new health system for the 21st century*. Washington, DC: I.o Medicine, 2001.
2 Brian DSA, Nelson AR. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. Washington, DC: I.o. Medicine, 2022.
3 Silber JH, Rosenbaum PR, Kelz RR, *et al*. Examining causes of racial disparities in general surgical mortality: Hospital quality versus patient risk. *Med Care* 2015;53:619:619–29.:.
4 Jha AK, Shlipak MG, Hosmer W, *et al*. Racial differences in mortality among men hospitalized in the Veterans Affairs health care system. *JAMA* 2001;285:297–303.
5 Neuman MD, Fleisher LA, Even-Shoshan O, *et al*. Nonoperative care for hip fracture in the elderly: the influence of race, income, and comorbidities. *Med Care* 2010;48:314:314–20.:.
6 Rangrass G, Ghaferi AA, Dimick JB. Explaining racial disparities in outcomes after cardiac surgery: the role of hospital quality. *JAMA Surg* 2014;149:223–7.
7 Groeneveld PW, Laufer SB, Garber AM. Technology diffusion, hospital variation, and racial disparities among elderly Medicare beneficiaries: 1989-2000. *Med Care* 2005;43:320–9.
8 Bailey ZD, Krieger N, Agénor M, *et al*. Structural racism and health inequities in the USA: evidence and interventions. *The Lancet* 2017;389:1453–63.
9 Asch DA, Islam MN, Sheils NE, *et al*. Patient and hospital factors associated with differences in mortality rates among black and white

us Medicare beneficiaries hospitalized with COVID-19 infection. *JAMA Netw Open* 2021;4:e2112842.
10 Hasnain-Wynia R, Baker DW, Nerenz D, *et al*. Disparities in health care are driven by where minority patients seek care: examination of the hospital quality alliance measures. *Arch Intern Med* 2007;167:1233–9.
11 Jha AK, Orav EJ, Li Z, *et al*. Concentration and quality of hospitals that care for elderly black patients. *Arch Intern Med* 2007;167:1177–82.
12 Dimick J, Ruhter J, Sarrazin MV, *et al*. Black patients more likely than whites to undergo surgery at low-quality hospitals in segregated regions. *Health Affairs* 2013;32:1046–53.
13 Lasater KB, Aiken LH, Sloane DM, *et al*. Chronic Hospital nurse understaffing meets COVID-19: an observational study. *BMJ Qual Saf* 2021;30:639–47.
14 Lake ET, Sanders J, Duan R, *et al*. A meta-analysis of the associations between the nurse work environment in hospitals and 4 sets of outcomes. *Med Care* 2019;57:353:353–61.:.
15 Lasater KB, McHugh M, Rosenbaum PR, *et al*. Valuing Hospital investments in nursing: multistate matched-cohort study of surgical patients. *BMJ Qual Saf* 2021;30:46–55.
16 Silber JH, Rosenbaum PR, McHugh MD, *et al*. Comparison of the value of nursing work environments in hospitals across different levels of patient risk. *JAMA Surg* 2016;151:527–36.
17 Kane RL, Shamliyan TA, Mueller C, *et al*. The association of registered nurse staffing levels and patient outcomes: systematic review and meta-analysis. *Med Care* 2007;45:1195–204.
18 Brooks Carthon JM, Kutney-Lee A, Jarrín O, *et al*. Nurse staffing and postsurgical outcomes in black adults. *J Am Geriatr Soc* 2012;60:1078–84. 10.1111/j.1532-5415.2012.03990.x Available: http://doi.wiley.com/10.1111/jgs.2012.60.issue-6
19 Lasater KB, McHugh MD. Reducing Hospital readmission disparities of older black and white adults after elective joint replacement: the role of nurse staffing. *J Am Geriatr Soc* 2016;64:2593–8.
20 Brooks-Carthon JM, Lasater KB, Rearden J, *et al*. Unmet nursing care linked to rehospitalizations among older black AMI patients: a cross-sectional study of US hospitals. *Med Care* 2016;54:457:457–65.:.
21 Brooks Carthon M, Brom H, McHugh M, *et al*. Better nurse staffing is associated with survival for black patients and diminishes racial disparities in survival after in-hospital cardiac arrests. *Med Care* 2021;59:169:169–76.:.
22 Dillman DA, Smyth JD, Christian LM. *Internet, phone, mail, and mixed-mode surveys: the tailored design method*. John Wiley & Sons, 2014.
23 Lasater KB, Jarrín OF, Aiken LH, *et al*. A methodology for studying organizational performance: a multistate survey of front-line providers. *Med Care* 2019;57:742:742–9.:.
24 Lake ET. Development of the practice environment scale of the nursing work index. *Res Nurs Health* 2002;25:176–88. 10.1002/nur.10032 Available: http://doi.wiley.com/10.1002/nur.v25:3
25 Rosenbaum PR. Coherence in observational studies. *Biometrics* 1994;50:368.
26 Lasater KB, McHugh MD, Rosenbaum PR, *et al*. Evaluating the costs and outcomes of hospital nursing resources: a matched cohort study of patients with common medical conditions. *J Gen Intern Med* 2021;36:84–91.
27 Daniel SR, Armstrong K, Silber JH, *et al*. An algorithm for optimal tapered matching, with application to disparities in survival. *J Comput Graph Stat* 2008;17:914–24.
28 Silber JH, Rosenbaum PR, Clark AS, *et al*. Characteristics associated with differences in survival among black and white women with breast cancer. *JAMA* 2013;310:389.
29 Silber JH, Rosenbaum PR, Ross RN, *et al*. Racial disparities in colon cancer survival: a matched cohort study. *Ann Intern Med* 2014;161:845–54.
30 Rosenbaum PR, Silber JH. Using the exterior match to compare two entwined matched control groups. *The American Statistician* 2013;67:67–75.
31 Gart JJ. An exact test for comparing matched proportions in crossover designs. *Biometrika* 1969;56:75–80.
32 Koch GG. 328. note: the use of none-parametric methods in the statistical analysis of the two-period change-over design. *Biometrics* 1972;28:577.
33 Silber JH, Rosenbaum PR, Romano PS, *et al*. Hospital teaching intensity, patient race, and surgical outcomes. *Arch Surg* 2009;144:113–20;
34 Lucas FL, Stukel TA, Morris AM, *et al*. Race and surgical mortality in the United States. *Ann Surg* 2006;243:281:281–6.:.
35 Deswal A, Petersen NJ, Souchek J, *et al*. Impact of race on health care utilization and outcomes in veterans with congestive heart failure. *J Am Coll Cardiol* 2004;43:778–84.
36 Volpp KG, Stone R, Lave JR, *et al*. Is thirty-day hospital mortality really lower for black veterans compared with white veterans? *Health Serv Res* 2007;42:1613–31.
37 Barnato AE, Lucas FL, Staiger D, *et al*. Hospital-Level racial disparities in acute myocardial infarction treatment and outcomes. *Med Care* 2005;43:308:308–19.:.
38 Mehtsun WT, Figueroa JF, Zheng J, *et al*. Racial disparities in surgical mortality: the gap appears to have narrowed. *Health Affairs* 2017;36:1057–64.
39 Kutney-Lee A, Stimpfel AW, Sloane DM, *et al*. Changes in patient and nurse outcomes associated with magnet Hospital recognition. *Med Care* 2015;53:550:550–7.:.
40 McHugh MD, Kelly LA, Smith HL, *et al*. Lower mortality in magnet hospitals. *Med Care* 2013;51:382:382–8.:.
41 McHugh MD, Brooks Carthon M, Sloane DM, *et al*. Impact of nurse staffing mandates on safety-net hospitals: lessons from California. *Milbank Q* 2012;90:160–86.
42 Aiken LH, Sloane DM, Cimiotti JP, *et al*. Implications of the California nurse staffing mandate for other states. *Health Serv Res* 2010;45:904–21.
43 Lasater KB, Aiken LH, Sloane D, *et al*. Patient outcomes and cost savings associated with Hospital safe nurse staffing legislation: an observational study. *BMJ Open* 2021;11:e052899.
44 Best MJ, McFarland EG, Thakkar SC, *et al*. Racial disparities in the use of surgical procedures in the US. *JAMA Surg* 2021;156:274–81.
45 Silber JH, Rosenbaum PR, Ross RN, *et al*. Disparities in breast cancer survival by socioeconomic status despite Medicare and Medicaid insurance. *Milbank Q* 2018;96:706–54.
46 Braveman P. Health disparities and health equity: concepts and measurement. *Annu Rev Public Health* 2006;27:167–94.