Title

The Impact of Patient-Physician Racial and Gender Concordance on Patient Satisfaction with Outpatient Clinic Visits.

Authors

Shaw, Nathan M
Hills, Nancy
Holler, Jordan
et al.

Publication Date

2023-06-01

DOI

10.1007/s40615-023-01676-5

Journal of Racial and Ethnic Health Disparities
https://doi.org/10.1007/s40615-023-01676-5

The Impact of Patient–Physician Racial and Gender Concordance on Patient Satisfaction with Outpatient Clinic Visits

Nathan M. Shaw[1,2] · Nancy Hills[3,4] · Jordan Holler[1,5] · Alicia Fernandez[6,7] · Denise Davis[6] · Nynikka R. Palmer[6,7] · Diane Sliwka[6] · Benjamin N. Breyer[1,3]

Received: 9 April 2023 / Revised: 10 May 2023 / Accepted: 5 June 2023

Abstract

Background Patient and provider race and gender concordance (patient and physician identify as the same race/ethnicity or gender) may impact patient experience and satisfaction.

Objective We sought to examine how patient and physician racial and gender concordance effect patient satisfaction with outpatient clinical encounters. Furthermore, we examined factors that changed satisfaction among concordant and discordant dyads.

Design Consumer Assessment of Healthcare Provider and Systems (CAHPS) Patient Satisfaction Survey Scores were collected from outpatient clinical encounters between January 2017 and January 2019 at the University of California, San Francisco.

Participants Patients who were seen in the eligible time period, who voluntarily provided physician satisfaction scores. Providers with fewer than 30 reviews and encounters with missing data were excluded.

Main Measures Primary outcome was rate of top satisfaction score. The provider score (1–10 scale) was dichotomized as "top score (9-10)" and "low scores (<9)."

Key Results A total of 77,543 evaluations met inclusion criteria. Most patients identified as White (73.5%) and female (55.4%) with a median age of 60 (*IQR* 45, 70). Compared to White patients, Asian patients were less likely to give a top score even when controlling for racial concordance (*OR*: 0.67; CI 0.63–0.714). Telehealth was associated with increased odds of a top score relative to in-person visits (*OR* 1.25; CI 1.07–1.48). The odds of a top score decreased by 11% in racially discordant dyads.

Conclusions Racial concordance, particularly among older, White, male patients, is a nonmodifiable predictor of patient satisfaction. Physicians of color are at a disadvantage, as they receive lower patient satisfaction scores, even in race concordant pairs, with Asian physicians seeing Asian patients receiving the lowest scores. Patient satisfaction data is likely an inappropriate means of determining physician incentives as such may perpetuate racial and gender disadvantages.

Keywords Patient satisfaction · Racial disparities · Gender

Introduction

The relationship between a patient and physician holds unique legal and ethical privileges and the quality of this relationship has been linked to health outcomes [1]. Studies

Diane Sliwka and Benjamin N. Breyer contributed equally to this work.

✉ Benjamin N. Breyer
Benjamin.Breyer@ucsf.edu

[1] Department of Urology, University of California San Francisco, San Francisco, CA, USA

[2] Department of Urology, MedStar Georgetown University Hospital, Washington, DC, USA

[3] Department of Epidemiology and Biostatistics, University of California San Francisco, 1001 Potrero Suite 3A, San Francisco, CA 94110, USA

[4] Department of Neurology, University of California San Francisco, San Francisco, CA, USA

[5] Harvard T.H. Chan School of Public Health, Boston, MA, USA

[6] Department of Medicine, University of California San Francisco, San Francisco, CA, USA

[7] Division on General Internal Medicine at Zuckerberg San Francisco General Hospital, University of California San Francisco, San Francisco, CA, USA

Published online: 21 June 2023

across multiple disciplines have sought to better understand fixed and modifiable factors that predict successful patient–physician relationships [2–5]. Specifically, there has been heightened focus on patient satisfaction as an objective assessment of the patient–physician relationship. Satisfaction scores are now routinely captured and available online for providers, administrators, and other patients. Despite limitations of online satisfaction scores including lower ratings for higher volume physicians [6], online ratings and reviews can be useful proxies to assess physician quality [7]. For example, physicians on probation are more likely to have lower online ratings [8]. Patient satisfaction has become a common metric to compare and evaluate physicians and is predicted to play an increasing role in physician incentives [9, 10]. In this environment, it is important to understand factors that may affect patient satisfaction with their physician.

Over 20 years ago, two studies suggested racial concordance between the patient and provider may impact patient satisfaction; and that satisfaction can be used as a metric to measure partnership in the physician–patient relationship [5, 11]. Since then, there have been conflicting studies on the impact of patient/physician race/ethnicity and satisfaction. Some studies demonstrate higher satisfaction [12–14] and better patient experience [15–17], and even improved clinical/surgical outcomes [18] based on patient and/or physician race and racial concordance. Despite these data, a review of 27 studies in 2009 failed to show conclusive evidence that racial concordance improved outcomes [19]. Meanwhile, similar data over the same period has demonstrated variation in patient satisfaction based on provider gender including patient and provider gender concordance (patient and provider are same gender) [20]. Specifically, female patients frequently rate their physicians lower across multiple domains [21–23]. In light of the conflicting data on race, gender, and patient satisfaction, Takeshita et al. (2020) in a large cross sectional study with a diverse patient and provider population demonstrated that racial concordance between physician and patient is associated with higher satisfaction scores [3]. The same study found that patient and provider gender concordance was not associated with patient satisfaction. While these data were derived from a diverse, high-volume academic center, they have yet to be replicated. Nor is it clear if there are other patient and physician factors that could overcome the bias of discordant dyads.

This study has two primary aims. First, we sought to examine the relationship between patient–physician concordance (race/ethnicity and gender) and patient satisfaction in a highly ethnically diverse patient population. We hypothesized that racially concordant dyads will be associated with higher patient satisfaction scores and patient male gender would be associated with higher scores, particularly in racially and gender concordant dyads. Second, we sought to identify additional factors which predict patient satisfaction. We hypothesized that younger patients and those seen in telehealth may provide higher satisfaction scores even among gender and/or racially discordant dyads.

Methods

Data Source

We utilized the Consumer Assessment of Healthcare Provider and Systems (CAHPS) Patient Satisfaction Survey Scores from the University of California, San Francisco (UCSF) Health outpatient clinics from January 2017 to January 2019.

Study Population

All available patient satisfaction scores during the study period were included across specialties in the UCSF health system. Patients can voluntarily evaluate any provider they have seen for clinical care. Demographic data is derived from the clinical encounter being evaluated. In order to create unique patient–physician dyads, only the first evaluation was included when a patient had provided multiple evaluations of the same provider. Patients who evaluated more than one physician were included. Only providers who had a minimum of 30 reviews were included to decrease variability associated with smaller samples.

Variables

Patient demographics were captured including age, gender, and self-identified race/ethnicity (White, Black/African American, Hispanic/Latinx, Asian), and primary language. Primary language for patients was based on medical record database. Physician age, gender, languages spoken, and race/ethnicity (White, Black/African American, Hispanic/Latinx, Asian) was assigned using publicly available data and reviewer consensus. A total of 4 reviewers (JH, BN, NH, MS) independently determined gender and race based on online provider profiles (https://www.ucsfhealth.org/find-a-doctor). The determination of provider race/ethnicity required consensus of 3 of 4 reviewers based on previously described methods [3].

The clinic location was determined by the location where the score generating encounter occurred. These were used to assign each physician into one of the following categories as aggregates for medical specialty: "Surgery" (general surgery and all surgical subspecialties); "Adult Primary Care" (internal medicine, adult family medicine); "Medical Subspecialties" (all adult medical specialties not categorized as surgery or "Adult Primary Care"); "Obstetrics/Gynecology"; and "Pediatrics" (including general pediatrics, pediatric surgery and medical specialties). Specialties

were grouped to ensure physician anonymity, particularly in small departments with limited providers.

Based on the significant skew of the patient satisfaction data (Fig. 1), scores were dichotomized as high (a rating of 9–10 out of 10 total) and low (a rating from 0 to 8 out of 10 total). This is similar to prior analyses done on a standard Likert scale where 5 and <5 (out of 5) were dichotomized [3, 20].

Statistical Analysis

The study population was then divided into dyads based on race and gender. The entire cohort was analyzed with racial or gender concordance as a variable. Furthermore, a subgroup analysis of racially or gender concordant (physician and patient identify as the same race/ethnicity or gender) dyads and racially discordant (physician and patient identify as different race/ethnicity or gender) was performed.

Descriptive statistics were used to characterize the study cohort. Continuous variables were reported as mean and standard deviation (SD). Univariate analysis was used to report the odds of patient and provider factors that were associated with "top score" among the entire cohort. An additional univariate analysis was carried out among racially and gender concordant dyads that received "low score" and racially and gender discordant dyads that received "top score." Predictors of "top score" were compared using chi-square or Fisher's exact test for categorical variables or Mann–Whitney *U* test for continuous variables. Finally, logistic regression was performed for all factors that predicted "top score." Statistical analysis was performed using Stata 17 (Stata Corp, College Station, TX, USA) and a 2-sided p value <0.05 considered statistically significant.

Finally, a Monte-Carlo simulation model was created of an identical patient pool of 1500 patients reflective of the gender, racial, and age breakdown of the San Francisco Bay Area based on 2020 census data with the following breakdown — White 55%, Black 5%, Asian 25%, and Hispanic, 15% with 50% male and 18% over the age of 65 [24]. The model was simulated on Microsoft Excel (Microsoft Corporation, Seattle WA). Model 1 assumed all in-person visits. The model utilized the probability of physician top rating described in Table 1 with modifiers of top score found in Table 3. Two additional models were run to include the same population assuming a provider saw 75% telehealth (model 2) and 75% men over age 60 (model 3).

Results

A total of 58,528 unique patients, 604 unique physicians, and 77,543 dyads met inclusion criteria for analysis. The majority of patients identified as White (73.5%) and female (55.4%) with a median age of 60 (Appendix Table 4). Physicians were nearly balanced by gender, male (51.3%), and majority White (64.1%) and had a median of 143 evaluations (Appendix Table 4).

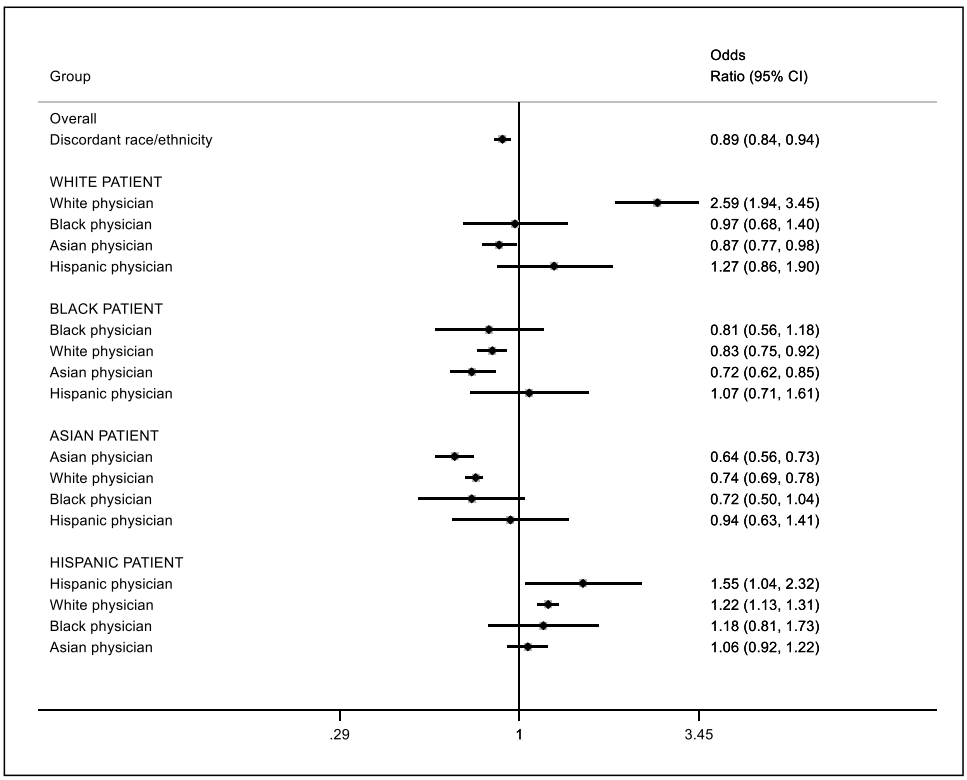

Fig. 1 Odds of receiving "top score" by patient–physician dyad. Patient satisfaction of 9 or 10 out of 10 was defined as "top score"

Table 1 Distribution of satisfaction scores by patient–physician racial/ethnic pairs*

Physician race/ethnicity and Press-Ganey score	Patient race (evaluations)								*p*-value
	White (*n*=57,397)		Black (*n*=2824)		Hispanic (*n*=6841)		Asian (*n*=10,481)		
	n	(%)	*n*	(%)	*n*	(%)	*n*	(%)	
Any race/ethnicity	57,397		2824		6841		10481		<0.0001
Topscore [9, 10]	49,765	(86.7)	2344	(83.0)	5927	(86.6)	8418	(80.3)	
Lowscore[<9]	7,632	(13.3)	480	(17.0)	914	(13.4)	2063	(19.7)	
White	36,711		1689		4240		6013		<0.0001
Topscore [9, 10]	31,992	(87.1)	1405	(83.2)	3692	(87.1)	4837	(80.4)	
Lowscore[<9]	4,719	(12.9)	284	(16.8)	548	(12.9)	1176	(19.6)	
Black	1,061		113		155		195		0.32
Topscore [9, 10]	899	(84.7)	98	(86.7)	132	(85.2)	156	(80.0)	
Lowscore[<9]	162	(15.3)	15	(13.3)	23	(14.8)	39	(20.0)	
Hispanic	769		41		185		116		0.19†
Topscore [9, 10]	677	(88.0)	37	(90.2)	162	(87.6)	94	(81.0)	
Lowscore[<9]	92	(12.0)	4	(9.8)	23	(12.4)	22	(19.0)	
Asian	18,856		981		2261		4157		<0.0001
Topscore [9, 10]	16,197	(85.9)	804	(82.0)	1941	(85.8)	3331	(80.1)	
Lowscore[<9]	2659	(14.1)	177	(18.0)	320	(14.2)	826	(19.9)	

*First score per unique pair. **p-values calculated using chi-square tests unless otherwise indicated. †p-value calculated using Fisher's exact test

Racial and Gender Concordance

All patient–physician dyads had a high density of satisfaction scores of 8 or greater on a 10-point scale (Appendix Fig. 3). There was a significant difference in physician satisfaction ratings across patient races, with White (86.7%) and Hispanic (86.6%) patients more likely to give top ratings compared to Black (83.0%) and Asian (80.3%) patients. When controlling for racial concordance, this difference remained significant for White and Hispanic physicians, but not among Black or Asian physicians (Table 1).

In univariate analyses, compared to their White colleagues, Asian physicians received 12.6% fewer top scores (odds ratio (OR) 0.87, 95% confidence interval (CI) 0.78, 0.98); p=0.02). Black and Asian patients were significantly less likely to give physicians a top score (adjusted OR (AOR) 0.80 95% CI 0.72, 0.89; p<0.001; *AOR* 0.67, 95% CI 0.63. 0.71; p=<0.001, respectively) (Table 3). The odds of receiving a top score, adjusted for physician and patient race/ethnicity, physician and patient age, and physician and patient gender, were highest between a White patient and a White physician (*OR* 2.59 95% CI 1.94, 3.45). The lowest odds of receiving a top score were between Asian patient and Asian physician (*OR* 0.64 95% CI 0.56, 0.73) (Fig. 1).

Factors other than race also demonstrated significant association with a top score in univariate analysis. Increasing patient age per decade demonstrated an increased odds of top score (*OR* 1.23; 95% CI 1.21, 1.25; p<0.001). Similarly, patient and physician male gender were each associated with increased odds of receiving top score (*OR* 1.36; 95% CI 1.30, 1.43; p<0.001) and (*OR* 1.14; 95% CI 1.02, 1.26; p=0.02) respectively. Relative to primary care, only obstetrics and gynecology was associated decreased odds of receiving a top score (*OR* 0.57; 95% CI 0.46, 0.71; p<0.001). Finally, telehealth visits (relative to in-person) were associated with increased odds of receiving a top score (*OR* 1.25; 95% CI 1.07, 1.48; p<0.001) (Table 2).

Univariate Predictors of Patient Satisfaction in Racially Concordant and Discordant Pairs

When the analysis was repeated by examining the cohort by racially concordant and discordant dyads, patient age and gender remained predictors of top score. Older patients (*OR* 1.20 per decade, 95% CI 1.17, 1.22; p<0.001) and male patients (*OR* 1.30; 95% CI 1.22, 1.38; p<0.001) were more likely to give physicians a top score among discordant dyads (Table 3). Similarly, physicians seeing Asian patients were more likely to receive low scores compared to those seeing white patients. Female physicians and physicians in obstetrics and gynecology were more likely to receive low scores relative to male colleagues and those in primary care respectively (Table 3).

Univariate Predictors of Patient Satisfaction Based on Gender Concordant and Discordant Pairs

When dyads were restructured as gender concordant/discordant, many of the same factors that were found to be

Table 2 Effect of patient and physician factors on physician top score

Characteristic	Provider top score Univariate analysis		
	OR; AOR**	(95% CI)	*p**
Racial discordance	0.89	(0.84, 0.94)	<0.001
Patients			
Age, decades	1.23	(1.21, 1.25)	<0.001
Gender male	1.36	(1.30, 1.43)	<0.001
Race			
White	Ref		
Black	0.77; 0.80**	(0.69, 0.84); (0.72–0.88)**	<0.001; <0.001**
Hispanic	1.03; 1.07**	(0.95, 1.11); (0.99–1.16)**	0.48; 0.073**
Asian	0.66; 0.67**	(0.62, 0.70); (0.63–0.714)**	<0.001; <0.001**
Physicians			
Age, decades	0.998	(0.95, 1.05)	0.93
Gender male	1.14	(1.02, 1.26)	0.02
Race			
White	Ref		
Black	0.87; 0.94**	(0.61, 1.24); (0.65–1.34)**	0.45; 0.74**
Hispanic	1.02; 1.09**	(0.75, 1.38); (0.80–1.50)**	0.90; 0.57**
Asian	0.87; 0.94**	(0.78, 0.98); (0.834–1.07)**	0.02; 0.38**
Department			
Primary care	Ref		
Medical subspecialty	1.15	(0.99, 1.33)	0.07
Surgery	1.02	(0.86, 1.20)	0.85
Obstetrics and gynecology	0.57	(0.46, 0.71)	<0.001
Pediatrics	0.87	(0.67, 1.14)	0.31
Visits			
Visit type			
In-person	Ref		
Telehealth	1.25	(1.07, 1.48)	0.006
Visit year			
2017	Ref		
2018	1.03	(0.98, 1.08)	0.24
2019	1.13	(1.07, 1.19)	<0.001

*p-values were calculated using univariate logistic regression models. **AOR, adjusted odds ratio. Adjusted for racial discordance

associated with high and low rating among racial dyads remained so (Table 3). Physician age, however, differed. Physician age in decades was associated with a decreased odds of receiving a top score (OR 0.92 95% CI 0.88, 0.97; p=0.002) in gender discordant dyads despite being more likely to received one in racially concordant dyads.

Simulation Model

In a simulated model of 1500 patients, there were differences in satisfaction scores based on physician race alone (Fig. 2). Black (82%) and Asian (84.7%) providers received fewer top scores compared to White (85.7%) and Hispanic/LatinX (86.6%) colleagues. Differences in satisfaction score by physician race increased in models with substantial telehealth (Fig. 2, model 2) or with predominance of older male patients (Fig. 2, model 3).

Discussion

Our study found that patient and physician race/ethnicity are significant predictors of patient satisfaction scores across a large multispecialty cohort. Older, White male patients seeing White physicians, particularly in telehealth, were among the most satisfied while providers seeing younger, Asian women were much less likely to receive top satisfaction scores. We also offer novel findings that older male patients and those seen in telehealth

Table 3 Factors that predict high and low ratings in racially and gender concordant and discordant dyads

Characteristic	Racially concordant pairs			Racially discordant pairs			Gender concordant pairs			Gender discordant pairs		
	Odds of receiving low* score			Odds of receiving high** score			Odds of receiving low* score			Odds of receiving high** score		
	OR	(95% CI)	*p*-value	*OR*	(95% CI)	*p*-value	*OR*	(95% CI)	*p*-value	*OR*	(95% CI)	*p*-value
Patients												
Age, decades	0.82	(0.80, 0.84)	<0.001	1.2	(1.17, 1.22)	<0.001	0.82	(0.80, 0.84)	<0.001	1.21	(1.18, 1.23)	<0.001
Gender male	0.71	(0.66, 0.75)	<0.001	1.3	(1.22, 1.38)	<0.001	-			-		
Race/ethnicity												
White	-			-			Ref			Ref		
Black	-			-			1.26	(1.10, 1.44)	<0.001	0.76	(0.65, 0.88)	<0.001
Hispanic	-			-			1.01	(0.92, 1.11)	0.8	1.1	(0.97, 1.25)	0.12
Asian	-			-			1.5	(1.38, 1.62)	<0.001	0.66	(0.60, 0.72)	<0.001
Age, decades	0.99	(0.94, 1.04)	0.75	0.9	(0.93, 1.03)	0.43	0.94	(0.89, 0.99)	0.01	0.92	(0.88 0.97)	0.002
Gender (male)	0.86	(0.76, 0.97)	0.02	1.1	(0.97, 1.23)	0.12	-			-		
Race/ethnicity												
White	-			-			Ref			Ref		
Black	-			-			1.32	(0.90, 1.92)	0.15	1.2	(0.73, 1.98)	0.47
Hispanic	-			-			0.8	(0.60, 1.07)	0.13	0.92	(0.65, 1.30)	0.65
Asian	-			-			1.22	(1.07, 1.37)	0.002	0.95	(0.83, 1.09)	0.51
Department												
Primary care	Ref			Ref			Ref			Ref		
Medical subspecialty	0.87	(0.73, 1.03)	0.11	1	(0.92, 1.27)	0.33	0.82	(0.70, 0.97)	0.02	1.1	(0.92, 1.32)	0.3
Surgery	1.17	(0.97, 1.41)	0.1	1.1	(0.93, 1.36)	0.19	0.93	(0.78, 1.12)	0.46	0.97	(0.79, 1.20)	0.79
Obstetrics and gynecology	1.91	(1.48, 2.46)	<0.001	0.6	(0.51, 0.82)	<0.001	1.63	(1.29, 2.07)	<0.001	0.51	(0.35, 0.74)	<0.001
Pediatrics	1.19	(0.88, 1.61)	0.25	0.9	(0.71, 1.25)	0.68	1	(0.72, 1.38)	1	0.79	(0.60, 1.04)	0.1
Visits												
Visit type												
In-person	Ref			Ref			Ref			Ref		
Telehealth	0.79	(0.64, 0.98)	0.04	1.3	(1.04, 1.62)	0.02	0.79	(0.65, 0.97)	0.025	1.23	(0.97, 1.57)	0.09
Visit year												
2017	Ref			Ref			Ref			Ref		
2018	1	(0.94, 1.07)	0.98	1	(0.99, 1.13)	0.09	0.99	(0.93, 1.06)	0.88	1.06	(0.98, 1.14)	0.14
2019	0.91	(0.85, 0.97)	0.005	1.1	(1.07, 1.24)	<0.001	0.87	(0.81, 0.94)	<0.001	1.09	(1.01, 1.18)	0.02

*Low=0-8/10; high=9-10/10. Racially and gender concordant pairs where the patient satisfaction was <9/10 were defined as "Concordant Pairs with Low Score." Racially and gender discordant pairs where the patient satisfaction was 9 or 10 out of 10 were defined as "Discordant Pairs with High Score." This is a univariate analysis, stratified by race/ethnicity and gender concordance/discordance. *Odds ratio indicates odds of low score in a concordant pair with the described variable. Value <1 indicates lower likelihood of receiving a low score. **Odds ratio indicates odds of high score in a discordant pair with the described variable. Value <1 indicates lower likelihood of receiving a high score

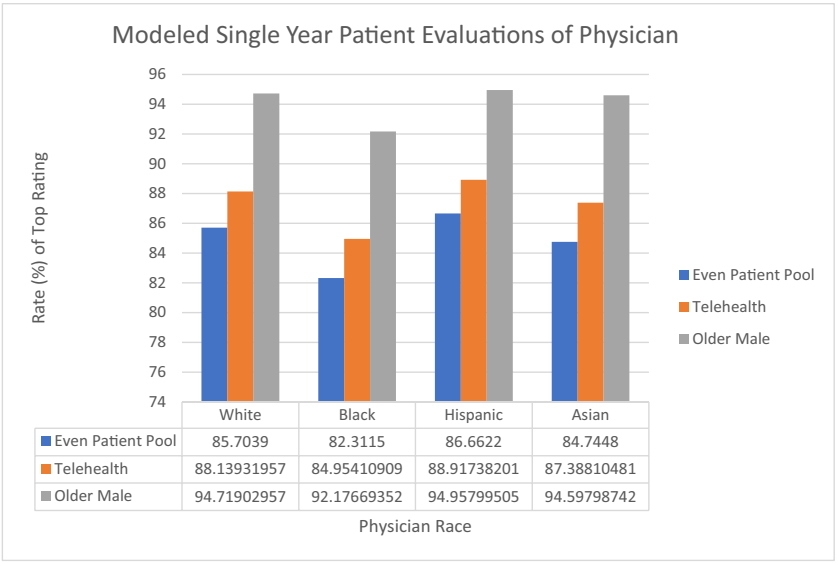

Fig. 2 Simulation of physician satisfaction score. Data presented as a percentage of ratings as "top score." In model 1 "Even Patient Pool," the rate of physician top score for each physician is expressed based on a mixed patient pool and odds of top score. In model 2 "Telehealth," the model assumes that 75% of patients are seen in telehealth. All other variables are kept constant. No adjustment was done for access to care. In model 3 "Older Male," the model assumes that 75% of patients are men in their 60s. All other variables are kept constant

were more likely to provide higher score regardless of provider race and racial concordance. Despite that, racial discordance alone resulted in an 11% decrease in the odds that visit will result in the provider receiving a high score in this study.

Interestingly, dyads in obstetrics and gynecology had significantly lower odds of a top score (*OR* 0.57). These data are in contrast with existing data which show that obstetrics/gynecology receive higher patient satisfaction across multiple metrics compared to other specialties [25]. No other specialty category resulted in a significant effect on dyad rating. These findings are likely linked to lower ratings offered by women across all specialties.

Even prior to the dramatic increase in telehealth during the COVID-19 pandemic, data presented here suggest that telehealth is associated with increased patient satisfaction. Studies both before and after the COVID-19 pandemic have found that telehealth is associated with improved patient satisfaction, fewer missed appointments, and lower cost [26, 27]. These benefits, however, are likely not shared equally by all patients. Patients who identify as Black or Hispanic are more likely to come in-person and have poor access to video visits [28, 29]. Furthermore, age, insurance status, and health literacy have all been shown to decrease the likelihood of video visit capability [30]. We were not able to determine if the telehealth used in our study was audio-only visits or video-visits. We are therefore unable to comment on satisfaction or utilization of telephone versus video; this distinction is particularly important in the era of reimbursement being substantively different between those visit types.

The fixed or non-modifiable factors in the patient–physician dyads examined in this study are also influenced by an ever-growing number of interventions to address patient satisfaction. These modifiable factors include interventions directed toward the provider [31, 32] (humility, cultural competency, and communication skills training) and healthcare delivery systems (patient focused therapy, improved scheduling application, and standardization of post-operative protocols) [33]. Trainings aligned toward providers including cultural humility, anti-racism, and anti-bias trainings are increasing in frequency, but there is a paucity of data on the effect of such trainings on patient outcomes (including satisfaction) [34].

This study, like others, found that patients who identify as Asian rated providers lower compared to White patients [35, 36]. Past data suggest that this may be a measure of response tendencies and not necessarily lower satisfaction with care received [37]. Despite this, a recent systematic review found that higher levels of cultural humility led to increased satisfaction among hospitalized Asian patients [38]. More generally, an analysis of 19 review articles found a modest impact of cultural humility on patient satisfaction, healthcare utilization, and patient outcome [39]. This underscores that interventions can be tailored to improve satisfaction and an interpretation of the data as "some patients are just less satisfied or unsatisfiable" is inaccurate.

The differences in provider satisfaction score by provider race in the modeled patient population emphasized the impact on a provider. While a difference in top score of a 2–3% may seem modest, thresholds to achieve resources (e.g., >85% top score) heighten the stakes for even small absolute differences. These thresholds already exist across major healthcare systems and within departments to assignment of crucial healthcare resources. Additionally, this effect is magnified by patient population seen. In our model, a Black physician who sees a mixed patient population (model 1) will have 12.4% fewer top scores compared to a White colleague who sees older men (model 3). This difference can be further complicated by specialty, access to telehealth, and other factors. This should serve as a warning to the notion of linking physician reimbursement or other incentives to patient satisfaction. If

satisfaction data is taken out of context or not controlled, multiple non-modifiable factors can disadvantage physicians — particularly physicians of color that see women. Thus, even as efforts are made to diversify the workforce, "objective" data on satisfaction may tell a very misleading tale on the quality of the patient–physician relationship and the physician.

Limitations

Our study has limitations. First, the significant skew of the ratings clustered most scores around 9 or 10 out of a total of 10. This may be related to selection bias, as patients who were pleased with their care may be motivated to review their provider. That said, patients displeased with their care are more likely to respond to surveys [40]. The choice was made to treat 9 and 10 out of 10 as a "top score" to address this limitation. As this metric was used across the entire study population, any impact would be non-differential across variables. As these data are derived from a single institution, findings may not be generalizable. This is made more complex by the aggregates of medical specialty and the small number of dyads in certain categories.

There are also methodical limitations inherent to the dataset including missing data; these were excluded. Scores beyond the first patient satisfaction score between a patient–physician dyad were excluded. This was done to reduce the bias of a patient who liked a physician and continued to see and rate that physician. Averaging the dyad score or collecting top score would not address this bias. The assigned specialty of physicians was determined based on the clinic where that provider saw the patient; it is therefore possible that clinician specialty was inappropriately categorized. To reduce the potential for mis-categorization and to ensure physician anonymity (particularly in small departments), the aggregate categories were created. This categorization has not previously been validated but aligns with clinical practice. Similarly, given the wide variety of specialties incorporated, it is possible that a greater effect could have been seen within a category (e.g., nephrology instead of the whole category "Medical Subspecialties"). While using reviewers to determine physician age, race/ethnicity, and gender has been described, it may be a poor proxy for physician self-identity and may have misrepresenting physician age.

Positionality Statement

Science and research are situated in a social context, and scientific inquiry is affected by the perspectives and biases of the individuals engaged in scholarly work. This research team was assembled intentionally to represent expertise in health disparities, healthcare communication, and patient experience; racially/ethnically diverse backgrounds; and multiple career stages and professional backgrounds. Four of the nine authors are people of color and two-thirds are women. However, we acknowledge identifies and perspectives are missing from this team; this may introduce biases that are challenging to measure and/or address.

Conclusion

Racial concordance remains a strong predictor of patient satisfaction. This disadvantages physicians of color and Asian physicians, particularly those that see a majority female and/or Asian population of patients as those patients score physicians lower across the board. While some modifiable factors (e.g., telehealth visits) appear to diminish the negative impact of racial discordance on patient satisfaction, race/ethnicity and gender continue to play a significant role in patient experience. Patient satisfaction measures that do not account for patient mix by gender and race/ethnicity are not appropriate as a means of determining physician incentives as they will likely perpetuate racial and gender disadvantages.

Appendix

Table 4 Patient and physician characteristics

	Unique patients (n=58,528)		Unique physicians (n=604)	
	n	(%)	n	(%)
Age in years				
Mean (SD)	56	(18.5)	50.5	(11.6)
Median (IQR)	60	(45, 70)	49	(41, 59)
Gender				
Male	26067	(44.5)	310	(51.3)
Female	32431	(55.4)	294	(48.7)
Missing	30	(0.1)		
Race/ethnicity				
White	43033	(73.5)	387	(64.1)
Black	2157	(3.7)	15	(2.5)
Hispanic	5383	(9.2)	14	(2.3)
Asian	7955	(13.6)	188	(31.1)
Preferred language				
English	57099	(97.6)		
Other	1429	(2.4)		
Number of surveys	[per patient]		[per physician]	
Median (IQR)	1	(1, 2)	143	(80, 245)
Range		(1, 34)		(35, 1175)

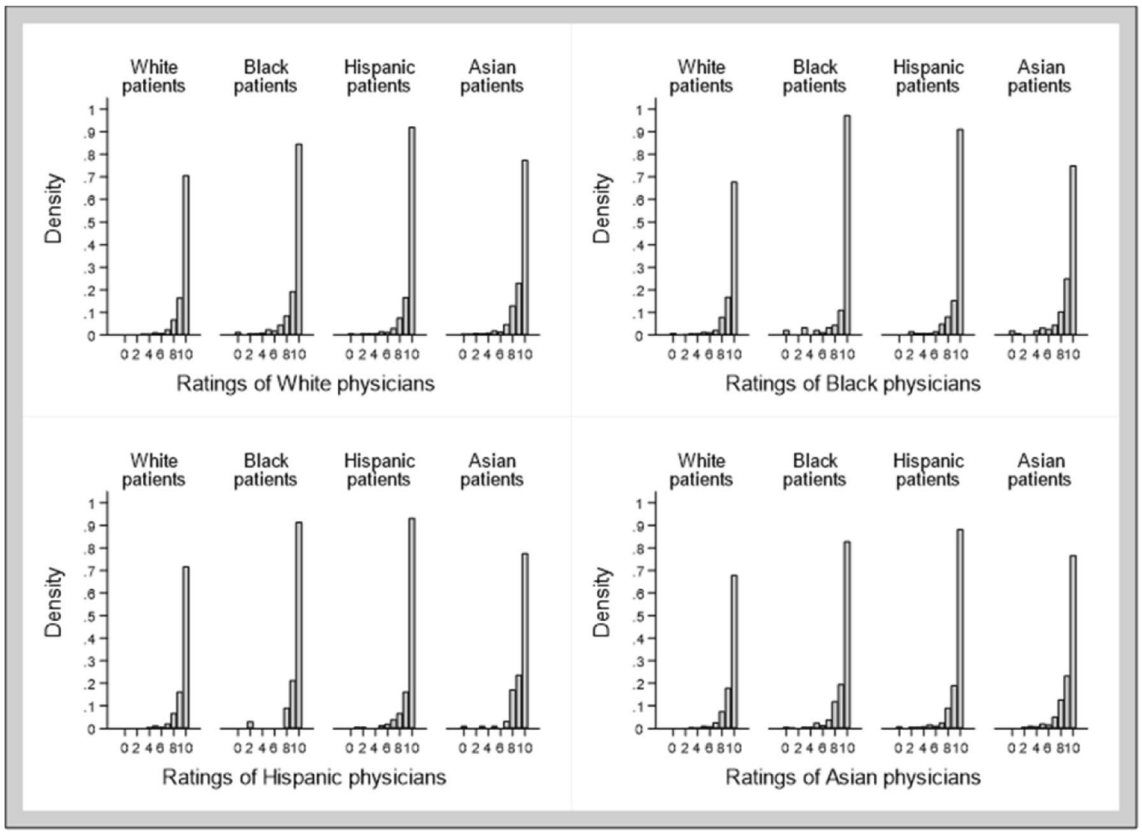

Fig. 3 Ratings by patient/physician racial pairing. Density of patient satisfaction scores (0–10) based on patient and physician race/ethnicity

Acknowledgements We gratefully acknowledge the assistance of Nizar Hakam, Behnam Nabavizadeh, Natalie Rios, and Michael Sadighian particularly for their roles in reviewing provider profiles. All represent the UCSF Department of Urology and were otherwise uncompensated for their work.

Author Contributions NS: data analysis/interpretation and manuscript drafting and editing. NH: data analysis/interpretation and manuscript editing. JH: project design and data collection and analysis. AF: project design, data analysis, and manuscript editing. DD: project design and manuscript editing. NP: project design and manuscript editing. DS: project design, data analysis, manuscript editing, and supervision. BB: project design, data analysis, manuscript editing, and supervision.

Data Availability Data will not be made available.

Code Availability Code can be made available without associated data.

Declarations

Ethics Approval Institutional Review Board (IRB) exemption was granted for this project. Data was primarily de-identified and/or otherwise publicly available.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Conflict of Interest The authors declare no competing interests.

References

1. Olaisen RH, Schluchter MD, Flocke SA, Smyth KA, Koroukian SM, Stange KC. Assessing the longitudinal impact of physician-patient relationship on functional health. Ann Fam Med. 2020;18(5):422–9. https://doi.org/10.1370/AFM.2554.
2. Martinez KA, Keenan K, Rastogi R, et al. The association between physician race/ethnicity and patient satisfaction: an exploration in direct to consumer telemedicine. J Gen Intern Med. 2020;35(9):2600–6. https://doi.org/10.1007/S11606-020-06005-8.
3. Takeshita J, Wang S, Loren AW, et al. Association of racial/ethnic and gender concordance between patients and physicians with patient experience ratings. JAMA Netw Open. 2020;3(11):e2024583. https://doi.org/10.1001/jamanetworkopen.2020.24583.
4. Jackson JL, Chamberlin J, Kroenke K. Predictors of patient satisfaction. Soc Sci Med. 2001;52(4):609–20. https://doi.org/10.1016/S0277-9536(00)00164-7.
5. Cooper-Patrick L, Gallo JJ, Gonzales JJ, et al. Race, gender, and partnership in the patient-physician relationship. JAMA. 1999;282(6):583–9. https://doi.org/10.1001/JAMA.282.6.583.
6. Murphy GP, Awad MA, Tresh A, et al. Association of patient volume with online ratings of California urologists. JAMA Surg. 2018;153(7):685–6. https://doi.org/10.1001/JAMASURG.2018.0149.
7. Murphy GP, Radadia KD, Breyer BN. Online physician reviews: is there a place for them? Risk Manag Healthc Policy. 2019;12:85–9. https://doi.org/10.2147/RMHP.S170381.
8. Murphy GP, Awad MA, Osterberg EC, et al. Web-based physician ratings for California physicians on probation. J Med Internet Res. 2017;19(8). https://doi.org/10.2196/JMIR.7488.

9. Petrullo KA, Lamar S, Nwankwo-Otti O, Alexander-Mills K, Viola D. The patient satisfaction survey: what does it mean to your bottom line? J Hosp Adm. 2013;2(2). https://doi.org/10.5430/jha.v2n2p1.
10. Siegrist RB. Patient satisfaction: history, myths, and misperceptions. Virtual Mentor. 2013;15(11):982–7. https://doi.org/10.1001/VIRTUALMENTOR.2013.15.11.MHST1-1311.
11. Saha S, Komaromy M, Koepsell TD, Bindman AB. Patient-physician racial concordance and the perceived quality and use of health care. Arch Intern Med. 1999;159(9):997–1004. https://doi.org/10.1001/ARCHINTE.159.9.997.
12. Haviland MG, Morales LS, Reise SP, Hays RD. Do health care ratings differ by race or ethnicity? Jt Comm J Qual Saf. 2003;29(3):134–45. https://doi.org/10.1016/s1549-3741(03)29016-x.
13. Murray-García JL, Selby JV, Schmittdiel J, Grumbach K, Quesenberry CPJ. Racial and ethnic differences in a patient survey: patients' values, ratings, and reports regarding physician primary care performance in a large health maintenance organization. Med Care. 2000;38(3):300–10. https://doi.org/10.1097/00005650-200003000-00007.
14. Kumar D, Schlundt DG, Wallston KA. Patient-physician race concordance and its relationship to perceived health outcomes. Ethn Dis. 2009;19(3):345–51.
15. LaVeist TA, Nuru-Jeter A, Jones KE. The association of doctor-patient race concordance with health services utilization. J Public Health Policy. 2003;24(3-4):312–23. https://doi.org/10.2307/3343378.
16. Laveist TA, Nuru-Jeter A. Is doctor-patient race concordance associated with greater satisfaction with care? J Health Soc Behav. 2002;43(3):296–306.
17. Taira DA, Safran DG, Seto TB, et al. Do patient assessments of primary care differ by patient ethnicity? Health Serv Res. 2001;36(6 Pt 1):1059–71.
18. Wallis CJD, Jerath A, Coburn N, et al. Association of surgeon-patient sex concordance with postoperative outcomes. JAMA Surg. 2021. https://doi.org/10.1001/JAMASURG.2021.6339.
19. Meghani SH, Brooks JM, Gipson-Jones T, Waite R, Whitfield-Harris L, Deatrick JA. Patient-provider race-concordance: does it matter in improving minority patients' health outcomes? Ethn Health. 2009;14(1):107–30. https://doi.org/10.1080/13557850802227031.
20. Rogo-Gupta LJ, Haunschild C, Altamirano J, Maldonado YA, Fassiotto M. Physician gender is associated with Press Ganey patient satisfaction scores in outpatient gynecology. Women's Heal issues Off Publ Jacobs Inst Women's Heal. 2018;28(3):281–5. https://doi.org/10.1016/j.whi.2018.01.001.
21. Woods SE, Heidari Z. The influence of gender on patient satisfaction. J gender-specific Med JGSM Off J Partnersh Women's Heal Columbia. 2003;6(4):30–5.
22. Okunrintemi V, Valero-Elizondo J, Patrick B, et al. Gender differences in patient-reported outcomes among adults with atherosclerotic cardiovascular disease. J Am Heart Assoc. 2018;7(24):e010498. https://doi.org/10.1161/JAHA.118.010498.
23. Alsaif A, Alqahtani S, Alanazi F, Alrashed F, Almutairi A. Patient satisfaction and experience with anesthesia: a multicenter survey in Saudi population. Saudi J Anaesth. 2018;12(2):304–10. https://doi.org/10.4103/sja.SJA_656_17.
24. U.S. Census Bureau QuickFacts: San Francisco County, California. https://www.census.gov/quickfacts/fact/table/sanfranciscocountycalifornia/POP010220. Accessed November 14, 2022.
25. Patel I, Chang J, Srivastava J, Feldman S, Levender M, Balkrishnan R. Patient satisfaction with obstetricians and gynecologists compared with other specialties: analysis of US self-reported survey data. Patient Relat Outcome Meas. 2011;2:21–6. https://doi.org/10.2147/PROM.S15747.
26. Kruse CS, Krowski N, Rodriguez B, Tran L, Vela J, Brooks M. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open. 2017;7(8):e016242. https://doi.org/10.1136/BMJOPEN-2017-016242.
27. Ramaswamy A, Yu M, Drangsholt S, et al. Patient satisfaction with telemedicine during the COVID-19 pandemic: retrospective cohort study. J Med Internet Res. 2020;22(9):e20786. https://doi.org/10.2196/20786.
28. Mueller BR, Lawrence S, Benn E, et al. Disparities in telehealth utilization in patients with pain during COVID-19. Pain Reports. 2022;7(3):E1001. https://doi.org/10.1097/PR9.0000000000001001.
29. Campos-Castillo C, Anthony D. Racial and ethnic differences in self-reported telehealth use during the COVID-19 pandemic: a secondary analysis of a US survey of internet users from late March. J Am Med Inform Assoc. 2021;28(1):119. https://doi.org/10.1093/JAMIA/OCAA221.
30. Brown SH, Griffith ML, Kripalani S, Horst SN. Association of health literacy and area deprivation with initiation and completion of telehealth visits in adult medicine clinics across a large health care system. JAMA Netw Open. 2022;5(7):e2223571–1. https://doi.org/10.1001/JAMANETWORKOPEN.2022.23571.
31. Chang ES, Simon M, Dong XQ. Integrating cultural humility into health care professional education and training. Adv Health Sci Educ Theory Pract. 2012;17(2):269–78. https://doi.org/10.1007/S10459-010-9264-1.
32. Tang C, Tian B, Zhang X, et al. The influence of cultural competence of nurses on patient satisfaction and the mediating effect of patient trust. J Adv Nurs. 2019;75(4):749–59. https://doi.org/10.1111/JAN.13854.
33. Davidson KW, Shaffer J, Ye S, et al. Interventions to improve hospital patient satisfaction with healthcare providers and systems: A systematic review. BMJ Qual Saf. 2017;26(7):596. https://doi.org/10.1136/BMJQS-2015-004758.
34. Argueza BR, Saenz SR, McBride D. From diversity and inclusion to antiracism in medical training institutions. Acad Med. 2021;96(6):798–801. https://doi.org/10.1097/ACM.0000000000004017.
35. Liao L, Chung S, Altamirano J, et al. The association between Asian patient race/ethnicity and lower satisfaction scores. BMC Health Serv Res. 2020;20(1):678. https://doi.org/10.1186/s12913-020-05534-6.
36. Butler C, Baskin AS, Mmonu NA, et al. Patient satisfaction in a safety net urology clinic. Urol Pract. 2021;8(2):168–75. https://doi.org/10.1097/UPJ.0000000000000177.
37. Saha S, Hickam DH. Explaining low ratings of patient satisfaction among Asian-Americans. Am J Med Qual Off J Am Coll Med Qual. 2003;18(6):256–64. https://doi.org/10.1177/106286060301800606.
38. Alfred M, Ubogaya K, Chen X, Wint D, Worral PS. Effectiveness of culturally focused interventions in increasing the satisfaction of hospitalized Asian patients: a systematic review. JBI Database Syst Rev Implement Reports. 2016;14(8):219–56. https://doi.org/10.11124/JBISRIR-2016-003048.
39. Truong M, Paradies Y, Priest N. Interventions to improve cultural competency in healthcare: a systematic review of reviews. BMC Health Serv Res. 2014;14(1):1–17. https://doi.org/10.1186/1472-6963-14-99/TABLES/1.
40. Dunsch F, Evans DK, Macis M, Wang Q. Bias in patient satisfaction surveys: a threat to measuring healthcare quality. BMJ Glob Heal. 2018;3(2):694. https://doi.org/10.1136/BMJGH-2017-000694.