DOI: 10.1111/j.1475-6773.2007.00770.x
RESEARCH ARTICLE

Testing for Statistical Discrimination by Race/Ethnicity in Panel Data for Depression Treatment in Primary Care

Thomas G. McGuire, John Z. Ayanian, Daniel E. Ford, Rachel E. M. Henke, Kathryn M. Rost, and Alan M. Zaslavsky

Objective. To test for discrimination by race/ethnicity arising from clinical uncertainty in treatment for depression, also known as "statistical discrimination."
Data Sources. We used survey data from 1,321 African-American, Hispanic, and white adults identified with depression in primary care. Surveys were administered every six months for two years in the Quality Improvement for Depression (QID) studies.
Study Design. To examine whether and how change in depression severity affects change in treatment intensity by race/ethnicity, we used multivariate cross-sectional and change models that difference out unobserved time-invariant patient characteristics potentially correlated with race/ethnicity.
Data Collection/Extraction Methods. Treatment intensity was operationalized as expenditures on drugs, primary care, and specialty services, weighted by national prices from the Medical Expenditure Panel Survey. Patient race/ethnicity was collected at baseline by self-report.
Principal Findings. Change in depression severity is less associated with change in treatment intensity in minority patients than in whites, consistent with the hypothesis of statistical discrimination. The differential effect by racial/ethnic group was accounted for by use of mental health specialists.
Conclusions. Enhanced physician–patient communication and use of standardized depression instruments may reduce statistical discrimination arising from clinical uncertainty and be useful in reducing racial/ethnic inequities in depression treatment.

Key Words. Statistical discrimination, racial/ethnic disparities, depression

In the Institute of Medicine (IOM) (2003) *Unequal Treatment* report, "discrimination" refers to a health care provider's treatment of patients with similar health care needs differently due to the patient's race or ethnicity. Investigators have studied three types of discrimination—affective bias, cognitive stereotyping, and statistical discrimination (Hilton and von Hippel 1996; Fiske 1998; Dovidio 1999; Balsa and McGuire 2003; IOM 2003; Escarce 2005; Fennell

2005), noting that each type of discrimination has distinct implications for appropriate remedial actions (Balsa and McGuire 2003).

Statistical discrimination appears to be a potent (if more difficult to observe) source of discrimination in health care use (Balsa, McGuire, and Meredith 2005; Lutfey and Ketcham 2005; Werner 2005). In statistical discrimination, providers apply correct information about a group to reduce their clinical uncertainty about an individual patient. While discrimination stemming from affective bias and cognitive stereotyping is generally expected to reduce treatment of minorities in comparison with whites, statistical discrimination may or may not exacerbate race/ethnicity gaps in treatment (Balsa and McGuire 2001). Statistical discrimination can be in the minority patient's best interest if physicians use reliable group differences in the absence of reliable data about the individual. Nonetheless, if physician must rely less on individual level information, treatment for minority patients is less likely to be matched well to their individual needs.

When poor communication exacerbates clinical uncertainty, providers give undue weight to group-based generalizations, taking the form of diminished responsiveness to the circumstances of an individual patient. Prior research has documented worse communication between white providers and minority patients compared with white patients (see, for example, Cooper-Patrick et al. 1999). Because providers primarily rely on patient report to judge the severity of depression, effective communication is particularly important for the physician to allocate severity-appropriate treatment resources for this condition.

The objective of this paper is to test the hypothesis that providers employ statistical discrimination in treating depressed patients over time. Specifically, we assess whether treatment intensity changes as depression severity changes comparably over 2 years for minority and white patients.

Address correspondence to Thomas G. McGuire, Ph.D., Professor of Health Economics, Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115. John Z. Ayanian, M.D., M.P.P., Associate Professor of Health Care Policy, Associate Professor of Medicine, Rachel E. M. Henke, Ph.D., and Alan M. Zaslavsky, Ph.D., Professor of Health Care Policy (Statistics), are with the Department of Health Care Policy, Harvard Medical School, Boston, MA. John Z. Ayanian, M.D., M.P.P., Associate Professor of Health Policy & Management, is also with the Harvard School of Public Health, and as Practicing General Internist, with the Department of Health Care Policy, Division of General Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA. Daniel E. Ford, M.D., M.P.H., is David M. Levine Professor of Medicine, at Johns Hopkins University School of Medicine, Baltimore, MD. Kathryn M. Rost, Ph.D., Professor, is with the Department of Medical Humanities and Social Sciences, Florida State University, Tallahassee, FL.

DATA

The Quality Improvement for Depression Study

The Quality Improvement for Depression study (QID) is a collaborative of four randomized controlled trials designed to evaluate the effectiveness of interventions to improve the quality of care for depressed patients in primary care practices: the Hopkins Quality Improvement for Depression (HQID) Project, the Mental Health Awareness Project (MHAP), the Quality Enhancement by Strategic Teaming (QuEST) Project, and the Partners in Care (PIC) Project. Each project recruited primary care practices from community-based health care organizations/practice networks. Practices were randomized to intervention or usual care conditions using a block randomization design. Participating clinicians within intervention practices were encouraged to adopt depression quality improvement models. Numerous publications have described the effectiveness of the interventions in improving quality of care and outcomes (Rost et al. 2002; Rubenstein et al. 2002; Wells et al. 2004).[1]

Consecutive adult patients were screened for eligibility at each primary care practice during pre-specified enrollment windows. Patients had to report 2 weeks or more during the past year and 1 week or more during the past month when they felt sad, empty, depressed, or lost interest in things they normally enjoyed and plan to continue to receive care in the practice to be eligible. Those who screened positive and met the inclusion criteria were evaluated for major depression over the past year using the Composite International Diagnostic Interview (CIDI). As the exclusion criteria varied somewhat across studies, a QID core sample was developed which included only the patients who met a common set of QID eligibility criteria.

All four projects used standardized survey measures to permit combined data analysis. Detailed clinical, treatment, and demographic information was collected from patients through telephone and mail surveys at baseline and at four periods after baseline: 6, 12, 18, and 24 months. Depression severity was assessed at each wave. Use of mental health services during the previous 6 months was assessed at the end of each period. We excluded the 116 subjects who did not self-identify as Hispanic, non-Hispanic white, or African American.

Mental Health Expenditures

Our primary dependent variable is a measure of mental health services used by each patient over a 6-month period. For each of the four 6-month periods in the study, we estimated each patient's mental health care expenditures using data collected from patients in the follow-up surveys on use of services and

antidepressants in the previous 6 months. Use of mental health services was measured by the number of mental health specialist visits and type of specialist, the number of primary care visits in which depression was discussed, and the approximate number of months (1, 2, or 3+ months) of use of an appropriate dose of antidepressant.[2] To define mental health expenditures we used national averages of prices from the Medical Expenditure Panel Survey (MEPS) using the same methodology described in McGuire et al. (2006).

Depression Severity and Mental Health Functioning

Depression severity was assessed during each 6-month period of the study using the modified Center for Epidemiologic Studies Depression scale (mCES-D), designed to provide an index of the number and frequency of DSM-IV depressive symptoms in the past week, with higher scores indicating more severe depression (Rost et al. 2001). Extensive research has confirmed the reliability and validity of the CES-D across diverse populations (Radloff 1977; Roberts 1980; Aneshensel, Clark, and Frerichs 1983; Fava 1983; Ross and Yinger 2002).

Mental health functioning was measured using the mental health component score (MCS) of the Medical Outcomes Study Short-Form 36 (SF-36; Ware and Sherbourne 1992). In the SF-36 survey, respondents are asked to report functional limitations experienced in the previous 4 weeks. To facilitate comparison with the mCES-D, we reversed the direction of the MCS score so higher MCS scores indicate worse mental health functioning.

Other Patient Information

Patient sociodemographic characteristics from the baseline survey that may be associated with depression treatment including race/ethnicity (Hispanic, African American, and non-Hispanic white), age, gender, marital status, education (highest grade completed), employment (full-time, part-time, temporarily laid off, unemployed–looking, unemployed–not looking, unemployed due to health reasons, student, homemaker, retired, and other), and insurance status (insured or uninsured) were included. We also include two health measures: physical health functioning score from the Medical Outcomes Study Short-Form Survey (PCS), and presence of any comorbid condition (diabetes, hypertension, arthritis, asthma, cancer, epilepsy, stroke, heart failure, angina, back pain, irritable bowel, thyroid disease, kidney failure, and eye disease).

METHODS

Testing for Statistical Discrimination Using Change in Severity of Illness

Statistical discrimination can emerge if: (a) clinicians do not directly observe severity, but only a severity signal such as a symptom report, and (b) this signal is less informative for minority patients than for their white counterparts. Communication difficulties between minority patients and white clinicians might make a signal of severity from a minority patient less informative to a clinician, particularly for an illness such as depression that relies upon subjective reports by the patient about emotional well being. A less informative signal can be thought of as having more "noise" intervening between the patient's true severity and the severity report heard by the clinician. Such noise could be introduced either by the way symptoms are described by the patient or interpreted by the clinician. In either case, the noisier the signal, the less weight the clinician should put on the signal in decisions about treatment. This argument is developed formally in an appendix available online. The argument applies in cross sectional and longitudinal models, as we explain below, and leads to the hypothesis we test in this paper: clinicians respond less to change in symptom severity for minority than for white patients.

Other mechanisms of discrimination involve antipathy toward an outside group or stereotyping, both of which imply negative main effects of minority status in a regression model (Balsa and McGuire 2003). Since race is fixed over time, our change models yield no evidence on the importance of these forms of discrimination. Differential symptom response, which we interpret as evidence for statistical discrimination, is also consistent with other motives, for example, racial stereotypes might only be activated against patients who fail to respond to initial treatment. Such alternative explanations cannot be excluded by our empirical analysis.

Empirical Implementation

Change in severity was calculated for both severity measures, the mCES-D and the MCS, by taking the difference between the scores measured in consecutive waves. Positive severity change indicates worsening depression symptoms/mental health functioning. Change in mental health expenditures was calculated similarly.

The motivation for our test using differencing in panel data can be seen readily in the context of a linear model of level of treatment. Suppose the level of treatment (Y) depends on health status (HS) (here, depression severity),

socioeconomic status (*SES*), and an indicator for minority status (*M*). We set out a model with main effects and interactions but linear in the parameters for patient *k* at time *t*:

$$Y_{kt} = \alpha_0 + \alpha_M M_k + \alpha_H HS_{kt} + \alpha_{HM}(HS_{kt} \cdot M_k) + \alpha_s SES_k + \alpha_{SM}(SES_k \cdot M_k) + \delta_k + \gamma_{kt}, \quad (1)$$

where coefficients α_M, α_H, α_{HM}, α_S, and α_{SM} are to be estimated. Note that group membership *M* (0 for whites, 1 for minorities) and *SES* are assumed to be fixed for any patient (at least over the duration of the study) and are indexed only by *k*, but *HS* may vary over time and is indexed by (*k, t*). The error terms δ_k and γ_{kt} represent, respectively, the persistent and occasion-specific variation in treatment for patient *k*. The γ_{kt} are assumed to be independent of each other and all other variables. The δ_k may not be independent of some right-hand side variables. There is evidence for statistical discrimination if the change in *Y* with *HS* is greater for whites than for minorities.

Under model (1), we would test the hypothesis $\alpha_{HM} < 0$. This test can be conducted with cross-sectional data, but only if we can assume that δ_k is uncorrelated with the covariates *M*, *SES*, and *HS*. This is more than a theoretical problem, since many elements of socioeconomic status and health status are typically unmeasured, and may be correlated with minority status. Thus, the "race/ethnicity effect" in cross-sectional regressions is subject to many interpretations.

Suppose, however, we have data for times *t* and *t*+1. Then, under our assumption that *M* and *SES* are fixed but *HS* can change, the change in level of treatment between the two periods is $\Delta Y_{kt} = \alpha_H \Delta HS_{kt} + \alpha_{HM}(\Delta HS_{kt} \cdot M_k) + (\gamma_{k,t+1} - \gamma_{kt})$, where the final error term is independent of the covariates by assumption. The unmeasured time-invariant factors in δ are differenced out and no longer a source of bias in estimates. Thus, in a change regression with the linear specification, a test of the interaction term between ΔHS and *M* is a test of statistical discrimination.

A similar strategy can be applied in nonlinear regressions. We can generalize model (1) to the GLM (generalized linear model) model $Y_{kt} = f(\eta_{kt}) + \delta_k + \gamma_{kt}$, where $\eta_{kt} = \alpha_0 + \alpha_M M_k + \alpha_H HS_{kt} + \alpha_{HM}(HS_{kt} \cdot M_k) + \alpha_s SES_k + \alpha_{SM}(SES_k \cdot M_k)$ is the same linear predictor as in model (1) and is related to the expectation of Y_{kt} through the possibly nonlinear inverse link function *f*. Typically the distributions of the error terms δ_k and γ_{kt} are specified as heteroskedastic with variance a function of the expectation $f(\eta_{kt})$, such as (Var $\gamma_{kt} = v(f(\eta_{kt}))$. Alternative specifications and diagnostics for such models applied to health care utilization have been explored elsewhere (Manning and

Mullahy 2001; Buntin and Zaslavsky 2004). In such a model the expected value of the difference ΔY cannot be expressed directly as a function of the coefficients as in the linear model. In the case of the logarithmic link function used here, a regression of *relative* changes in treatment on changes in health status constitutes the corresponding test of statistical discrimination in a nonlinear model. See the online Appendix for details.

The Role of SES in Statistical Discrimination by Race/Ethnicity

A more complex specification of the regression model would allow for interactions between *HS* and *SES* as well as between *HS* and M, thus modeling the relationship of responsiveness of treatment to *SES* as well as race. Perhaps members of lower *SES* groups (of any race) communicate less well with health care providers and therefore, for the reasons laid out above in the case of minorities, experience less response in treatment to changes in their health status. Recognition of the possibility of statistical discrimination due to *SES* requires us to be clear about what we mean when assessing statistical discrimination by race/ethnicity. Statistical discrimination is something that *happens to* members of a group (e.g., African Americans); it may or may not be *caused by* group membership. Suppose, for example, that communication patterns were worse for members of lower *SES* groups, but there were no independent effect of race on communication. However, suppose also that African Americans were disproportionately in the lower *SES* groups. Then, while statistical discrimination might be due to *SES*, it would still be accurate to say that African Americans were subject to statistical discrimination in relation to whites. This issue is analogous to the debate about the interpretation of racial/ethnic disparities in health care. The IOM, in its *Unequal Treatment* (IOM 2003) report on health care disparities, argued that the definition of disparity should include racial differences mediated by *SES* differences, as we do in our analyses.

In a linear regression model that includes terms for *M* and *SES*, the IOM definition of disparity includes both the residual race effect (coefficient of *M*) and the effect mediated through *SES* (coefficient of *SES* multiplied by the difference in mean *SES* between whites and minorities) (Cook et al. 2006). The coefficient of *M* in the model that omits *SES* usually yields a similar disparity estimate, absorbing most of the effect mediated through the omitted *SES* variable. The same argument applies to assessing a disparity in *treatment responsiveness*, which we have operationalized here as an interaction with *HS*: the coefficient of *M* in a change model omitting *SES* factors absorbs

effects mediated through both observed and unobserved *SES* factors, closely approximating the full disparity in responsiveness associated with race/ethnicity.

Analyses

We assessed sociodemographic, health status, and site differences for white, Hispanic, and African-American subjects at baseline using χ^2 tests. Racial/ethnic differences in depression severity and mental health expenditures across waves were assessed with F tests. To test for statistical discrimination, we estimated cross-sectional and change models using patient report at baseline and follow-up periods. Our preliminary cross-sectional model includes patient covariates and dummy variables for study site to capture some of the many patient and provider-side factors influencing utilization. These time invariant factors are differenced out, however, in the change models.

We used generalized estimating equations with an exchangeable correlation structure to account for repeated measurements and robust variance estimates to account for clustering of patients within practices. Because mental health expenditures were not normally distributed, we used a modified Park test and methods described in Buntin and Zaslavsky (2004) to identify the best fitting variance distribution and link function. The Poisson variance function and log link offered the best fit and we applied this specification. We added covariates to the cross-sectional models sequentially, in two groups: (1) patient demographics, health status, and intervention variables, (2) site indicators and their interactions with racial/ethnic group, to check if unmeasured factors associated with site affect the results. We specified OLS models for change. We repeated the change models for each component of spending.

To check whether differences in adherence to provider recommendation across racial/ethnic groups could explain differences in mental health expenditures, we examined the association between provider recommendation and utilization across groups. In these models, race/ethnicity, provider recommendation, and an interaction between race/ethnicity and recommendation were specified as main effects and patient and study variables were added as covariates. A smaller association of recommendation and utilization for a group would suggest that differential adherence contributes to the differential relationship of treatment expenditures to severity change.

All analyses were performed with *Stata* version 9.0 (StataCorp 2005) and pairwise deletion of missing data.

RESULTS

Descriptive Information

Our study cohort included 1,321 patients. Table 1 describes patients' characteristics by race/ethnicity. The groups differed by age, marital status, employment, education, health insurance, number of comorbid condition, and health status. African Americans and Hispanics were as likely as whites to be employed and insured. Socioeconomic differences between whites and minorities in our sample appear to be less than one would find in a community sample. The population groups were unevenly distributed across study sites: African Americans were well-represented in all four sites, whites in all but the HQID Project, and Hispanics were concentrated in the MHAP and PIC Project.

Table 2 describes depression severity and mental health expenditures by race and ethnicity across the five waves of data collection. There were no significant differences between the groups in depression symptom severity (mCES-D score) or mental health functioning (MCS score) at baseline. At 6 months, there was a significant difference in depression severity among groups, with African Americans reporting the worst average mCES-D scores and mental health functioning levels (MCS) of the three groups. There were no differences in severity among groups at the 12-, 18-, and 24-month waves. Overall, mental health improved to month 18, then stabilized. Mean mental health expenditures for 6 months before baseline were $633 for all subjects, but only $397 for Hispanic subjects. Mental health expenditures did not differ significantly across groups during any of the four follow-up time periods.

Testing Minority and Severity Interactions

Table 3 shows the cross-sectional association between depression severity and mental health expenditures for each race/ethnicity group. Depression severity (mCES-D or MCS) was a significant predictor of mental health expenditures for the entire group. Panel 3a and 3b show the association between depression severity and mental health expenditures by race/ethnicity (grouping African Americans and Hispanics together to enhance power). The interaction between depression severity (measured by mCES-D) and race/ethnicity was negative, as would be expected with statistical discrimination, but not significant. The interaction between depression severity (measured by MCS) and race/ethnicity was negative and significant ($p < .05$) indicating that treatment is about half as responsive to MCS for minorities as for whites. The point estimates for the Hispanic and African-American interactions with MCS

Table 1: Description of the Sample by Race/Ethnicity

	Whites (n = 994)	*Hispanics (n = 200)*	*African Americans (n = 127)*	*All (n = 1,321)*
Demographics				
Gender				
Male	0.27	0.24	0.21	0.26
Female	0.73	0.76	0.79	0.74
*Age**				
18–24	0.24	0.35	0.22	0.25
35–49	0.40	0.42	0.51	0.41
50–64	0.26	0.20	0.25	0.25
65+	0.10	0.04	0.02	0.08
*Marital status**				
Married	0.47	0.47	0.33	0.46
Not married	0.53	0.53	0.67	0.54
*Work status**				
Full-time	0.50	0.55	0.62	0.52
Part-time	0.12	0.11	0.06	0.11
Temporarily laid off	0.01	0.00	0.00	0.00
Unemployed, looking	0.04	0.07	0.03	0.04
Unemployed, health reasons	0.08	0.12	0.14	0.09
Unemployed, not looking	0.02	0.01	0.02	0.01
Retired	0.11	0.03	0.07	0.09
Student	0.02	0.03	0.02	0.02
Homemaker	0.07	0.06	0.02	0.06
Other	0.27	0.31	0.25	0.03
*Education**				
Less than high school	0.11	0.29	0.13	0.14
High school graduate	0.35	0.26	0.30	0.33
Some college	0.40	0.42	0.47	0.41
College graduate	0.14	0.03	0.10	0.12
*Health insurance**				
Insured	0.95	0.98	0.97	0.96
Not insured	0.05	0.02	0.03	0.04
Health Status				
*Comorbid conditions**				
One or more	0.75	0.74	0.82	0.76
None	0.25	0.26	0.18	0.24
*Health status**				
Excellent	0.05	0.03	0.06	0.05
Very good	0.21	0.16	0.07	0.19
Good	0.40	0.40	0.43	0.40
Fair	0.25	0.29	0.31	0.26
Poor	0.09	0.13	0.13	0.10
Distribution over studies				
Intervention	0.63	0.66	0.62	0.63
Control	0.37	0.34	0.38	0.37

Table 1: *Continued*

	Whites ($n = 994$)	*Hispanics* ($n = 200$)	*African Americans* ($n = 127$)	*All* ($n = 1,321$)
*QID study**				
Mental Health Awareness Project ($n = 449$)	0.35	0.34	0.28	0.34
Quality Enhancement by Strategic Teaming ($n = 342$)	0.30	0.06	0.24	0.26
Hopkins Quality Improvement for Depression ($n = 70$)	0.03	0.03	0.26	0.05
Partners in Care ($n = 460$)	0.32	0.58	0.22	0.35

Source: Baseline data from Quality Improvement for Depression (QID) core including whites, Hispanics, and African Americans.

*Difference across groups significant at the 0.05 level in chi-squared analysis.

estimated separately are very similar in magnitude. We checked this interaction in models with few and many covariates, and the finding was robust to these alternative specifications.

Table 4 shows the longitudinal relationship between changes in depression severity and mental health expenditures for each race/ethnicity group, comparable to the cross-sectional models. Worsening severity (mCES-D or MCS) predicted greater expenditures for the entire group. Panels 4a and 4b (column [3]) show the longitudinal relationship between depression severity and mental health expenditures by race/ethnicity. The interaction between minority race/ethnicity (grouping African Americans and Hispanics together) and change in depression severity (measured by mCES-D) was negative, as would be expected with statistical discrimination, but not significant. The interaction between minority race/ethnicity and change in depression severity (measured by MCS) was negative and significant ($p < .05$) indicating that treatment is about half as responsive to change in MCS for minorities as for whites. Although point estimates for separate African-American and Hispanic indicators are similar (column [4]), the reduced sample size for separate estimation of the effects weakens the significance. Interestingly, the estimates from the cross-sectional model are very close to the estimates from the longitudinal analysis, implying that, at least in this application, unmeasured factors may not be biasing cross-sectional findings.

Table 5 displays the results for one of the three components of total mental health expenditures, mental health specialist visits. The negative interaction between race/ethnicity and change in mental health was evident for the both mCES-D ($p < .10$) and MCS ($p < .05$). Interactions were negative for

Table 2: Depression Severity and Mental Health Services Expenditures across Waves, by Race/Ethnicity

	Baseline	*6 Months*	*12 Months*	*18 Months*	*24 Months*
Depression severity (mCES-D)					
All	50.60 (20.87)	39.31 (24.96)*	37.11 (24.81)	36.33 (24.14)	37.64 (24.36)
White	50.43 (21.00)	39.12 (24.90)	36.62 (24.87)	36.87 (24.08)	38.04 (24.45)
Hispanic	51.59 (19.46)	36.81 (24.24)	39.13 (24.26)	31.68 (23.99)	36.67 (23.67)
African American	50.36 (22.02)	44.28 (25.95)	38.04 (25.17)	37.63 (24.38)	35.86 (24.58)
Mental health functioning (MCS)					
All	38.44 (9.83)	30.56 (11.99)	29.86 (12.34)	29.19 (12.12)	30.14 (12.54)
White	38.32 (9.69)	30.47 (11.71)	29.43 (12.38)	29.26 (12.14)	30.41 (12.56)
Hispanic	39.08 (9.88)	29.36 (12.48)	31.15 (12.19)	27.81 (12.07)	29.51 (12.27)
African American	38.41 (10.86)	32.95 (13.07)	31.32 (12.10)	30.17 (11.98)	28.92 (12.72)
Expenditures on mental health care					
All	633.07 (1,449.44)*	709.70 (1,120.90)	590.45 (1,126.59)	487.2 (979.18)	515.89 (1,167.54)
White	672.23 (1,592.40)	722.42 (1,168.50)	576.33 (1,103.86)	517.6 (1,069.44)	517.50 (1,155.02)
Hispanic	396.57 (711.60)	656.94 (948.44)	517.68 (869.30)	414.04 (653.73)	488.22 (945.32)
African American	699.01 (1,058.26)	689.55 (985.17)	780.78 (1,503.12)	345.87 (402.90)	534.83 (1,454.41)

Source: All waves data from QID core including only whites, Hispanics, and African Americans.

The number of observations varies across waves due to loss to follow-up and missing data.

Expenditures calculated using MEPS price weights as follows: PCP visit = $69; Psychiatrist visit = $100; Non-Psychiatrist MH specialty visit = $85; Antidepressant 1 month = $75; Antidepressant 2 months = $150; Antidepressant 2+months = $300.

mCES-D, modified Center for Epidemiologic Studies Depression scale; QID, Quality Improvement for Depression; MCS, mental health component score; MEPS, Medical Expenditure Panel Survey.

*F-test significant across groups at 5 percent level.

Table 3: Depression Severity and Total Mental Health Expenditures by Race/Ethnicity: Cross-Sectional Analysis

	1	*2*	*3*	*4*	*5*	*6*
Panel (a): mCES-D						
Hispanic	−0.089	0.000	0.010	0.097	0.096	0.041
African American	−0.010	0.085	0.109	−0.028	0.009	0.034
mCES-D	0.013**	0.014**	0.013**	0.014**	0.013**	0.013**
mCES-D × Minority		−0.002	−0.002			
mCES-D × Hispanic				−0.004	−0.004	−0.004
mCES-D × African American				0.000	0.000	0.000
Panel (b): MCS						
Hispanic	−0.100	−0.556*	−0.572**	−0.587*-	−0.574*	0.617*
African American	−0.021	−0.459*	−0.458*	−0.428^{+}	−0.456^{+}	−0.387^{+}
MCS	0.025**	0.028**	0.027**	0.028**	0.027**	0.028**
MCS × Minority		−0.012*	−0.013*			
MCS × Hispanic				−0.013^{+}	−0.013^{+}	−0.012^{+}
MCS × African American				−0.011	−0.013^{+}	−0.011

Source: All waves data from QID core including only whites, Hispanics, and African Americans ($n = 4{,}000$).

Expenditures calculated using MEPS price weights.

Dependent variable: Total mental health expenditures.

All models control for wave of data collection.

Models 3, 5, and 6 adjusted for patient covariates: gender, marital status, employment, education, insurance, physical health functioning, age, comorbitidies, intervention status.

Model 6 also includes QID study indicators and interaction between QID study and race/ethnicity.

mCES-D, modified Center for Epidemiologic Studies Depression scale; QID, Quality Improvement for Depression; MCS, mental health component score.

**$p < 0.01$,

*$p < 0.05$,

$^{+}p < 0.10$.

both African Americans and Hispanics separately. The negative interaction between Hispanic and change in mCES-D was significant ($p < .05$). In contrast, in results available in the online appendix, comparable interaction terms for primary care visits and antidepressant medications showed no clear pattern and were not statistically significant.

Sensitivity Analyses

We conducted a number of sensitivity analyses of the change regressions to examine if the magnitude and significance of the minority-severity interactions varied according to the way we defined change in expenditures. First, we analyzed the raw, rather than scaled, expenditure difference for total expen-

Table 4: Influence of Change in Depression Severity on Change in Total Mental Health Expenditures by Race/Ethnicity: Longitudinal Analysis

	1	2	3	4
Panel (a): Change in mCES-D (ΔmCES-D)				
Hispanic	0.124^{+}	0.135^{+}	0.118^{+}	0.117^{+}
African American	−0.018	−0.014	−0.029	−0.027
ΔmCES-D		0.013**	0.014**	0.014**
ΔmCES-D × Minority			−0.004	
ΔmCES-D × Hispanic				−0.004
ΔmCES-D × African American				−0.004
Panel (b): Change in MCS (ΔMCS)				
Hispanic	0.132^{+}	0.144*	0.112	0.108
African American	0.021	0.029	−0.001	0.004
ΔMCS		0.025**	0.028**	0.028**
ΔMCS × Minority			−0.013*	
ΔMCS × Hispanic				−0.014^{+}
ΔMCS × African American				0.011

Source: All Waves Data from QID Core including whites, Hispanics, and African Americans ($n = 3{,}983$).

Expenditures calculated using MEPS price weights.

Analysis: GEE with exchangeable correlation structure.

Dependent variable: Standardized change in total mental health expenditures.

All models control for wave of data collection.

mCES-D, modified Center for Epidemiologic Studies Depression scale; QID, Quality Improvement for Depression; MCS, mental health component score.

**$p < 0.01$,

*$p < 0.05$,

$^{+}p < 0.10$.

ditures. Our results were "stronger" in the sense that for both the change in mCES-D and MCS we found significant negative interactions ($p < .05$), and for minority ($p < .05$) for change in MCS. The sign of all interactions was negative, and in the case of MCS, being minority offset about half of the effect of a change of MCS on expenditures as compared with whites. We next examined the impact of an alternative scaling of expenditures, dividing by the sum of actual expenditures in the two periods defining a change rather than using the fitted values. This alternative has the effect of diminishing the impact of large unpredictable changes in spending. With this alternative, our results "weaken." The negative interaction with change in MCS approaches significance only for Hispanics ($p < .10$) and not for African Americans.

We also checked whether specialist referral was as predictive for minorities on use of the specialist as for whites. We found that 64 percent

Table 5: Influence of Change in Depression Severity on Change in Expenditures for Mental Health Professional Visits by Race/Ethnicity: Longitudinal Analysis

	1	*2*	*3*	*4*
Panel (a): Change in mCES-D (ΔmCES-D)				
Hispanic	0.131	0.137	0.111	0.106
African American	− 0.022	− 0.016	− 0.04	− 0.034
ΔmCES-D		0.015**	0.016**	0.016**
ΔmCES-D × Minority			− 0.007^{+}	
ΔmCES-D × Hispanic				− 0.008
ΔmCES-D × African American				0.005
Panel (b): Change in MCS (ΔMCS)				
Hispanic	0.137	0.144	0.097	0.084
African American	0.033	0.043	− 0.003	0.014
ΔMCS		0.030**	0.034**	0.034**
ΔMCS × Minority			− 0.020*	
ΔMCS × Hispanic				− 0.025*
ΔMCS × African American				0.013

Source: All Waves Data from Quality Improvement for Depression (QID) Core including whites, Hispanics, and African Americans ($n = 3{,}991$).

Expenditures calculated using MEPS price weights.

Analysis: GEE with exchangeable correlation structure.

Dependent variable: Standardized change in expenditures for mental health specialists.

All models control for wave of data collection.

mCES-D, modified Center for Epidemiologic Studies Depression scale; QID, Quality Improvement for Depression; MCS, mental health component score.

**$p < 0.01$,

*$p < 0.05$,

$^{+}p < 0.10$.

of patients who received a referral reported at least one visit to a mental health specialist. The rate did not vary significantly by race/ethnicity (64 percent whites, 62 percent Hispanics, and 63 percent African American) after controlling for all patient and study covariates. This provides evidence against the role of differential patient adherence—at least at this stage—as an explanation for the different patterns of expenditures across race/ethnicity groups.

DISCUSSION

This study used an innovative analytic approach to assess whether changes in the severity of depression were associated with comparable changes in mental

health expenditures for Hispanic, African-American, and white patients followed over 2 years in primary care practices in multiple U.S. regions. Expenditures responded to changes in severity, but the degree of response overall was about half as large for minority patients as for whites. The degree of lower response is similar for Hispanics and African Americans. The effect was more consistently significant for Hispanics, presumably due to the larger number of Hispanics in our cohort.

The longitudinal analysis produced similar findings to the cross-sectional analysis, indicating that unmeasured sociodemographic factors did not substantially confound cross-sectional conclusions about differential treatment by race/ethnicity. The minimal importance of unmeasured sociodemographic factors echoes previous findings that measured sociodemographic factors do not explain disparities in depression care for Hispanics (Lagomasino et al. 2005).

Statistical discrimination is one plausible mechanism for health care disparities, arising if health care providers are less confident of the clinical information they obtain for minority patients. In these situations, providers become more reliant on group characteristics and less sensitive to individual factors, such as changes in patient burden of illness. Our results suggest that statistical discrimination affected specialty care for depression care among minority patients but we lacked detailed data to explain how this pattern arose. Specialty mental health care depends upon a primary care physician referral, patient adherence with the referral, and specialty care decisions about the number of treatment visits. Differences by race/ethnicity at any of these steps might reduce specialty mental health expenditures for minorities. In our sensitivity analyses, we did not find evidence for differential adherence by race/ethnicity to specialist referrals.

Previous studies have demonstrated that Hispanic and African-American patients are less satisfied with health care provider communication and access to specialty care (Morales et al. 1999; Lurie et al. 2003; Ayanian et al. 2005) than their white counterparts, factors which may increase statistical discrimination in minority patients. Improving communication between minority patients and their health care providers may help to reduce reliance on group characteristics in decision making for depression care and thereby improve the mental health outcomes of minority patients. It is intriguing that the MCS, a measure of severity that captures functioning as well as symptoms, is a more sensitive measure of the information that is differentially observed in communication between minority patients and their physicians, compared with the symptom-driven mCES-D, suggesting that minority patients may feel

less social permission than their white counterparts to discuss how a devastating disease like depression is affecting their ability to function. Because language differences may contribute to statistical discrimination for Spanish-speaking Hispanic patients (Baker, Hayes, and Fortier 1998; Woloshin et al. 2001), greater availability of bilingual providers and professional translators may improve their communication with providers about the severity and treatment of depression (Office of Minority Health 2001).

Our findings also suggest that for depression care these efforts should focus on optimizing the frequency of visits to mental health specialists, as the responsiveness of primary care visits and antidepressant medications was similar across racial/ethnic groups. Potential policy solutions include greater use of standardized depression severity measures by providers of depression care, as well as the use of trained navigators to help patients overcome barriers to seeing mental health specialists—a model currently being implemented in some centers to improve access to cancer specialists (Dohan and Schrag 2005).

There are several limitations associated with the use of the QID data. First, patient race and ethnicity information were collected in broad, heterogeneous categories (e.g. African-American, Hispanic, white) and degree of acculturation, language, and patient-provider concordance was not assessed. Information about mental health service utilization relied on patient self-report. This data collection strategy may have led to under-reporting of service use from recall bias. However, recall bias has not been shown to vary by race/ethnicity and thus any under-report should be similar across patient groups and should not affect our results. Using data from patients voluntarily participating in a research study rather than data from a population sample limits the generalizability of findings. Finally, our measure of antidepressant expenditures did not differentiate the cost of generic vs. nongeneric drugs or specify an additional cost associated with use of multiple antidepressants at one time. However, as our interest lies in the response of severity to mental health treatment received and not in an exact quantification of resource use, our use of a single price per month for any type or number of antidepressants achieves this objective.

A particular strength of the QID cohort was that nearly all participants were insured and had some access to primary care, so the typically higher rates of uninsurance and corresponding access barriers for Hispanic and African-American patients were unlikely to have played a major role in the findings of our study.

Evidence is accumulating that statistical discrimination, emerging as a consequence of heightened clinical uncertainty about treatment for racial/

ethnic minorities, plays some role in explaining the discriminatory treatment received by these groups. These findings point to the importance of supporting clinicians by improving decision making tools, including training for clinicians in collecting and interpreting information from members of minority groups, and formal decision supports that may facilitate better communication and elicitation of pertinent information. Improving patient–physician communication may improve the quality of care for all groups, and may be particularly beneficial for minority populations.

ACKNOWLEDGMENTS

This research was supported by grants R01 MH068260 and P50 MH073469 from the National Institute of Mental Health. We thank Virginia Wilcox-Gok for comments on an earlier draft. This research was approved by the Institutional Review Board of Harvard Medical School, file M1240.

Disclosures: None.

Disclaimers: None.

NOTES

1. For more details on the QID study design and administration, refer to Rost et al. (2001). Previous research on race/ethnic differences using data from the QID studies has found lower rates of mental health service use by minorities at baseline (Miranda and Cooper 2004), differential clinical and functional outcome improvement for Hispanics compared with whites (Miranda et al. 2003, Miranda et al. 2004), and disparities in depression care at baseline for Hispanics not attributable to clinic, sociodemographic, or clinical characteristics (Lagomasino et al. 2005).
2. We define appropriate dose as receipt of guideline daily dosage (i.e., at or above the minimum dosage recommended in the AHRQ Guidelines, with comparable criteria for newer antidepressant medications for 25 days of more) (Miranda et al. 2003).

REFERENCES

Aneshensel, C. S., V. A. Clark, and R. R. Frerichs. 1983. "Race, Ethnicity, and Depression: A Confirmatory Analysis." *Journal of Personality and Social Psychology* 44 (2): 385–98.

Ayanian, J. Z., A. M. Zaslavsky, E. Guadagnoli, C. S. Fuchs, K. J. Yost, C. M. Creech, R. D. Cress, L. C. O'Connor, D. W. West, and W. E. Wright. 2005. "Patients'

Perceptions of Quality of Care for Colorectal Cancer by Race, Ethnicity, and Language." *Journal of Clinical Oncology* 23 (27): 6576–86.

Baker, D. W., R. Hayes, and J. P. Fortier. 1998. "Interpreter Use and Satisfaction with Interpersonal Aspects of Care for Spanish-Speaking Patients." *Medical Care* 36 (10): 1461–70.

Balsa, A. I., and T. G. McGuire. 2001. "Statistical Discrimination in Health Care." *Journal of Health Economics* 20 (6): 882–907.

———. 2003. "Prejudice, Clinical Uncertainty and Stereotyping as Sources of Health Disparities." *Journal of Health Economics* 22 (1): 89–116.

Balsa, A. I., T. G. McGuire, and L. S. Meredith. 2005. "Testing for Statistical Discrimination in Health Care." *Health Services Research* 40 (1): 227–52.

Buntin, M. B., and A. M. Zaslavsky. 2004. "Too Much Ado about Two-Part Models and Transformation? Comparing Methods of Modeling Medicare Expenditures." *Journal of Health Economics* 23 (3): 525–42.

Cook, B. L., T. G. McGuire, E. Meara, and A. Zaslavsky. 2006. "Adjusting for Health Status in Non-Linear Models of Health Care Disparities." Unpublished manuscript.

Cooper-Patrick, L., J. J. Gallo, J. J. Gonzales, H. T. Vu, N. R. Powe, C. Nelson, and D. E. Ford. 1999. "Race, Gender, and Partnership in the Patient–Physician Relationship." *Journal of the American Medical Association* 282 (6): 583–9.

Dohan, D., and D. Schrag. 2005. "Using Navigators to Improve Care of Underserved Patients: Current Practices and Approaches." *Cancer* 104 (4): 848–55.

Dovidio, J. F. 1999. "Stereotyping." In *The MIT Encyclopedia of the Cognitive Sciences*, edited by R. A. Wilson and F. C. Keil. Cambridge, MA: MIT Press.

Escarce, J. J. 2005. "How Does Race Matter, Anyway?" *Health Services Research* 40 (1): 1–7.

Fava, G. A. 1983. "Assessing Depressive Symptoms across Cultures: Italian Validation of the CES-D Self-Rating Scale." *Journal of Clinical Psychology* 39 (2): 249–51.

Fennell, M. L. 2005. "Racial Disparities in Care: Looking beyond the Clinical Encounter." *Health Services Research* 40 (6 Part 1): 1713–21.

Fiske, S. 1998. "Stereotyping, Prejudice, and Discrimination." In *The Handbook of Social Psychology*, edited by D. T. Gilbert, S. T. Fiske, and G. Lindzey, pp. 357–411. New York: Guilford Press.

Hilton, J. L., and W. von Hippel. 1996. "Stereotypes." *Annual Review of Psychology* 47: 237–71.

Institute of Medicine (IOM). 2003. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care.* Washington, DC: National Academies Press.

Lagomasino, I. T., M. Dwight-Johnson, J. Miranda, L. Zhang, D. Liao, N. Duan, and K. B. Wells. 2005. "Disparities in Depression Treatment for Latinos and Site of Care." *Psychiatric Services* 56 (12): 1517–23.

Lurie, N., C. Zhan, J. Sangl, A. S. Bierman, and E. S. Sekscenski. 2003. "Variation in Racial and Ethnic Differences in Consumer Assessments of Health Care." *American Journal of Managed Care* 9 (7): 502–9.

Lutfey, K. E., and J. D. Ketcham. 2005. "Patient and Provider Assessments of Adherence and the Sources of Disparities: Evidence from Diabetes Care." *Health Services Research* 40 (6, part 1): 1803–17.

Manning, W. G., and J. Mullahy. 2001. "Estimating Log Models: To Transform or Not to Transform?" *Journal of Health Economics* 20 (4): 461–94.

McGuire, T. G., M. Alegria, B. L. Cook, K. B. Wells, and A. M. Zaslavsky. 2006. "Implementing the Institute of Medicine Definition of Disparities: An Application to Mental Health Care." *Health Services Research* 41 (5): 1979–2005.

Miranda, J., and L. A. Cooper. 2004. "Disparities in Care for Depression among Primary Care Patients." *Journal of General Internal Medicine* 19 (2): 120–6.

Miranda, J., N. Duan, C. Sherbourne, M. Schoenbaum, I. Lagomasino, M. Jackson-Triche, and K. B. Wells. 2003. "Improving Care for Minorities: Can Quality Improvement Interventions Improve Care and Outcomes for Depressed Minorities? Results of a Randomized, Controlled Trial." *Health Services Research* 38 (2): 613–30.

Miranda, J., M. Schoenbaum, C. Sherbourne, N. Duan, and K. Wells. 2004. "Effects of Primary Care Depression Treatment on Minority Patients' Clinical Status and Employment." *Archive of General Psychiatry* 61 (8): 827–34.

Morales, L. S., W. E. Cunningham, J. A. Brown, H. Liu, and R. D. Hays. 1999. "Are Latinos Less Satisfied with Communication by Health Care Providers?" *Journal of General Internal Medicine* 14 (7): 409–17.

Office of Minority Health. 2001. *National Standards for Culturally and Linguistically Appropriate Services in Health Care.* Washington, DC: Department of Health and Human Services.

Radloff, L. S. 1977. "The CES-D Scale: A Self-Report Depression Scale for Research in the General Population." *Applied Psychology Measurements* 1: 385–401.

Roberts, R. E. 1980. "Reliability of the CES-D Scale in Different Ethnic Contexts." *Psychiatry Research* 2 (2): 125–34.

Ross, S. L., and J. Yinger. 2002. *The Color of Credit: Mortgage Discrimination, Research Methodology and Fair-Lending Enforcement.* Boston: MIT Press.

Rost, K. M., N. Duan, L. V. Rubenstein, D. E. Ford, C. D. Sherbourne, L. S. Meredith, and K. B. Wells. 2001. "The Quality Improvement for Depression Collaboration: General Analytic Strategies for a Coordinated Study of Quality Improvement in Depression Care." *General Hospital Psychiatry* 23 (5): 239–53.

Rost, K. M., P. Nutting, J. L. Smith, C. E. Elliott, and M. L. Dickinson. 2002. "Managing Depression as a Chronic Disease: A Randomized Trial of Ongoing Primary Care Depression Treatment." *British Medical Journal* 325: 934–9.

Rubenstein, L. V., L. E. Parker, L. S. Meredith, A. Altschuler, E. dePillis, J. Hernandez, and N. P. Gordon. 2002. "Understanding Team-Based Quality Improvement for Depression in Primary Care." *Health Services Research* 37 (4): 1009–29.

StataCorp. 2005. *Stata Statistical Software (Release 9) Software.* College Station, TX: StataCorp LP.

Ware, J. E. Jr., and C. D. Sherbourne. 1992. "The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual Framework and Item Selection." *Medical Care* 30 (6): 473–83.

Wells, K., C. Sherbourne, M. Schoenbaum, S. Ettner, N. Duan, J. Miranda, J. Unutzer, and L. Rubenstein. 2004. "Five-Year Impact of Quality Improvement for Depression: Results of a Group-Level Randomized Controlled Trial." *Archive of General Psychiatry* 61 (4): 378–86.

Werner, R. 2005. *Testing Theories of Discrimination in Health Care: Evidence from New York's CABG Report Card.* Ph.D. dissertation, University of Pennsylvania, Philadelphia.

Woloshin, S., L. M. Schwartz, J. Tremmel, and H. G. Welch. 2001. "Direct-to-Consumer Advertising for Prescription Drugs: What Are Americans Being Sold?" *Lancet* 358 (9288): 1141–6.

SUPPLEMENTARY MATERIAL

The following supplementary material for this article is available online:

Table 6. Influence of Change in Depression Severity on Change in Expenditures for Primary Care Visits by Race/Ethnicity: Longitudinal Analysis.

Table 7. Influence of Change in Depression Severity on Change in Expenditures for Antidepressants by Race/Ethnicity: Longitudinal Analysis.

Table 8. Depression Severity (CES-D) and Total Mental Health Expenditures by Race/Ethnicity: Cross-Sectional Analysis.

Table 9. Depression Severity (MCS) and Total Mental Health Expenditures by Race/Ethnicity: Cross-Sectional Analysis.

This material is available as part of the online article from: http://www.blackwellsynergy.com/doi/abs/10.1111/j.1475-6773.2007.00770.x (this link will take you to the article abstract).

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.