ORIGINAL ARTICLE

ONLINE FIRST

Influence of Patient Race and Ethnicity on Clinical Assessment in Patients With Affective Disorders

Michael A. Gara, PhD; William A. Vega, PhD; Stephan Arndt, PhD; Michael Escamilla, MD; David E. Fleck, PhD; William B. Lawson, MD, PhD; Ira Lesser, MD; Harold W. Neighbors, PhD; Daniel R. Wilson, MD, PhD; Stephen M. Strakowski, MD

Context: Rates of clinical diagnoses of schizophrenia in African American individuals appear to be elevated compared with other ethnic groups in the United States, contradicting population rates derived from epidemiologic surveys.

Objective: To determine whether African American individuals would continue to exhibit significantly higher rates of clinical diagnoses of schizophrenia, even after controlling for age, sex, income, site, and education, as well as the presence or absence of serious affective disorder, as determined by experts blinded to race and ethnicity. A secondary objective was to determine if a similar pattern occurred in Latino subjects.

Design: Ethnicity-blinded and -unblinded diagnostic assessments were obtained in 241 African American individuals (mean [SD] age, 34.3 [8.1] years; 57% women), 220 non-Latino white individuals (mean [SD] age, 32.7 [8.5] years; 53% women), and 149 Latino individuals (mean [SD] age, 33.5 [8.0] years; 58% women) at 6 US sites. Logistic regression models were used to determine whether elevated rates of schizophrenia in African American individuals would persist after controlling for various confounding variables including blinded expert consensus diagnoses of serious affective illness.

Settings: Six academic medical centers across the United States.

Participants: Six hundred ten psychiatric inpatients and outpatients.

Main Outcome Measure: Relative odds of unblinded clinical diagnoses of schizophrenia in African American compared with white individuals.

Results: A significant ethnicity/race effect (χ^2_2=10.4, P=.01) was obtained when schizophrenia was narrowly defined, controlling for all other predictors. The odds ratio comparing African American with non-Latino white individuals was significant (odds ratio=2.7; 95% CI, 1.5-5.1). Similar differences between African American and white individuals occurred when schizophrenia was more broadly defined (odds ratio=2.5; 95% CI, 1.4-4.5). African American individuals did not differ significantly from white individuals in overall severity of manic and depressive symptoms but did evidence more severe psychosis.

Conclusions: African American individuals exhibited significantly higher rates of clinical diagnoses of schizophrenia than non-Latino white subjects, even after controlling for covariates such as serious affective disorder.

Arch Gen Psychiatry.
Published online February 6, 2012.
doi:10.1001/archgenpsychiatry.2011.2040

Author Affiliations are listed at the end of this article.

In the United States, African American individuals in clinical venues are disproportionately diagnosed with schizophrenia compared with white individuals.[1-7] These clinical discrepancies contrast with the Epidemiologic Catchment Area Study in which similar prevalence rates of nonaffective psychoses were observed among African American, Latino, and non-Latino white subjects.[8,9] In clinical practice, the degree of diagnostic bias and its potential causes are unsettled questions. Unfortunately, there are no pathognomonic markers for schizophrenia that would move case assignment onto firmer ground than the present reliance on clinical assessment.[10] Schizophrenia is a complex condition that presents with an extensive range of symptoms and signs that also occur in other serious mental disorders especially mood disorders.[11] Consequently, distinguishing schizophrenia from other conditions can be difficult, although by definition these other conditions must be ruled out before diagnosing schizophrenia.

Racial or ethnic bias in assigning clinical diagnoses is not simply an academic exercise. Failure to identify affective syndromes in psychotic patients may lead to inadequate or inappropriate treatment assignment. Moreover, incorrectly assigning a diagnosis of schizophrenia to a pa-

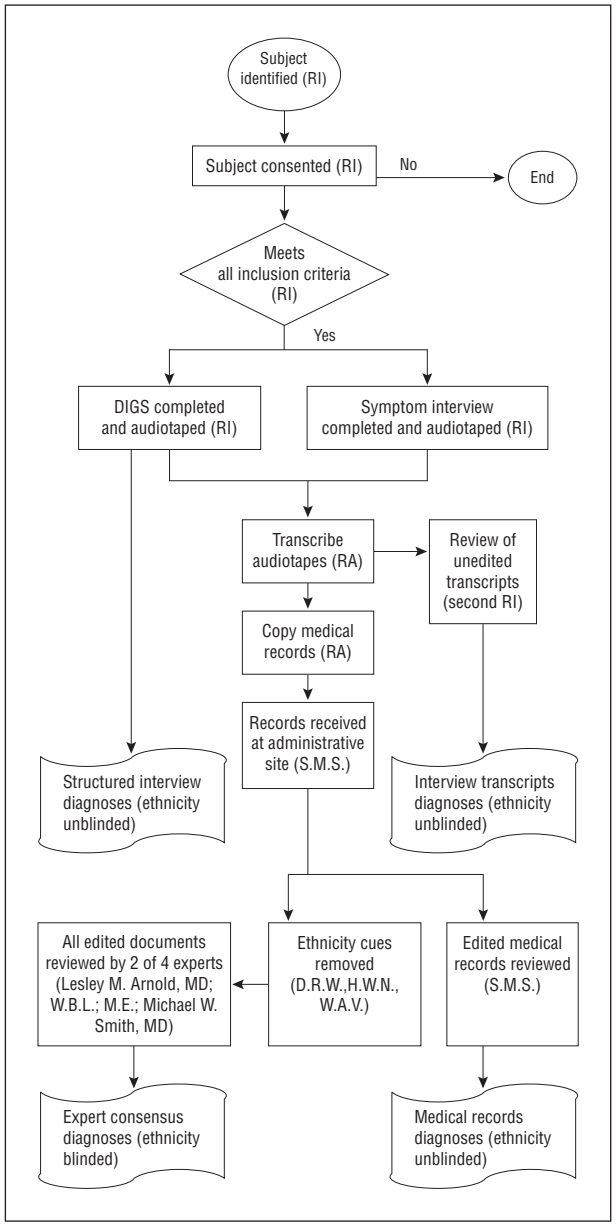

Figure. Flowchart depicting study procedures. DIGS indicates Diagnostic Interview for Genetic Studies; RA, research assistant; and RI, research interviewer.

tient with another condition may lead to assumptions about a poor prognosis that limits treatment options or defeats expectations for a good outcome. Consequently, understanding how race/ethnicity impacts clinical assessment in these conditions can help to eliminate racial disparities in health care.

For decades, there has been an appreciation that individual mental experience is affected by a matrix of factors including sociocultural and racial background. More recently, Kleinman et al[12] and others have emphasized how different ethnic groups have unique explanatory models of mental experience. Because of divergent explanatory models, patients and clinicians can be seen to sometimes inhabit different cultural contexts that introduce misunderstanding of patient concerns or misconstrual of patient history.

Previously, Strakowski and colleagues[3,13] suggested that racial bias in the diagnosis of schizophrenia results from several factors, one of which is attributable to clinicians overemphasizing the relevance of psychotic symptoms relative to affective symptoms in African American individuals. However, previous studies of racial effects on diagnosis are vulnerable to effects attributable to ethnic or racial bias in both clinical and research assessments due to a failure to blind clinicians and researchers to the ethnicity of subjects.[14] Additionally, prior studies typically occurred at a single clinical site, limiting generalizability.

To overcome these limitations, we made multiethnic comparisons at 6 regional sites in the United States using both ethnicity-blinded and -unblinded assessments. We included a Latino population to determine whether biases in diagnostic assignment occur generally in minorities or are specific to African American subjects. The study hypothesis was that African American individuals would exhibit significantly higher rates of clinical diagnoses (unblinded) in the schizophrenia spectrum than would non-Latino white subjects, even after controlling for confounding factors such as the presence of serious affective disorder as well as age, sex, income, site, and education. A corollary to this prediction is that African American individuals would receive higher ratings of psychotic symptoms and lower ratings of affective symptoms.

METHODS

STUDY OVERVIEW

The various data collection and extraction processes are shown in the **Figure**. Diagnostic data were collected at 6 sites from 4 sources: (1) medical records, (2) structured interviews, (3) interview transcripts, and (4) blinded expert consensus. The medical record provided the first data source. Racial and ethnic cues were deleted from medical records (as described later), and these edited records were reviewed at the University of Cincinnati to identify the primary psychiatric diagnoses as formulated by treating clinicians, who nearly always were psychiatrists. Medical records were blinded to subjects' race and ethnicity to guard against investigator bias when extracting this information. However, the clinical diagnoses themselves were, of course, unblinded.

Structured interviews provided the second set of diagnostic data. At each site, trained clinically experienced interviewers completed the Diagnostic Interview for Genetic Studies (DIGS)[15] for each subject, supplemented with additional symptom questionnaires. The interviewers also reviewed all available medical records to assist with this assessment and were instructed to be as complete as possible to make the best assessment. Demographic data including the subject's self-ascribed racial and ethnic assignment were obtained during these interviews, which were audiotaped and transcribed. Subsequently, a second independent interviewer at each site reviewed the transcripts and unblinded medical records and provided an independent diagnostic assessment and symptom ratings. Hence, the raw interview transcripts provided an assessment of the potential effects of making the diagnosis from a transcription rather than a direct interview.

Medical records and transcripts were sent to the University of Cincinnati for archiving and distribution to editors. Editors carefully reviewed each set of medical records and transcripts and removed racial/ethnic cues. Interview text was edited to eliminate references to people, names, places, and linguistic idiosyncra-

sies, phrasing, or syntax describing personal histories or situations that could cue the diagnostic rater to specific ethnicities. These redacted documents were then sent to a second editor. Secondary editors compared the modified text with the original transcript to ensure both adequate blinding and equivalence of meaning. If not, the second editor completed a second level of redaction and then sent it to a third editor for final approval.

The redacted records and transcripts were then distributed to 2 independent diagnostic experts, who rated symptoms and diagnosis independently and then discussed the case by telephone to derive a consensus assessment. The experts reviewed diagnostic criteria and the basis for each of their decisions and then discussed disagreements until a consensus could be reached. Expert consensus diagnoses and ratings composed the fourth source of study data.

STUDY DETAILS

Sites

The 6 sites and number of subjects per site were Harbor–University of California at Los Angeles Medical Center (n=118 patients), University of Texas Health Science Center at San Antonio (n=105), Howard University (n=101), University of Michigan (n=99), University of Medicine and Dentistry of New Jersey University Behavioral HealthCare (n=92), and University of Cincinnati (n=95), which also served as the lead administrative site.

Subjects

Potential subjects were identified from inpatient and outpatient clinical venues at each site and had been clinically diagnosed with either significant affective or psychotic symptoms. Subjects met the following criteria:

1. Subjects were aged 16 to 45 years.
2. Subjects exhibited significant affective symptoms, defined as a score of 17 or greater on the Montgomery-Asberg Depression Rating Scale[16] or a score of 12 or greater on the Young Mania Rating Scale,[17] or exhibited psychotic symptoms defined as ratings of 2 or greater on 1 or more of the Scale for the Assessment of Positive Symptoms[18] global score items.
3. Subjects exhibited no history of mental retardation or IQ of less than 70.
4. Subjects were able to complete a structured psychiatric interview.
5. Subjects exhibited psychiatric symptoms that were not secondary to medical illness as determined by medical evaluation and the absence of major medical or neurologic illnesses.
6. Subjects exhibited psychiatric symptoms that were not secondary to acute drug or alcohol intoxication or withdrawal.
7. Subjects were English speaking and at least second-generation residents of the United States.
8. Subjects were racially self-identified as African American or white and ethnically as Latino or non-Latino. The study was restricted to these large ethnic and racial categories to specifically test the primary hypothesis and maximize statistical power.
9. Subjects provided written informed consent that was approved by the institutional review boards at each site.

Procedure and Measures

Once informed consent was obtained, a research interviewer completed a semistructured evaluation of affective and psychotic symptoms using the Montgomery-Asberg Depression Rating Scale,[16] the Young Mania Rating Scale,[17] and the Scale for the Assessment of Positive Symptoms.[18] Each symptom was rated according to the maximal severity during the current (index) episode by interviewing the subject and reviewing clinical records. These ratings ensured that subjects met inclusion criteria and provided symptom measures for analyses. These rating scales are widely used in psychiatric research, have face and construct validity, and have interrater reliability established in diverse settings including within our own research groups.[2,19,20]

The interviewer then completed the DIGS.[15] The interviewers had broad educational backgrounds to simulate the variability of real-life clinical venues including several doctoral-level clinicians. Interviewers were kept blind to the specific hypotheses of the study. The DIGS consists of modules to elicit the common symptoms and signs of specific psychiatric disorders and syndromes; some modules that were not relevant to the primary questions of the study were not completed. Interviewers ignored the stop rules in each diagnostic section to ensure that all symptoms were reviewed. In this study, *DSM-IV* diagnostic criteria were applied to the DIGS results to assign diagnoses.

Establishing and Maintaining Reliability Among Interviewers

To establish and maintain reliability of assessments across sites, the sites' principal investigators first worked with their interviewers to train them on the instruments. Then, all clinical interviewers participated in a joint 3-day training session in which each interviewer performed interviews in a group setting with live patients. Finally, each clinical interviewer rated 10 videotaped interviews and provided results to the University of Cincinnati, which approved him or her as a study interviewer based on performance relative to standards determined by the group for each video. Specifically, each case was reviewed by multiple investigators at the University of Cincinnati, who discussed the cases and assigned a correct diagnosis that was used for training. Interviewers were expected to get more than 80% agreement with these diagnoses to qualify for the study.

Overview of Statistical Analyses

A Pearson χ^2 test or 1-way analysis of variance was used to compare demographic categorical variables across the 3 groups: white (non-Latino), African American (non-Latino), and Latino (of any race; non-white Latino subjects were too infrequent to separately analyze). Percentage correct and κ statistics were also performed to assess diagnostic agreement among sources. Logistic regressions were used to test the major hypothesis that African American subjects would exhibit significantly higher rates of clinical diagnoses (unblinded) in the schizophrenia spectrum than would white subjects, even after controlling for confounding factors. The logistic regressions treated unblinded clinical diagnoses of schizophrenia spectrum disorder (yes or no) as the dependent variable, whereas the independent variables were blinded expert-consensus–defined affective illness (yes or no) as well as the covariates of race/ethnicity, sex, age, site, income level, and education. Regressions were conducted first using a narrow definition of schizophrenia (excluding schizoaffective disorder) and then a broad definition (including schizoaffective disorder). We had 80% power to detect an odds ratio (OR) of 1.26 or greater given this sample size. Power was 90% for ORs of 1.3 or greater.

Because subjects were nested within site, there was the potential of within-site correlations among observations. To ensure that this did not create a problem, we conducted 2 additional procedures repeating the logistic regression analyses using

Table 1. Demographics for African American, Non-Latino White, and Latino Subjects[a]

Demographic Variable	No. (%)		
	African American (n = 241)	White (Non-Latino) (n = 220)	Latino (Any Race) (n = 149)
Sex			
Male	103 (43)	103 (47)	62 (42)
Female	138 (57)	117 (53)	87 (58)
Annual income, $			
0-15 000	191 (79)	160 (73)	123 (82)
15 001-25 000	25 (10)	32 (14)	16 (11)
25 001-40 000	17 (7)	16 (7)	6 (4)
40 001-55 000	6 (3)	6 (3)	3 (2)
≥55 001	2 (1)	6 (3)	1 (1)
Education, mean (SD), y			
Less than high school	69 (28)	40 (18)	64 (43)
High school	91 (38)	63 (29)	46 (31)
Some college	74 (31)	105 (48)	31 (21)
Postgraduate	7 (3)	12 (5)	8 (5)
Age, mean (SD) [range], y	34 (8.1) [18-53]	33 (8.5) [18-51]	34 (8.0) [18-54]

[a]Significant difference among groups: $\chi^2_6 = 44.6$; $P = .001$.

Table 2. Diagnostic Agreement by Ethnic/Racial Group Between Various Diagnostic Assessments

Diagnostic Assessment	African American (n = 241)	White (Non-Latino) (n = 220)	Latino (Any Race) (n = 149)	All Subjects (N = 610)
κ For diagnoses				
Clinical vs structured interview	0.39	0.45	0.56	0.45
Structured vs transcript	0.56	0.67	0.54	0.59
Transcript vs blinded expert	0.46	0.45	0.52	0.48
Clinical vs blinded expert	0.41	0.38	0.55	0.44
Structured vs blinded expert	0.40	0.46	0.63	0.48
Independent experts				
Agreement, %	65	70	69	68
κ For expert 1 vs expert 2	0.56	0.60	0.60	0.59

robust sandwich estimators of the error[21] and conditional logistic regression, conditioning on site.[22] The results from these analyses did not differ appreciably from the simple logistic regression (eg, no predictor variables that were found to be statistically significant in the simple regression lost significance in either alternative analysis) nor did variables previously found not to be significant emerge as significant. Consequently, we present the simpler logistic analyses.

Finally, to compare the racial/ethnic groups on symptom scales, multiple regression with covariates and dummy variables for discrete variables (eg, race) was used. All analyses were conducted by the statistician (S.A.) using SAS software version 9.1 (SAS Institute) with the exception of those involving robust sandwich estimators and conditional logistic regression, which were conducted using Stata/MP12 (StataCorp).

RESULTS

DEMOGRAPHICS, DIAGNOSTIC AGREEMENT, AND ACCURACY OF GUESSING SUBJECT RACE/ETHNICITY

Demographic data are reported in **Table 1**. Demographic differences among the 3 ethnic groups were not significant with the exception of education ($\chi^2_6 = 44.6$; $P = .001$). **Table 2** presents results for the 4 sources of diagnostic data and shows agreement among blinded experts prior to their generating consensus diagnoses. Diagnostic agreement among all combinations was, in general, modest and did not differ significantly among groups. Importantly, experts guessed ethnicity (Latino vs non-Latino) and race of subjects at chance levels, as determined by κ statistics, both individually prior to consensus ratings as well as in consensus assessments (**Table 3**). These results suggest that ethnicity effects were not present in the expert assessments.

LOGISTIC REGRESSION RESULTS

The first logistic regression treated narrowly defined, unblinded clinical diagnoses of schizophrenia spectrum disorder (yes or no) as the dependent variable. The independent variables included blinded expert-consensus–defined serious affective illness (yes or no; major depression, bipolar disorder, or schizoaffective disorder) as well as the covariates of race/ethnicity, sex, age, site, income level, and education. The results of the regression found the race effect to be significant ($\chi^2_2 = 10.4$; $P = .006$) as well as blinded expert-consensus–defined affective illness ($\chi^2_1 = 34.6$;

Table 3. Accuracy of Experts' Guessing Subject Ethnicity/Race For All Subjects[a]

Variable	Ethnicity, % Correct (κ)	Race, % Correct (κ)
Expert 1 (prior to consensus)	71 (0.17)	51 (0.11)
Expert 2 (prior to consensus)	69 (0.14)	51 (0.13)
Consensus of experts	73 (0.16)	59 (0.19)

[a] The κ values were not statistically significant.

$P=.001$). Additionally, a diagnosis in the schizophrenia spectrum was associated with lower income ($\chi^2_2=4.7$; $P=.03$), male sex ($\chi^2_2=4.6$; $P=.03$), and site ($\chi^2_5=24.4$; $P=.001$) with higher rates at the University of Michigan and University of California at Los Angeles. The full set of results presented in **Table 4** shows ORs and other statistics for each predictor variable. The pattern of results in Table 4 supported the primary predictions; namely, the OR comparing rates of schizophrenia-spectrum clinical diagnoses among African American with non-Latino white subjects was significant as predicted ($\chi^2_1=6.7$; $P=.01$; OR, 2.7; 95% CI, 1.5-5.1), controlling for all other factors. The OR comparing Latino with white subjects in an exploratory analysis was not ($\chi^2_1=1.9$; $P=.16$).

A similar but nonsignificant pattern held when unblinded diagnoses made by DIGS interviewers were treated as the dependent variable in logistic regression using the same set of independent variables as in Table 4. Namely, comparing African American with white subjects, an OR of 2.3 was not significant ($\chi^2_1=3.5$; $P=.06$); the OR comparing Latino with white subjects was also not significant ($\chi^2_1=0.9$; $P=.34$). However, the ethnicity-related patterns held, this time significantly so ($\chi^2_2=9.1$; $P=.01$) when unblinded diagnoses made from transcripts of the DIGS interviews were examined; the hypothesized contrast was obtained when comparing African American with non-Latino white subjects ($\chi^2_1=5.3$; $P=.02$; OR, 2.8; 95% CI, 1.4-5.6) but not Latino with white subjects ($\chi^2_1=1.3$; $P=.25$).

We then conducted a parallel set of logistic regressions, broadly defining schizophrenia to include schizoaffective disorder. Results for unblinded clinical diagnoses of schizophrenia are reported in **Table 5**. Significant differences in rates of clinical diagnoses of schizophrenia spectrum diagnoses, controlling for all variables in Table 5, were again observed between African American and white subjects ($\chi^2_1=7.5$; $P=.006$; OR, 2.5; 95% CI, 1.4-4.5) but not between Latino and white subjects ($\chi^2_1=3.1$; $P=.08$). In other regressions involving this broad definition of schizophrenia, significant effects comparing African American with white but not Latino with white subjects were observed for unblinded diagnoses made by DIGS interviewers ($\chi^2_1=12.0$; $P=.001$; OR, 2.9; 95% CI, 1.6-5.3) and for unblinded diagnoses based on interview transcripts ($\chi^2_1=20.4$; $P=.001$; OR, 4.2; 95% CI, 2.3-7.6).

CONSENSUS SYMPTOM RATINGS OF BLINDED EXPERTS BY ETHNICITY

Consensus symptom ratings on 5 scales (**Table 6**) were analyzed by ethnicity in multiple regressions with dummy variables for the covariates previously described. Multiple comparisons were controlled using the Holm method. Means, adjusted for all covariates, are shown in Table 6. Differences between African American and white subjects were not significant for affective symptoms but were for psychotic symptoms ($P<.05$).

Table 4. Demographic and Clinical Predictors of Clinical Diagnoses of Schizophrenia Narrowly Defined

Independent Variable	OR (95% CI)	Wald χ^2	*P* Value
Expert diagnosis[a]	0.23 (0.14-0.38)	34.60	.001
Ethnicity/race			
Non-Latino white	1 [Reference]		
African American	2.75 (1.49-5.10)	6.68	.01
Latino	1.62 (0.79-3.32)	0.01	.94
Sex			
Male	1.67 (1.05-2.68)	4.63	.03
Female	1 [Reference]		
Age	1.003 (0.975-1.033)	0.52	.82
Education			
Postgraduate	1 [Reference]		
Less than high school	1.83 (0.37-8.99)	0.03	.86
High school	2.14 (0.44-10.46)	1.56	.21
Some college	1.48 (0.30-7.30)	0.38	.54
Site			
UCLA	1 [Reference]		
University of Cincinnati	0.48 (0.22-1.04)	0.62	.43
University of Michigan	0.52 (0.25-1.09)	1.36	.24
UMDNJ	0.24 (0.11-0.57)	1.77	.18
University of Texas	0.35 (0.16-0.81)	0.05	.82
Howard University	0.14 (0.06-0.36)	7.36	.01
Annual income, $			
<15 000	1 [Reference]		
≥15 000	0.61 (0.39-0.96)	4.67	.03

Abbreviations: OR, odds ratio (from logistic regression); UCLA, University of California at Los Angeles; UMDNJ, University of Medicine and Dentistry of New Jersey.

[a] Binary diagnosis: major depression, bipolar disorder, and schizoaffective disorder coded as 1 and all others coded as 0.

COMMENT

Supporting our predictions, we found that African American subjects exhibited significantly higher rates (OR, 2.7) of clinical diagnoses in the schizophrenia spectrum than did non-Latino white subjects, even after controlling for the presence or absence of serious affective disorder and other clinical and demographic confounds. Effects were robust in that they persisted whether the schizophrenia spectrum did or did not include schizoaffective disorder and across different types of assessments. Despite these diagnostic differences, African American and white subjects did not differ significantly in blinded expert ratings of affective symptoms, but African American subjects did receive higher ratings of psychosis. These observations suggest that in African American subjects, psychotic symptoms may be overvalued by clinicians, skewing diagnoses toward schizophrenia-spectrum conditions, even with similar levels of affective symptoms as white subjects. The exploratory analysis suggests that this effect was specific to African American subjects and did not extend to Latino subjects.

Table 5. Demographic and Clinical Predictors of Clinical Diagnoses of Schizophrenia Broadly Defined

Independent Variable	OR (95% CI)	Wald χ^2	*P* Value
Expert diagnosis[a]	0.06 (0.04-0.10)	124.97	.001
Ethnicity/race			
Non-Latino white	1 [Reference]		
African American	2.51 (1.41-4.47)	7.46	.01
Latino	1.34 (0.71-2.55)	0.28	.59
Sex			
Male	1.32 (0.86-2.03)	1.66	.20
Female	1 [Reference]		
Age	0.99 (0.96-1.01)	0.86	.35
Education			
Postgraduate	1 [Reference]		
Less than high school	1.35 (0.32-5.67)	0.30	.58
High school	1.97 (0.47-8.23)	0.99	.32
Some college	2.18 (0.52-9.14)	1.92	.17
Site			
UCLA	1 [Reference]		
University of Cincinnati	0.82 (0.39-1.73)	0.01	.91
University of Michigan	1.35 (0.65-2.82)	3.26	.07
UMDNJ	0.64 (0.30-1.33)	1.12	.29
University of Texas	1.27 (0.61-2.64)	2.04	.15
Howard University	0.42 (0.19-0.92)	6.09	.01
Annual income, $			
<15 000	1 [Reference]		
≥15 000	0.73 (0.54-0.99)	3.93	.05

Abbreviations: OR, odds ratio (from logistic regression); UCLA, University of California at Los Angeles; UMDNJ, University of Medicine and Dentistry of New Jersey.

[a] Binary diagnosis: major depression and bipolar disorder coded as 1 and all others coded as 0.

For decades, higher rates of schizophrenia among African American vs white subjects in clinical settings have been reported in the medical literature, despite the lack of corresponding epidemiologic differences by race and ethnicity in the community. Our findings suggest that these differences persist even after controlling for multiple factors. Results also suggest that the expression of hallucinations and delusions by African American subjects in clinical venues may have preempted a more comprehensive review of diagnoses other than schizophrenia, even though schizophrenia is intended to be treated as a diagnosis of exclusion by *DSM-IV*. Therefore, in African American subjects, clinicians appeared to minimize the possibility of mood disorder diagnoses or failed to carefully apply the diagnostic criteria for these disorders. These effects occurred across all 6 US sites. Why African American subjects with mood disorders present with higher levels of psychosis or symptoms that are interpreted as psychotic cannot be answered by this study, but it may reflect cultural differences in worldview based on previous discriminatory experiences and reactions to them, ie, healthy paranoia, as well as cultural mistrust, cultural differences in expressing illness, or delayed treatment seeking leading to more severe illness at the time of clinical presentation.

The transcript editing process suggested that factors that are ambiguous such as idioms of distress including intense ruminations, complex linguistic patterns (eg, modes of expressing mood and cognitive and behavioral information) and personal experiences may be taken as evidence of psychosis during clinical evaluation. Given the complexity of these interactions, structured diagnostic interviews based on established diagnostic criteria may be insensitive to cultural variation; consequently, future studies might examine whether factoring ethnically sensitive aspects of symptom expression into diagnostic criteria might improve the clinical assessment process.

So why does it matter whether a subject with a psychotic mood disorder is incorrectly diagnosed with schizophrenia? Although sharing many symptoms, these 2 diagnostic groups warrant different treatments; in particular, affective illness is best treated by including thymoleptics (ie, antidepressants or mood stabilizers) and different types of therapy. Failure to do so may worsen treatment response and course of illness. Moreover, the prognosis of psychotic affective illness is typically more positive than that of schizophrenia; consequently, failure to recognize the former may lower treatment expectations, leading to inadequate breadth and depth of therapeutic interventions. Correctly diagnosing these complicated cases is the first step toward maximizing outcome. Moreover, bias in the assessment of diagnosis may reflect a more general bias in how different patient groups are managed, identified, and treated, contributing to an overall racial or ethnic disparity in health care for these subjects extending beyond the initial clinical assessment.

Notably, agreement across all combinations of diagnostic processes was modest, suggesting that it is difficult for clinicians to consistently interpret symptoms within the context of current diagnostic criteria independent of the approach. This finding is not unique to this study by any means and represents an ongoing challenge to the field of psychiatry. Although changes in diagnostic criteria might improve these processes, it is likely that significant improvement in diagnostic agreement of psychiatric patients will not occur until objective (eg, biological) markers for specific conditions are identified.

Several limitations must be considered when interpreting these results. Although ethnicity- and racially blinded medical records and transcripts control for these factors with the expert consensus review, these experts did not actually see the subjects, thereby losing potentially important visual data. However, there is no *a priori* evidence to expect differential effects of the loss of these data on diagnoses among the ethnic groups. Although identification of race and ethnicity was limited to respondents' self-reports, this information was verified by interviewers and reports of parental ethnicity. Moreover, there is considerable heterogeneity in the 3 broadly designated ethnic/racial groups studied, and we did not attempt to distinguish among ethnic subgroups because of insufficient statistical power even in this large data set. Some ethnic subgroups may therefore exhibit different patterns than observed for these more general designations. Also, although we included a broad geographic representation among our sites, these sites are academically affiliated settings generally serving urban, low-income populations, which may not be representative of all specialty care clinical venues. On the other hand, these particular academic sites may be more sensitive to the

Table 6. Adjusted Means Among African American, Non-Latino White, and Latino Subjects on 5 Symptom Scales According to Consensus of Blinded-expert Raters

	Adjusted Means (SD)			
Symptom Scale	African American (n = 241)	White (Non-Latino) (n = 220)	Latino (Any Race) (n = 149)	Significant Differences
MADRS, depression	23 (11)	26 (12)	26 (14)	
YMRS, mania	20 (14)	19 (12)	16 (13)	African American, non-Latino white>Latino
SAPS, total psychosis	1.8 (1.4)	1.4 (1.3)	1.3 (1.3)	African American>non-Latino white, Latino
SAPS, bizarre behavior	1.7 (1.6)	1.4 (1.5)	1.2 (1.5)	African American>Latino
SAPS, hall/del	2.0 (1.7)	1.4 (1.6)	1.3 (1.7)	African American>non-Latino white, Latino

Abbreviations: Del, delusion; Hall, hallucination; MADRS, Montgomery-Asberg Depression Rating Scale; SAPS, Scale for the Assessment of Positive Symptoms; YMRS, Young Mania Rating Scale.

potential impact of race and ethnicity on clinical assessment; therefore, any magnitude of differences observed herein may be relatively smaller than would be found more generally.

Despite limitations, this study is the first and largest to our knowledge to blind investigators to subject ethnicity and race to examine how these factors impact clinical assessment in mental health care. The results strongly support the need for a national approach to enhance quality of mental health care by eliminating disparities in psychiatric diagnoses that occur systematically in this vulnerable population. Moreover, these results remind clinicians to consistently challenge their own diagnostic assessments particularly in patients from other ethnic groups or in those who are failing to respond to treatment. Careful reconsideration of the criteria underlying diagnoses in all patients and examining each patient for these criteria over time may help minimize racial disparities in psychiatric practice.

Submitted for Publication: June 27, 2011; final revision received October 14, 2011; accepted December 2, 2011.
Published Online: February 6, 2012. doi:10.1001/archgenpsychiatry.2011.2040
Author Affiliations: University of Medicine and Dentistry of New Jersey University Behavioral HealthCare and Robert Wood Johnson Medical School, Piscataway, New Jersey (Dr Gara); University of Southern California (Dr Vega) and Harbor–University of California at Los Angeles Medical Center and Los Angeles Biomedical Research Institute (Dr Lesser), Los Angeles; University of Iowa, Iowa City (Dr Arndt); Paul L. Foster School of Medicine, Texas Tech University Health Science Center, El Paso (Dr Escamilla); University of Cincinnati College of Medicine, Cincinnati, Ohio (Drs Fleck and Strakowski); Howard University, Washington, DC (Dr Lawson); University of Michigan, Ann Arbor (Dr Neighbors); and University of Florida Health Science Center, Jacksonville (Dr Wilson).
Correspondence: Michael A. Gara, PhD, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, 671 Hoes Lane, Piscataway, NJ 08855 (garami@umdnj.edu).
Author Contributions: Drs Strakowski, Gara, and Arndt had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Financial Disclosure: Dr Strakowski has received research support in the form of grants to the University of Cincinnati Academic Health Center from Eli Lilly, Janssen Pharmaceuticals, AstraZeneca, Martek Biosciences, Nutrition 21, Repligen, the National Institute on Drug Abuse, National Institute on Alcohol Abuse and Alcoholism, National Institute of Mental Health, and National Alliance for Research on Schizophrenia and Depression. He has also chaired a symposium for Consensus Medical Communications and mentored the young investigator meeting at the American Psychiatric Association. Dr Wilson has received grant and research support from the National Institute of Mental Health, Dianippon, Johnson & Johnson, Lilly Roche, Bristol-Meyers Squibb, GlaxoSmithKline, Pfizer, AstraZeneca, Wyeth, and Quintiles.
Funding/Support: The parent study, Ethnicity and the Diagnosis of Affective Illness, and each site were funded by the National Institute of Mental Health from July 1, 2005, to June 30, 2009. Grant numbers by site are as follows: MH068801, University of Cincinnati; MH068795, Harbor–University of California at Los Angeles Medical Center; MH068819, University of Medicine and Dentistry of New Jersey; MH068797, Howard University; MH068804, University of Michigan; and MH068807, University of Texas at San Antonio.
Role of the Sponsors: The National Institute of Mental Health, which funded this study, played no role in the design and conduct of the study, data collection, management, analysis, review, or approval of the manuscript.
Additional Information: This article is dedicated to Michael W. Smith, MD, and Victor R. Adebimpe, MD. Dr Smith was the original principal investigator at the University of California at Los Angeles site but died in 2006. Dr Adebimpe was a consultant on this project and a trailblazer in the study of ethnicity effects on diagnoses in mental illness. He died in 2007. Both gentlemen have been sorely missed.
Additional Contributions: We acknowledge The Effects of Ethnicity on Diagnostic Assessment in Affective Illness Study group, which was a 6-site collaboration funded by the National Institute of Mental Health. The 6 sites, principal investigators, and coinvestigators were

Harbor–University of California at Los Angeles Medical Center and the Los Angeles Biomedical Research Institute, Torrance, (Lesser, principal investigator; Benjamin Furst, MD; Deborah Flores, MD; and Michael W. Smith); University of Texas Health Science Center at San Antonio (Dr Escamilla, principal investigator; Salvador Contreras, MD; Mercedes Ramirez, MD; and Juan Zavala, MD); Howard University, Washington, DC (Dr Lawson, principal investigator; and Tonya L. Seward, MA); University of Michigan, Ann Arbor (Dr Neighbors; Cheryl Munday, PhD; and Jamie Abelson, MSW); University of Medicine and Dentistry of New Jersey University Behavioral HealthCare and Robert Wood Johnson Medical School (Drs Gara and Vega, co–principal investigators; Lesley A. Allen, PhD; Humberto Marin, MD; and Theresa Miskimen, MD); and the University of Cincinnati, Ohio, under the direction of Dr Strakowski (overall study principal investigator), which functioned as the lead administrative site and data management storage center. Also participating at the University of Cincinnati were Lesley M. Arnold, MD; Michael A. Cerullo, MD; Dr Fleck; Quinton Moss, MD; and Henry Nasrallah, MD. Dr Arnold deserves special acknowledgment for serving as an expert reviewer of all edited documents (Figure 1) and for her participation in multiple other aspects of the project. Dr Wilson and Sriram Ramaswamy, MD, participated from Creighton University. Dr Munday is now at the University of Detroit–Mercy.

REFERENCES

1. Williams DR, Earl TR. Commentary: race and mental health: more questions than answers. *Int J Epidemiol*. 2007;36(4):758-760.
2. Arnold LM, Keck PE Jr, Collins J, Wilson R, Fleck DE, Corey KB, Amicone J, Adebimpe VR, Strakowski SM. Ethnicity and first-rank symptoms in patients with psychosis. *Schizophr Res*. 2004;67(2-3):207-212.
3. Strakowski SM, Hawkins JM, Keck PE Jr, McElroy SL, West SA, Bourne ML, Sax KW, Tugrul KC. The effects of race and information variance on disagreement between psychiatric emergency service and research diagnoses in first-episode psychosis. *J Clin Psychiatry*. 1997;58(10):457-463, quiz 464-465.
4. Strakowski SM, Shelton RC, Kolbrener ML. The effects of race and comorbidity on clinical diagnosis in patients with psychosis. *J Clin Psychiatry*. 1993;54 (3):96-102.
5. Bresnahan M, Begg MD, Brown A, Schaefer C, Sohler N, Insel B, Vella L, Susser E. Race and risk of schizophrenia in a US birth cohort: another example of health disparity? *Int J Epidemiol*. 2007;36(4):751-758.
6. Chung H, Mahler JC, Kakuma T. Racial differences in treatment of psychiatric inpatients. *Psychiatr Serv*. 1995;46(6):586-591.
7. Abebimbe VR. A second opinion on the use of white norms in psychiatric diagnosis of black patients. *Psychiatr Ann*. 2004;34:543-551.
8. McGrath J, Saha S, Chant D, Welham J. Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiol Rev*. 2008;30:67-76.
9. Robins LN, Regier DA. *Psychiatric Disorders in America*. New York, NY: The Free Press; 1991.
10. Vega WA, Lewis-Fernández R. Ethnicity and variability of psychotic symptoms. *Curr Psychiatry Rep*. 2008;10(3):223-228.
11. Walker E, Kestler L, Bollini A, Hochman KM. Schizophrenia: etiology and course. *Annu Rev Psychol*. 2004;55:401-430.
12. Kleinman A, Eisenberg L, Good B. Culture, illness, and care: clinical lessons from anthropologic and cross-cultural research. *Ann Intern Med*. 1978;88(2):251-258.
13. Strakowski SM, Keck PE Jr, Arnold LM, Collins J, Wilson RM, Fleck DE, Corey KB, Amicone J, Adebimpe VR. Ethnicity and diagnosis in patients with affective disorders. *J Clin Psychiatry*. 2003;64(7):747-754.
14. Neighbors HW, Trierweiler SJ, Ford BC, Muroff JR. Racial differences in *DSM* diagnosis using a semi-structured instrument: the importance of clinical judgment in the diagnosis of African Americans. *J Health Soc Behav*. 2003;44(3): 237-256.
15. *Diagnostic Interview for Genetic Studies version 4.0*. Rockville, MD: National Institute of Mental Health, 2004.
16. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134:382-389.
17. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry*. 1978;133:429-435.
18. Andreasen NC. *The Scale for the Assessment of Positive Symptoms (SAPS)*. Iowa City: University of Iowa; 1984.
19. Strakowski SM, Flaum M, Amador X, Bracha HS, Pandurangi AK, Robinson D, Tohen M. Racial differences in the diagnosis of psychosis. *Schizophr Res*. 1996; 21(2):117-124.
20. Strakowski SM, McElroy SL, Keck PE Jr, West SA. Racial influence on diagnosis in psychotic mania. *J Affect Disord*. 1996;39(2):157-162.
21. Huber PJ. *The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. Vol 1. Berkeley: University of California Press, 1967:221-233.
22. Hosmer DW Jr, Lemeshow S. *Applied Logistic Regression*. 2nd ed. New York: Wiley; 2000.