Ethnicity Effects on Clinical Diagnoses Compared to Best-Estimate Research Diagnoses in Patients With Psychosis: A Retrospective Medical Chart Review

Deidre M. Anglin, Ph.D., and Dolores Malaspina, M.D., M.P.H.

Objective: Ethnicity effects on diagnoses are frequently reported and have variably been attributed to diagnostic biases versus ethnic differences in environmental exposures, and other factors.

Method: We compared best-estimate gold standard research diagnoses to clinical diagnoses (DSM-III-R and DSM-IV criteria) among 129 white, 57 African American, and 50 Hispanic patients with psychosis admitted to an inpatient research unit from 1990 to 2003.

Results: Clinical and research diagnoses showed greater agreement in Hispanic than in African American patients (white patients were intermediate). Diagnostic agreement for paranoid schizophrenia was likewise the best in Hispanic patients. While paranoid schizophrenia tended to be overdiagnosed in African American patients, it was underdiagnosed in white patients. Patterns of diagnostic agreement for schizoaffective disorder and "other" diagnoses were similar among the 3 ethnic groups.

Conclusions: Diagnostic unreliability may explain the excess of paranoid schizophrenia reported for African Americans. Further research is needed to elucidate the influence of ethnicity on clinical diagnosis before other theories to explain group differences can be reasonably proposed and reliably tested.

(J Clin Psychiatry 2008;69:941–945)

Received Jan. 15, 2007; accepted Oct. 25, 2007. From the Department of Psychiatry, Columbia University, New York State Psychiatric Institute (Dr. Anglin) and the Department of Psychiatry, New York University School of Medicine (Dr. Malaspina), New York, N.Y.

This research was supported by National Institute of Mental Health grant 2K24 MH01699 (Dr. Malaspina) and a Columbia University Department of Psychiatry Frontier Fund (Dr. Anglin).

The authors report no additional financial or other relationships relevant to the subject of this article.

Corresponding author and reprints: Deidre M. Anglin, Ph.D., 100 Haven Ave., Tower 3, Room 31F, New York, NY 10032 (e-mail: dma2105@columbia.edu).

Treated prevalence studies consistently find that African Americans are more likely than whites to be diagnosed with schizophrenia[1–3] and less likely to receive psychotic affective and bipolar diagnoses.[4] For example, although African American patients were less likely to self-report psychotic symptoms, they were more likely to be diagnosed with schizophrenia in a large sample of 19,219 inpatients and outpatients from a behavioral health system in New Jersey.[5] Conversely, Latinos in this sample were less likely to be diagnosed with schizophrenia and more likely to receive affective disorder diagnoses, despite reporting more psychotic symptoms.[5] Explanations for such discrepancies include biases and lack of cultural awareness by clinicians or the differential reporting of symptoms by ethnic subgroups, if there are no actual group differences, or racial/ethnic differences in genetic or environmental factors that influence the risks for psychiatric disorders, if such reports are valid.

Diagnostic practices could conceivably explain the group differences. Indeed, the findings of Trierweiler and colleagues[6–8] suggest that even when standardized diagnostic criteria are used, clinical judgment and clinician characteristics play a differential role in how symptoms are attributed to African American and white patients. In their analyses, Strakowski and colleagues find higher frequencies of schizophrenia among African American patients compared to whites, despite similar rates of affective symptoms,[9] and specifically psychotic mania.[10]

In comparison to the above described clinical diagnoses that varied with ethnicity, epidemiologic community research studies generally find no differences in schizophrenia diagnoses by ethnicity in controlled analyses.[11,12] Thus, various studies suggest that racial biases exist in the diagnostic process.[13] Questions concerning diagnostic accuracy must be resolved before genetic and environmental theories for these effects can be examined, since testing these hypotheses will rest on diagnostic validity and reliability.

One approach to investigating the possibility of clinician diagnostic errors is to compare diagnoses generated by treating clinicians with best-estimate research diagnoses that are generated during the same admission. Best-estimate research diagnostic procedures have greater diagnostic accuracy than clinical diagnoses[14] and are appropriately used as "gold standard" assessments for research

purposes.[15,16] Discordance between clinical diagnoses and best-estimate research diagnoses may indicate areas of inaccuracy for clinician-generated diagnoses. Previous studies find lower diagnostic agreement between the hospital diagnosis and the research diagnosis for African American patients compared to whites.[17,18] Furthermore, the pattern of disagreement suggests that clinicians often failed to identify affective symptoms in African American patients.[9] Sohler and Bromet[19] also found that African American patients had increased odds of receiving a discordant diagnosis compared to white patients; however, the magnitude of the effect was reduced in controlled analyses. To our knowledge, few findings have explicitly examined concordance between clinical and research diagnoses among *both* Hispanic patients and African American patients (but see Lawson et al.[20]).

We examined the agreement between clinical and best-estimate research diagnoses by race/ethnicity and further probed the sensitivity and specificity of psychotic diagnoses within racial/ethnic groups. We hypothesized that African American patients would receive the most discordant diagnoses and white patients the least. We hypothesized that Hispanic patients will also have lower concordance than white patients.

METHOD

The study involves a retrospective medical chart review of inpatients admitted to the Schizophrenia Research Unit (SRU) at the New York State Psychiatric Institute from 1990 to 2003. The SRU is a 12-bed inpatient unit that is part of a research program that provides inpatient evaluation and clinical treatment for patients with severe mental illness. The average length of stay on the unit is 3 to 6 months, and during this time, patients may participate in any number of research studies designed to study diagnostic, neurologic, biochemical, physiologic, and psychosocial aspects of severe mental illness. Patients who met the study criteria and provided written informed consent for the study were evaluated by the research team (approximately 64% of the admitted cases). They participated in research studies under a National Institute for Mental Health Clinical Research Center Grant.[21] Typically, the patients screened into the SRU were English speaking or bilingual (Spanish), did not have primary active substance abuse problems or significant histories of violence, and had a psychotic condition. The present study is focused on the comparison of clinical hospital diagnoses made by attending psychiatrists, or by psychiatry residents under their supervision, with the best-estimate consensus diagnoses made using research assessment data and clinical data.

Procedure

A comprehensive list of all patients evaluated on the SRU receiving a best-estimate consensus diagnosis from 1990 to 2003 was obtained from the computerized database. The medical charts available at New York State Psychiatric Institute were matched to this list using unique patient ID numbers and pulled. The total number of matched patients was 267. The *clinical hospital diagnoses* obtained from the medical charts were made with DSM-III-R or DSM-IV criteria using typical clinical data obtained in hospital settings (i.e., patient report of current symptoms, past psychiatric records obtained from previous hospitalizations at other hospitals in New York City, and any accompanying family or friends' report of patients' functioning). Patients who participate in the research protocols also participate in *best-estimate consensus diagnostic procedures*. These diagnoses are largely based on information from face-to-face structured psychiatric interviews with the Diagnostic Interview for Genetic Studies (DIGS).[22] The DIGS is used to assess lifetime and current psychiatric diagnoses. Its interrater reliability is $\kappa = 0.95$ for DSM-IV diagnosis and $\kappa = 0.80$ for individual symptoms.[22] The consensus diagnosis is typically ascertained using these data, as well as admission information on age, gender, education, ethnicity, the age at onset of positive symptoms, the patient's age at first treatment, global assessments of functioning for the worst period in the current episode and for the last month, past psychiatric records, and family interviews when possible. Typically, the best-estimate consensus diagnosis was made in a meeting that included the unit chief from the clinical team and the diagnostic experts of the research team. With a few exceptions, most of the diagnosticians on the SRU have been white, but since 1998, Hispanic representation among the team has increased.

This secondary analysis study was approved by the human subjects committee and all patients provide written informed consent to participate in research studies. The treating clinician's primary discharge diagnosis was obtained from the discharge summary note of each chart located in medical records, and the best-estimate research diagnosis was obtained from the research database maintained on all patients admitted to the SRU. Patients' race/ethnicity was obtained from this research database. Medical records were linked with the research database using unique patient ID numbers given to all SRU patients.

Data Analyses

The κ statistic was used to determine diagnostic agreement between the clinical hospital diagnosis and the best-estimate research diagnosis. Chi square tests were used to test whether the degree of agreement differed significantly between African American, Hispanic, and white patients. Sensitivity and specificity of the clinical diagnoses of psychotic disorders were estimated using the best-estimate consensus research diagnosis as the gold-standard, and compared in each racial/ethnic group. Finally, we conducted logistic regression analyses to

Table 1. Demographic and Clinical Characteristics of a Sample of 238 White, African American, and Hispanic Inpatients With Psychosis

Characteristic	White (N = 130)	African American (N = 58)	Hispanic (N = 50)	Analysis		
				F or χ^2	df	p Value
Years of education, mean (SD)	13.6 (2.7)	12.3 (2.3)	12.1 (3.2)	F = 7.13	2	.001
Age, mean (SD), y	34.5 (10.2)	29.0 (9.1)	31.9 (10.0)	F = 6.40	2	.002
Age at onset, mean (SD), y	20.7 (5.9)	21.3 (5.5)	21.7 (8.2)	F = 0.416	2	.660
Age at first treatment, mean (SD), y	19.3 (7.4)	21.6 (7.3)	22.9 (7.8)	F = 4.78	2	.009
GAF at admission, mean (SD)	37.3 (11.3)	35.1 (9.0)	36.1 (10.1)	F = 0.943	2	.391
GAF at discharge, mean (SD)	48.7 (13.7)	52.1 (11.6)	54.2 (12.0)	F = 3.67	2	.027
Gender, male, N (%)	83 (63.8)	35 (60.3)	28 (56.0)	χ^2 = 0.97	2	.616

Abbreviation: GAF = Global Assessment of Functioning.

Table 2. Comparison of the Frequency of Clinical and Best-Estimate Research Diagnoses in a Sample of 236 Inpatients With Psychosis by Racial/Ethnic Group[a]

Diagnostic Category	Clinical Diagnoses, N	Research Diagnoses, N	Sensitivity	Specificity	Overall κ
White	129	129			0.488
Schizoaffective	38	35	0.743	0.872	
Schizophrenia-paranoid	30	29	0.517	0.850	
Schizophrenia-other	50	55	0.673	0.824	
Other	11	10	0.500	0.950	
African American	57	57			0.362
Schizoaffective	6	5	0.800	0.962	
Schizophrenia-paranoid	27	24	0.625	0.636	
Schizophrenia-other	17	23	0.522	0.853	
Other	7	5	0.400	0.904	
Hispanic	50	50			0.599
Schizoaffective	17	17	0.824	0.909	
Schizophrenia-paranoid	17	17	0.706	0.849	
Schizophrenia-other	12	14	0.643	0.917	
Other	4	2	0.500	0.938	

[a]Analyses based on patients with both research and clinical diagnoses.

determine whether relevant demographic and clinical variables significantly influenced the relationship between race/ethnicity and diagnostic discordance.

RESULTS

The original study sample consisted of 267 ethnically diverse patients: 51% (N = 135) were white; 22% (N = 59) were African American; 20% (N = 54) were Hispanic; 6% (N = 17) were Asian/Pacific Islander; and 1% (N = 2) had "other" ethnicities. The present study focused on the white, African American, and Hispanic patients (N = 248). While there are 7 diagnostic categories represented among the consensus research diagnoses, 3 were excluded in the present study due to too small numbers of cases represented in each category (i.e., schizophrenia-catatonic, schizophrenia-residual, and substance abuse). The remaining diagnostic groupings are schizoaffective (N = 61), schizophrenia-paranoid (N = 74), schizophrenia-other (N = 79), and other (N = 22). The schizoaffective disorder category includes depressed and manic subtypes. The schizophrenia-other category includes both undifferentiated and disorganized subtypes. The most frequent diagnosis in the "other" psychotic disorder category was psychosis not otherwise specified (NOS), but this category also includes a variety of other diagnoses including major depression with psychotic features, depressive disorder NOS, and personality disorder.

Table 1 shows the demographic and clinical data for each racial/ethnic group. As indicated, whites were significantly older and more educated than Hispanics and African Americans. The gender distribution of the patient sample was similar across racial/ethnic groups. Clinically, there was no significant racial/ethnic difference on mean age at onset of symptoms; however, whites on average entered treatment at significantly younger ages than Hispanics. There was no significant difference in the Global Assessment of Functioning (GAF) mean score upon admission between the racial/ethnic groups; however, on average, Hispanics were discharged with a significantly higher GAF score than whites.

Table 2 shows the distribution of the 4 clinical hospital discharge and research diagnostic categories for white, African American, and Hispanic patients. The distribution was significantly different for both clinical hospital diagnoses ($\chi^2 = 17.18$, df = 6, $p < .01$) and best-estimate research diagnoses ($\chi^2 = 16.44$, df = 6, $p < .05$). With regard to clinical hospital diagnoses, while 34.0% of Hispanics (N = 17) and 29.5% of whites (N = 38) received a diagnosis of schizoaffective disorder, only 10.5% of

African Americans (N = 6) received this diagnosis (χ^2 = 9.58, df = 2, p < .01). The frequency of paranoid schizophrenia was significantly different by group (χ^2 = 9.76, df = 2, p < .01). While only a quarter of whites (23.3%, N = 30) and a third of Hispanics (34.0%, N = 17) were considered to have paranoid schizophrenia, almost half of African Americans (47.4%, N = 27) were diagnosed with paranoid schizophrenia. Likewise, with regard to best-estimate research diagnoses, the racial/ethnic groups differed in their frequency of schizoaffective disorder (χ^2 = 10.5, df = 2, p < .01) and paranoid schizophrenia (χ^2 = 7.67, df = 2, p < .05) (Table 2).

Overall agreement between the research and clinical diagnoses was moderate across the 4 diagnostic groups (i.e., schizoaffective disorder, schizophrenia-paranoid, schizophrenia-other, and "other" disorders), κ = 0.49 (95% CI = 0.41 to 0.57). Overall diagnostic agreement was higher among Hispanic patients than among African American patients (χ^2 = 3.13, df = 2, p = .06). While 24 African Americans (42.1%) had discordant clinical hospital and best-estimate research diagnoses, only 14 Hispanics (28.0%) had discrepant diagnoses. Forty-six whites (35.7%) had discordant diagnoses.

Using the best-estimate research diagnosis as the gold standard, sensitivity and specificity analyses were conducted and used as indicators of under- and overdiagnosis (See Table 2 for details). Percentages of diagnostic agreement for the schizoaffective and "other" diagnostic categories were similar across all 3 racial/ethnic groups. The majority of the discrepancies in clinical and research diagnoses for African American and white patients were due to the paranoid schizophrenia and schizophrenia-other categories. For example, of 29 white patients who received a research diagnosis of paranoid schizophrenia, only 51.7% (N = 15) also received this clinical diagnosis, indicating that whites were underdiagnosed with paranoid schizophrenia. In contrast, of 33 African American patients who did not receive a research diagnosis of paranoid schizophrenia, 36.4% (N = 12) received this clinical diagnosis; African Americans were overdiagnosed with paranoid schizophrenia. In terms of schizophrenia-other diagnoses, African Americans tended to be underdiagnosed; of 23 African American patients who received these research diagnoses, only 52.2% (N = 12) also received the same clinical diagnosis.

We used logistic regression to compute the odds ratio of receiving a discordant diagnosis in adjusted models. Given the significantly different κ values for Hispanic and African American groups, we focused on effects for these 2 groups. The odds ratio in the unadjusted model for African Americans versus Hispanics was 1.87 (95% CI = 0.83 to 4.21), and in adjusted models (i.e., controlling for mean age, level of education, and age at first treatment) was 2.18 (95% CI = 0.93 to 5.14), indicating a marginally significant effect.

DISCUSSION

The purpose of the present study was to compare diagnostic reliability in white, African American, and Hispanic patients with psychosis. We found diagnostic concordance was unexpectedly remarkable among Hispanic patients. Clinicians were most accurate in diagnosing psychotic conditions in this group, which had a higher accuracy than the groups of white and African American patients. The diagnostic agreement for the African American cases was below the accepted threshold of adequate reliability.[23] The kinds of diagnostic errors were related to the race and ethnicity of the patient. Clinicians overdiagnosed paranoid schizophrenia in African American cases, while they underdiagnosed paranoid schizophrenia in white cases. They also underdiagnosed undifferentiated or disorganized schizophrenia in African American patients, identifying such cases as having paranoid subtypes.

Our findings are consistent with previous studies that demonstrate a greater potential for clinician biases for African American patients.[24] It is not clear why clinicians diagnosed Hispanic patients more reliably given the greater potential for cultural and language differences between diagnosticians and Hispanic patients. One possibility is that the location of the research unit in a predominantly Latino community has sensitized the clinical staff to Latino diagnostic issues, which are emphasized in the clinical training. The Hispanic patients on this unit tended to be discharged with a higher GAF score than the African American and white patients, which was not evidenced upon admission. It is possible that the Hispanic patients improved the most because they had the correct diagnosis. Another possibility is the greater comfort of the Latino population in disclosing mental health issues. There may be less cultural mistrust among the Hispanics in this patient population than among the African American patients. This might be especially true given that after 1998, Hispanic diagnosticians and staff were better represented on the unit.

Limitations

One major limitation is that the base rate of diagnoses was not equally prevalent across ethnic groups. Thus, the relatively high percentage of diagnostic agreement among African Americans of the schizoaffective diagnosis could be misleading because very few African American patients were given that diagnosis by either the clinicians or the gold standard procedures. The small sample size also limits our ability to conduct time-sensitive analyses to assess the impact of the long event horizon under which the data were gathered. This finding needs to be replicated on a larger sample with more representation within each racial/ethnic group and among the diagnostic groupings. Notwithstanding, the present study highlights the importance of moving toward a truly multicultural paradigm in

psychiatric research, and away from making just African American/white comparisons. The increased reliability found among Hispanic patients needs to be replicated on a larger more representative sample.

It is unclear whether our findings would generalize to other hospital populations given that our patient sample was drawn from a specialized research unit. The fact that African American clinicians are not adequately represented on either the clinical or research diagnostic teams may have played a role in the low diagnostic reliability of African American patients. Indeed, Trierweiler et al.[6,7] find that the race of the clinician influences how symptoms are attributed to different racial/ethnic groups. This possibility should be empirically tested in future studies and compared to how well having Hispanic clinicians improves diagnostic reliability. After 1998, the SRU incorporated Hispanic Spanish-speaking diagnosticians and staff; thus it is possible that diverse representation among diagnosticians and staff improves diagnostic accuracy.

Conclusions and Implications

In sum, our results suggest the diagnostic process is particularly reliable for Hispanic psychotic patients in structured clinical settings. This reliability seems to be driven by the ability of clinicians to detect schizoaffective diagnoses consistently. For African American patients, the relatively lower diagnostic agreement seems to be driven by difficulties distinguishing paranoid schizophrenia from other psychotic disorders. Indeed, diagnostic unreliability may explain the excess of paranoid schizophrenia reported for African Americans. Our results suggest that more empirical attention needs to be placed on examining the diagnostic process including provider variables, as misclassification of disease by specific racial/ethnic groups can mask or bias relationships between race/ethnicity and psychiatric illness. A thorough analysis that identifies whether differential symptoms predict diagnostic discordance differently in Hispanic and African American patients will help clinicians and researchers understand the meaning behind the high and low reliability, respectively. Psychiatry, while a field of medicine, may be greatly impacted by cultural differences among patients. In our ever-changing multicultural society, it will be very important to tease apart diagnostic inconsistencies that are more prevalent among groups of color.

REFERENCES

1. Blow FC, Zeber JE, McCarthy JF, et al. Ethnicity and diagnostic patterns in veterans with psychoses. Soc Psychiatry Psychiatr Epidemiol 2004;39:841–851
2. Lawson WB, Hepler N, Holladay J, et al. Race as a factor in inpatient and outpatient admissions and diagnosis. Hosp Community Psychiatry 1994;45:72–74
3. Neighbors HW, Jackson JS, Campbell L, et al. The influence of racial factors on psychiatric diagnosis: a review and suggestions for research. Community Ment Health J 1989;25:301–311
4. Strakowski SM, Keck PE Jr, Arnold LM, et al. Ethnicity and diagnosis in patients with affective disorders. J Clin Psychiatry 2003;64:747–754
5. Minsky S, Vega W, Miskimen T, et al. Diagnostic patterns in Latino, African American, and European American psychiatric patients. Arch Gen Psychiatry 2003;60:637–644
6. Trierweiler SJ, Neighbors HW, Munday C, et al. Differences in patterns of symptom attribution in diagnosing schizophrenia between African American and non–African American clinicians. Am J Orthopsychiatry 2006;76:154–160
7. Trierweiler SJ, Muroff JR, Jackson JS, et al. Clinician race, situational attributions, and diagnoses of mood versus schizophrenia disorders. Cultur Divers Ethnic Minor Psychol 2005;11:351–364
8. Trierweiler SJ, Neighbors HW, Munday C, et al. Clinician attributions associated with the diagnosis of schizophrenia in African American and non–African American patients. J Consulting Clin Psychol 2000;68:171–175
9. Strakowski SM, Flaum M, Amador X, et al. Racial differences in the diagnosis of psychosis. Schizophr Res 1996;21:117–124
10. Strakowski SM, McElroy SL, Keck PE, et al. Racial influence on diagnosis in psychotic mania. J Affect Disord 1996;39:157–162
11. Robins L, Reiger DA. Psychiatric Disorders in America: The Epidemiologic Catchment Area Study. New York, NY: The Free Press; 1991
12. Kessler RC, Birnbaum H, Demler O, et al. The prevalence and correlates of nonaffective psychosis in the National Comorbidity Survey Replication (NCS-R). Biol Psychiatry 2005;58:668–676
13. Whaley AL. Cultural mistrust and clinical diagnosis of paranoid schizophrenia in African American patients. J Psychopath Behav Assess 2001;23:93–100
14. Kosten TA, Rounsaville BJ. Sensitivity of psychiatric diagnosis based on the best-estimate procedure. Am J Psychiatry 1992;149(9):1225–1227
15. Roy MA, Lanctot G, Merette C, et al. Clinical and methodological factors related to reliability of the best-estimate diagnostic procedure. Am J Psychiatry 1997;154(12):1726–1733
16. Weiser M, Kanyas K, Malaspina D, et al. Sensitivity of ICD-10 diagnosis of psychotic disorders in the Israeli National Hospitalization Registry compared with RDC diagnoses based on SADS-L. Compr Psychiatry 2005;46:38–42
17. Simon RJ, Fleiss JL, Gurland BJ, et al. Depression and schizophrenia in hospitalized black and white mental patients. Arch Gen Psychiatry 1973;28:509–512
18. Strakowski SM, Hawkins JM, Keck PE Jr, et al. The effects of race and information variance on disagreement between psychiatric emergency service and research diagnoses in first-episode psychosis. J Clin Psychiatry 1997;58:457–463
19. Sohler NL, Bromet EJ. Does racial bias influence psychiatric diagnoses assigned at first hospitalization? Soc Psychiatry Psychiatr Epidemiol 2003;38:463–472
20. Lawson WB, Herrera JM, Costa J. The dexamethasone suppression test as an adjunct in diagnosing depression. J Assoc Acad Minor Phys 1992;3:17–19
21. Malaspina D, Goetz RR, Yale S, et al. Relation of familial schizophrenia to negative symptoms but not to the deficit syndrome. Am J Psychiatry 2000;157(6):994–1003
22. Nurnberger JI, Blehar MC, Kaufmann CA, et al, and collaborators from the NIMH Genetics Initiative. Diagnostic interview for genetic studies: rationale, unique features, and training. Arch Gen Psychiatry 1994;51:849–859
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159–174
24. Whaley AL, Geller PA. Ethnic/racial differences in psychiatric disorders: a test of four hypotheses. Ethn Dis 2003;13(4):499–512