

## Diagnostic, Symptom, and Functional Assessments of Hispanic Outpatients in Community Mental Health Practice

Luis H. Zayas · Luis R. Torres · Leopoldo J. Cabassa

Received: 5 November 2007 / Accepted: 8 July 2008 / Published online: 22 July 2008  
© Springer Science+Business Media, LLC 2008

**Abstract** With increased US Hispanic diversity come diagnostic challenges associated with culture, language, and expression of mental disorders. In a community-based clinic, we compared diagnostic agreement between Hispanic and non-Hispanic clinicians, and a structured diagnostic instrument, in live and videotaped interviews with Hispanic adults. Percentage agreement and kappas show low diagnostic reliability between clinicians, and between clinicians and instrument. Significant differences appeared in rates of various diagnoses. Non-Hispanic clinicians rated patients' functional capacity and symptom-severity as significantly worse than Hispanic clinicians. Findings match past research with Hispanic patients and raise questions about diagnostic reliability in multi-cultural community mental health practice.

**Keywords** Psychiatric diagnosis · Diagnostic reliability · Hispanics

In an era of vast human migration around the world, psychiatric clinicians in industrialized countries, such as the United States, are called on to diagnose and treat patients from a heterogeneous range of ethnic, racial, cultural, social and linguistic backgrounds. Pronounced challenges

to diagnostic accuracy and agreement in clinical practice come with these demographic differences. The growth of the US Hispanic population—projected to rise from 35 million in 2000 to over 47 million in 2010, and up to 87 million in 2040 (US Census Bureau 2004)—represents one such challenge. Recent Hispanic immigrants are joining large established communities of Hispanics in the US, and together they face linguistic and cultural barriers in communicating with providers for their psychiatric needs. Some speak only Spanish, others are bilingual, and still others monolingual English speakers. There are also variations in their conceptions of mental illness, its assessment and treatments based on a diverse set of cultural beliefs and traditions (Cabassa et al. 2007). Communication problems will curtail access to care and prolong the under-utilization and premature termination of services (Alegria and McGuire 2003; Cabassa et al. 2006; Vega and Alegria 2001; Vega and Lopez 2001). To assure accuracy and consistency and to improve service delivery, our diagnostic strategies need examination.

Because of the makeup of our psychiatric labor force, Hispanic immigrants are very likely to be assigned non-Hispanic, English-speaking clinicians who, while wanting to provide high-quality services, may have difficulty communicating with and understanding patients' linguistic and cultural nuances. Ethnic similarity with patients may enhance clinicians' ability to identify cultural modes of expressing symptoms; comprehend meanings of experiences; and understand variations in thought and expression, but it does not guarantee that clinicians will agree on their diagnostic decisions (Malgady et al. 1987; Strakowski et al. 2003). Clinicians render significantly different diagnoses of patients who are ethnically similar or different from them, raising questions about diagnostic accuracy and reliability (i.e., agreement by different clinicians using a common

---

L. H. Zayas (✉)  
Washington University in St. Louis, Campus Box 1196,  
One Brookings Drive, St. Louis, MO 63130-4899, USA  
e-mail: lzayas@wustl.edu

L. R. Torres  
University of Houston, Houston, TX, USA

L. J. Cabassa  
University of Southern California, Los Angeles, CA, USA

diagnostic system; Langenbucher et al. 1996). Structured diagnostic instruments may not resolve this problem since they typically do not capture the more nuanced information that patients provide and clinicians elicit with good skills and observation of cultural features.

This pilot study was designed to explore agreement between clinicians of Hispanic and non-Hispanic background and between clinicians and a structured diagnostic measure, in interviews with Hispanic patients in an urban clinic. Our operating assumption is that in community mental health, patients' diagnoses should not differ dramatically from one clinician to another due to their or their clinicians' ethnicities or cultures. Furthermore, diagnoses by objective measures and clinicians should be relatively concordant. Patients should expect to get the same, accurate diagnosis regardless of the clinician they visit. Diagnosis should not be unduly influenced by race, culture or ethnicity, although it should be sensitive to their influences.

While this assumption may have some face validity, contemporary research on diagnostic practice does not seem to support it. Minsky et al. (2003), for example, found that African Americans are more likely to be diagnosed with schizophrenia, while Hispanics are more likely to be diagnosed with major depressive disorders rather than schizophrenia-spectrum disorders. Hispanics tend to self-report more psychotic symptoms than African Americans, and yet do not receive schizophrenia-spectrum diagnoses at the rates African Americans do. Strakowski et al. (1996) report that African American patients are significantly more likely than White patients to show severe psychotic symptoms and be diagnosed with schizophrenia, and less likely to be diagnosed with psychotic depression. The cultural gap between clinicians and patients may result "in assigning pathology where it [does] not belong" (Strakowski et al. 1996, p. 122) such as misinterpreting culturally based jargon as thought disorder. Diagnostic agreement between emergency room clinicians and a structured research interview is less common for non-White minority patients than White patients (Strakowski et al. 1997). In fact, overall agreement between emergency room clinicians and structured-interview clinicians occurred only in 42% of patients ( $\kappa = 0.25$ ). In sum, racial differences in diagnoses are not entirely eliminated by using structured clinical interviews (Neighbors et al. 2003).

Signs of mental illness may be assessed differently by racial and ethnic groups, and culture adds variations in how persons communicate symptoms (Arnold et al. 2004). Trierweiler et al. (2000) conclude that African Americans as a group may elicit different responses from clinicians than do non-African Americans. Diagnostic differences may be the result of clinicians' attributions of certain behaviors to particular groups because they weigh their

observations of behaviors differently based on patient ethnicity, race, or culture, and even gender and age (Alegría and McGuire 2003). Clinicians' assessments of clients' orientation, symptoms, judgment, and other mental functions will be affected by their interpretations of clients' verbal and nonverbal communication. Sociocultural theory frames the psychiatric diagnostic process as a series of cultural interpretations made by patients and clinicians. The series begins with patients' interpretations of their symptoms and problems, which they then report in a manner consistent with the cultural categories, words, images, and feelings they have for expressing distress. Clinicians then interpret patients' interpretations from their own ethnic cultures and the culture of diagnosing in psychiatry (Kleinman 1996). Diagnostic differences among clinicians of ethnic and racial minority patients may be a result of attributions, ambiguous patient behaviors, and clinicians' interpretive bias (Minsky et al. 2003; Strakowski et al. 1996; Trierweiler et al. 2000).

We generally presume that the greater the cultural and ethnic distance, the higher the potential for misunderstandings, misattributions, and under-, over-, or misdiagnosis of patients. However, past research with Hispanics shows that cultural, linguistic, and ethnic proximity does not reduce potential misunderstanding. Some studies report that Spanish-speaking patients appear more psychotic during Spanish interviews than English ones (Del Castillo 1970). Others report that Spanish-dominant persons with schizophrenia are rated as more symptomatic when interviewed in English than in Spanish (Marcos et al. 1973). Even when clinician ethnicity is held constant by using bilingual Hispanic clinicians to evaluate Hispanic patients in both English and Spanish, symptoms are rated as more severe in Spanish interviews (Price and Cuellar 1981; Malgady and Costantino 1998).

Much of the diagnostic research on Hispanics predates the advent of multi-axial, criteria-driven diagnostic systems and DSM-based structured clinical interviews (American Psychiatric Association [APA] 2000). The research to date has not included objective diagnostic measures and has incorporated limited diagnostic categories, mostly, schizophrenia and major depression (Marcos et al. 1973; Price and Cuellar 1981; Malgady and Costantino 1998). Most studies have also differed on patients' illness severity (i.e., acute, recent admissions; chronically ill community-dwelling outpatients). Clinical studies with Hispanic service-seekers in typical community settings that include a range of diagnostic categories and symptom severity are lacking, despite the need for research reflecting real-world, everyday community practice (Hohman and Shear 2002). Everyday realities in community diagnostic practice include heterogeneity in patients' social, cultural, ethnic, racial, economic, and clinical profiles; diagnosing without

structured instruments or objective measures; and a relatively homogeneous clinician group.

Within this context, we compared levels of diagnostic agreement between Hispanic and non-Hispanic clinicians and between clinicians and a structured clinical research interview when assessing adult Hispanic outpatients. We also examined clinicians' agreement on ratings of patients' symptom severity and assessment of general functioning.

## Method

### Participants

We screened US or foreign-born Hispanic patients 18 years or older without a history of psychiatric treatment during the preceding 12 months, requesting psychiatric services for the first time in the adult outpatient clinics of a large urban general hospital. New patients were selected to avoid the influence of past psychiatric records and patients with experience in the diagnostic process. Eligible patients were required to (a) complete a brief capacity-to-consent screener (Zayas et al. 2005) to determine participation or not; (b) be videotaped during their intake interview with the assigned clinician; (c) complete a structured clinical interview administered by a research clinician; and (d) permit a second clinician to view the videotaped interview. Upon showing capacity and giving informed consent, participants were interviewed by a Hispanic or non-Hispanic clinician in the adult psychiatric services. The clinicians had volunteered and also given informed consent to conduct live interviews or to render diagnoses from videotaped interviews. All procedures in this project were approved by the human subject committees of Washington University in St. Louis and data collection sites.

### Instruments

As the objective measure against which to compare clinician diagnoses, we used the Structured Clinical Interview for DSM-IV-TR Research Version (SCID; First et al. 2002). The SCID is a widely used clinical instrument with good validity and reliability for Axis I. We used English and Spanish versions. Clinicians rated symptom severity with the brief psychiatric rating scale (BPRS; Overall and Gorham 1988) which consists of 18 items rated on a seven-point scale. Each item is anchored by definitions and descriptions of expected symptoms and problems for each of the seven intensity rating options and has good reliability and validity (Lachar et al. 2001). In the analyses, we used patients' total BPRS scores.

### Procedures

We used a "quasi-random" approach to assigning Hispanic (monolingual Spanish or English, or bilingual) walk-in patients or those with appointments to diagnostic interviews regardless of clinicians' ethnicity or Spanish abilities. Assignment was based on clinic appointment schedule, on-call, or walk-in schedule. When non-Hispanic clinicians encountered a Spanish-speaking patient for the live interview or watched a Spanish-language videotaped interview, interpretation was provided by the SCID administrator (a bilingual, bicultural clinically experienced master's degree-level psychologist). The SCID was administered after the live interview in every case except one. After conducting live interviews or watching videotaped interviews, clinicians completed questionnaires requiring diagnoses and responses to both quantitative and qualitative questions. Clinicians were instructed to rank-order up to three diagnoses in Axes I and II. In all instances, one clinician was Hispanic by self-identification and one was non-Hispanic by self-identification, regardless of whether they conducted live diagnostic interviews or watched the videotapes. In this report, we present findings based on Axis I (clinical disorders and other conditions that may be the focus of clinical attention) and V (global assessment of functioning), and symptom severity data.

### Data Analyses

Potential differences between Hispanic and non-Hispanic clinicians on basic demographic and professional characteristics were assessed via independent samples *t*-tests. To examine diagnostic reliability between SCID, Hispanic and non-Hispanic clinicians, we computed both percentage agreements (the most basic form of diagnostic reliability) and kappa statistics for broad levels of disorders (e.g., mood disorders, substance-related disorders) and narrow diagnostic levels (e.g., dysthymic disorder, alcohol use disorder). For clinical samples with higher base rates of disorders than the general population, kappa statistics are an appropriate measure of diagnostic reliability (Langebucher et al. 1996; Spitznagel and Helzer 1985). To determine whether the number of diagnoses rendered by SCID, Hispanic and non-Hispanic clinicians differed significantly, we computed difference chi-square and McNemar chi-square statistics. The difference chi-square assesses whether the distribution of obtained responses differs from that expected by chance, while the McNemar chi-square evaluates whether one "diagnostician" (i.e., SCID, Hispanic, or non-Hispanic clinician) diagnosed significantly more or less than another. Finally, we used general linear modeling to compare mean differences between Hispanic and non-Hispanic clinicians' ratings of

patients' symptom severity and global assessment of functioning while controlling for interview language and live versus video condition.

## Results

We identified 150 eligible patients, 96 of whom agreed to participate (64%). The most frequent reason given by the 54 who declined to participate was concern about the videotaping of interviews. Independent samples *t*-tests revealed that refusers did not differ significantly on any demographic characteristic (e.g., age and gender) from participants. Eight of the 96 who agreed to participate (seven men, one woman) failed to demonstrate capacity to consent on a brief screener developed for this research protocol, mostly from apparent cognitive deficits.

Eighty-eight patients (57% male) were enrolled in the study after providing informed consent. Their average age was 41 years ( $SD = 13$ , range 18–83 years) and they were mostly of Dominican (36%) or Puerto Rican (22%) descent, the two largest Hispanic groups in the local community. Mexicans and Ecuadorians each comprised 7% of the sample and those of other Latin American ancestry were 26%. Most patients (81%) had a high school education or more and 91% were US citizens or legal aliens.

Forty-seven clinicians volunteered to participate in the study and gave informed consent. Most were psychiatrists (40%) and psychiatric social workers (40%). Two thirds of clinicians were non-Hispanic of any race and twice as many participating clinicians were females (68%). Clinicians averaged 10 years of experience ( $SD = 8.7$ ) in adult psychiatric practice. The average years in practice ranged from 4 years among psychiatric residents to 12 years among social workers to 19 years among psychologists. The only significant difference between Hispanic and non-Hispanic clinicians was the average years of adult psychiatric practice, with Hispanic psychiatrists having more than double the experience of non-Hispanic psychiatrists ( $P = 0.013$ ).

Our analyses used all Axis I diagnoses assigned by clinicians regardless of rank order. As is customarily found in community samples, mood disorders were most frequently diagnosed (72 of 88 or 82%; Table 1). Substance-related disorders were the next most frequent diagnoses, followed by anxiety and adjustment disorders. As expected in a community sample seeking services for the first time, schizophrenia and other psychotic disorders were least often diagnosed. In all instances except alcohol use and adjustment disorders, non-Hispanic clinicians diagnosed more patients with Axis I disorders than did Hispanic clinicians.

Table 1 also shows the agreement between Hispanic and non-Hispanic clinicians on patients' Axis I disorders. Discrepancies in agreement are evident in patients who received a particular diagnosis from one clinician but not the other. The total agreement between Hispanic and non-Hispanic clinicians ranged from a low of 8% for schizophrenia and other psychotic disorders to a high of 69% for mood disorders. Disagreement was pronounced in substance-related disorders, adjustment disorder, and schizophrenia. For patients given a diagnosis of alcohol use disorders ( $n = 20$ ) and generalized anxiety ( $n = 16$ ), clinicians failed to agree on any patients with these disorders. The highest level of agreement between clinicians occurred in the diagnosis of substance use disorders (69%).

We compared diagnostic agreement between clinicians and SCID diagnoses on Axis I. Table 1 shows the rates at which both Hispanic and non-Hispanic clinicians agreed with the SCID on the same patients, and when the clinicians agreed with the SCID but disagreed with each other. At the high end of agreement, the SCID diagnosed 36 patients with mood disorders and both Hispanic and non-Hispanic clinicians agreed with the SCID on 21 of these patients (58%). Forty-eight patients with substance related disorders were identified by the SCID, yet both clinicians agreed with the SCID on less than half of these patients ( $n = 20$ ; 42%). Clinicians agreed with the SCID on only 28% of patients with anxiety disorders and 8% with schizophrenia and other psychotic disorders.

We found that neither Hispanic nor non-Hispanic clinicians agreed with the SCID on alcohol use and psychotic disorders not otherwise specified, and only once on panic disorders. At the high end of agreement, both Hispanic and non-Hispanic clinicians agreed with the SCID on major depressive disorders half the time (50%) and they were nearly evenly split on the remaining 50% of patients diagnosed by the SCID (four by the Hispanic clinicians and SCID and five by the non-Hispanic clinicians and SCID). In diagnosing substance use disorders, Hispanic clinicians failed to agree with the SCID on a single case and non-Hispanic clinicians agreed with the SCID on six patients. Total agreement (all clinicians) with the SCID was less than half (46% of patients).

Kappa statistics assess reliability of diagnoses, with Kappas  $< 0.40$  being poor to fair, between 0.40 and 0.80 moderate to substantial, and  $> 0.80$  exceptional (Landis and Koch 1977). As shown in Table 2, kappa statistics were overall low to moderate, ranging from  $-0.13$  to 0.74. For SCID versus Hispanic diagnoses, kappa statistics ranged from  $-0.07$  to 0.35. For SCID versus non-Hispanic clinician, kappa statistics ranged from  $-0.06$  to 0.52. Negative Kappas are rare and indicate that the diagnostic reliability was even less than what was expected by chance.

**Table 1** Percentage of agreement between clinicians, and between clinicians and SCID

Diagnostic category	Total patients diagnosed by either clinician <sup>a</sup>		Total agreement on patients Dx <sup>b</sup>		Diagnosed by Hispanic only <sup>c</sup>		Diagnosed by non-hisp only <sup>c</sup>	
	<i>n</i>	(%)	<i>n</i>	(%)	<i>n</i>	(%)	<i>n</i>	(%)
<i>Agreement between Hispanic and non-Hispanic clinicians</i>								
Mood disorders	72	(82)	46	(64)	10	(14)	16	(22)
Major depressive	56	(64)	33	(59)	8	(14)	15	(27)
Bipolar	12	(14)	2	(17)	7	(58)	3	(25)
Other mood	21	(24)	1	(5)	8	(38)	12	(57)
Substance related disorders	44	(50)	24	(55)	5	(11)	15	(34)
Alcohol use	20	(23)	0	(0)	11	(55)	9	(45)
Substance use	32	(37)	22	(69)	1	(3)	9	(28)
Anxiety disorders	39	(44)	12	(31)	11	(28)	16	(41)
Generalized anxiety	16	(19)	0	(0)	6	(37)	10	(63)
Other anxiety	12	(14)	1	(9)	5	(41)	6	(50)
Adjustment disorder	19	(22)	5	(26)	9	(47)	5	(26)
Schizophrenia/other psychotic	13	(15)	1	(8)	3	(23)	9	(70)
Diagnostic category	Total No. Pts Dx by SCID <sup>d</sup>		SCID, Hisp, and non-Hisp <sup>e</sup>		SCID and Hisp only <sup>e,f</sup>		SCID and non-Hisp only <sup>e,f</sup>	
	<i>n</i>	(%)	<i>n</i>	(%)	<i>n</i>	(%)	<i>n</i>	(%)
<i>Agreement between Hispanic and non-Hispanic clinicians and SCID</i>								
Mood disorders	36	(50)	21	(58)	6	(17)	4	(11)
Major depressive	30	(42)	15	(50)	4	(14)	5	(17)
Substance related disorders	48	(67)	20	(42)	2	(4)	10	(21)
Alcohol use	34	(48)	0	(0)	8	(24)	5	(15)
Substance use	42	(59)	19	(46)	0	(0)	6	(15)
Anxiety disorders	25	(35)	7	(28)	3	(12)	7	(28)
Panic	14	(20)	1	(8)	0	(0)	1	(8)
Schizophrenia/other psychotic	14	(20)	1	(8)	2	(15)	3	(22)
Psychotic disorder NOS	13	(18)	0	(0)	0	(0)	1	(18)

<sup>a</sup> Total number of patients diagnosed by either clinician (not necessarily the same patients). Percentages based on full sample of 88 patients

<sup>b</sup> Total number of patients on which both clinicians agreed on the diagnosis. Percentages based on total patients diagnosed by either clinician (column 1)

<sup>c</sup> Total number of patients diagnosed in which clinicians did not agree on the diagnosis. Columns three and four represent diagnoses given to different patients. Percentages based on total patients diagnosed by both clinicians (column 1)

<sup>d</sup> Percentages based on 72 patients with SCID diagnoses

<sup>e</sup> Percentages based on total patients diagnosed with the disorder by SCID (column 1)

<sup>f</sup> Clinicians agreed with SCID but disagreed with each other

Next, we examined whether the distribution of assigned diagnoses significantly differed from chance using difference chi-squares, and whether one clinician diagnosed at a significantly different rate than the other, using McNemar chi-squares (Table 2). When comparing mood disorder diagnoses rendered by Hispanic versus non-Hispanic clinicians, the difference chi-square was significant ( $P = 0.0015$ ) and the McNemar was not. For major depressive disorder the difference chi-square was significant ( $P < 0.0001$ ) and the McNemar was again non-significant. For other mood disorders (including dysthymia

and depressive disorder NOS) neither chi-square was significant.

In the substance-related disorder and substance use disorder categories, the difference ( $P < 0.0001$  for both) and McNemar ( $P = 0.0253$ ,  $P = 0.0114$ , respectively) chi-squares were all significant, indicating that non-Hispanic clinicians assigned the diagnosis at higher rates than Hispanic clinicians. For alcohol use neither chi-square was significant.

For anxiety disorders, the difference chi-square was significant ( $P = 0.0147$ ) and the McNemar chi-square was

**Table 2** Comparisons between clinicians and between clinicians and SCID

Disorder category	Kappas (95% CI)	Chi-square statistics	
		Difference	McNemar
<i>Hispanic and non-Hispanic clinicians</i>			
Mood disorders	0.33 (0.13–0.54)	$\chi^2(1) = 10.11^{**}$	$\chi^2(1) = 1.38$
Major depressive	0.48 (0.30–0.66)	$\chi^2(1) = 20.84^{***}$	$\chi^2(1) = 2.12$
Bipolar	0.23 (–0.09–0.55)	Fisher's exact	$\chi^2(1) = 1.60$
Other mood	–0.03 (–0.22–0.15)	Fisher's exact	$\chi^2(1) = 0.80$
Substance related disorders	0.52 (0.35–0.70)	$\chi^2(1) = 25.90^{***}$	$\chi^2(1) = 5.00^*$
Alcohol use	–0.13 (–0.18–0.07)	Fisher's exact	$\chi^2(1) = 0.20$
Substance use	0.74 (0.58–0.89)	$\chi^2(1) = 49.83^{***}$	$\chi^2(1) = 6.40^*$
Anxiety disorders	0.26 (0.04–0.47)	$\chi^2(1) = 5.95^*$	$\chi^2(1) = 0.93$
Generalized anxiety	–0.09 (–0.015 to –0.04)	Fisher's exact	$\chi^2(1) = 1.00$
Other anxiety	0.09 (–0.19–0.37)	Fisher's exact	$\chi^2(1) = 0.09$
Adjustment disorder	0.33 (0.06–0.60)	Fisher's exact**	$\chi^2(1) = 1.14$
Schizophrenia/other psychotic	0.08 (–0.17–0.34)	Fisher's exact	$\chi^2(1) = 3.00$
<i>Hispanic clinicians and SCID</i>			
Mood disorders	0.17 (–0.05–0.38)	$\chi^2(1) = 2.25$	$\chi^2(1) = 4.80^*$
Major depressive	0.25 (0.02–0.47)	$\chi^2(1) = 4.46^*$	$\chi^2(1) = 0.93$
Bipolar	0.15 (–0.19–0.49)	Fisher's exact	$\chi^2(1) = 2.00$
Other mood	–0.07 (–0.13 to –0.007)	Fisher's exact	$\chi^2(1) = 3.00$
Substance related disorders	0.33 (0.17–0.49)	$\chi^2(1) = 12.78^{***}$	$\chi^2(1) = 23.15^{***}$
Alcohol use	0.22 (0.06–0.38)	Fisher's exact*	$\chi^2(1) = 23.15^{***}$
Substance use	0.41 (0.25–0.57)	$\chi^2(1) = 18.44^{***}$	$\chi^2(1) = 23.00^{***}$
Anxiety disorders	0.25 (0.01–0.48)	$\chi^2(1) = 4.60^*$	$\chi^2(1) = 2.13$
Generalized anxiety	–0.06 (–0.10 to –0.02)	Fisher's exact	$\chi^2(1) = 0.00$
Other anxiety	–0.02 (–0.06–0.01)	Fisher's exact	$\chi^2(1) = 1.80$
Adjustment disorder	0.35 (0.03–0.66)	Fisher's exact*	$\chi^2(1) = 5.44^*$
Schizophrenia/other psychotic	0.30 (0.03–0.58)	Fisher's exact**	$\chi^2(1) = 11.00^{**}$
<i>Non-Hispanic clinicians and SCID</i>			
Mood disorders	–0.06 (–0.26–0.15)	$\chi^2(1) = 0.28$	$\chi^2(1) = 6.74^{**}$
Major depressive	0.16 (–0.06–0.37)	$\chi^2(1) = 1.98$	$\chi^2(1) = 3.90^*$
Bipolar	0.21 (–0.20–0.62)	Fisher's exact	$\chi^2(1) = 0.67$
Other mood	0.11 (–0.17–0.40)	Fisher's exact	$\chi^2(1) = 3.60$
Substance related disorders	0.46 (0.28–0.65)	$\chi^2(1) = 19.01^{***}$	$\chi^2(1) = 12.80^{***}$
Alcohol use	0.10 (–0.05–0.24)	Fisher's exact	$\chi^2(1) = 23.52^{***}$
Substance use	0.52 (0.35–0.70)	$\chi^2(1) = 23.95^{***}$	$\chi^2(1) = 14.22^{***}$
Anxiety disorders	0.38 (0.15–0.60)	$\chi^2(1) = 10.19^{**}$	$\chi^2(1) = 0.20$
Generalized anxiety	0.28 (–0.07–0.63)	Fisher's exact	$\chi^2(1) = 2.00$
Other anxiety	–0.02 (–0.07–0.02)	Fisher's exact	$\chi^2(1) = 3.57$
Adjustment disorder	0.12 (–0.20–0.44)	Fisher's exact	$\chi^2(1) = 1.00$
Schizophrenia/other psychotic	0.23 (–0.04–0.51)	Fisher's exact	$\chi^2(1) = 1.67$

Note: All Fisher's exact are two-sided: \*  $P \leq 0.05$ , \*\*  $P \leq 0.01$ , \*\*\*  $P \leq 0.001$

not. Finally, in the adjustment disorder category, once again the difference chi-square was significant ( $P = 0.0080$ ) and the McNemar chi-square was not. For all remaining diagnoses, no significant differences were found between Hispanic and non-Hispanic clinicians. It is worth noting that for substance related diagnoses, the video versus live condition appears to be a confound, with all five diagnoses rendered by Hispanic clinicians from video.

For SCID versus Hispanic clinicians, kappas ranged from –0.07 for other mood disorders, to 0.41 for substance use disorders. Difference chi-squares were significant in the major depressive disorders, anxiety disorders, schizophrenia and other psychotic disorders, and all substance-related categories. McNemar chi-squares were significant for the mood disorders, adjustment disorders, schizophrenia and other psychotic disorders, panic disorder, and all

**Table 3** Clinician ratings of patients' symptom severity and global assessment of functioning

Clinician ethnicity	Least square means	Adjusted <i>t</i>	<i>P</i> value
Symptom severity ( <i>N</i> = 87) <sup>a</sup>			
Hispanic	30.73	3.81	0.0003
Non-Hispanics	36.54		
<i>Global assessment of functioning</i> <sup>b</sup>			
Current ( <i>N</i> = 85)			
Hispanics	55.76	−5.49	<0.0001
Non-Hispanics	47.53		
Highest level past year ( <i>N</i> = 55)			
Hispanics	63.99	−3.43	0.0014
Non-Hispanics	54.34		

<sup>a</sup> Reflect mean total scores given by clinicians on brief psychiatric rating scale controlling for video/live condition and language. Higher scores indicate greater level of severity

<sup>b</sup> Scores based on the DSM-IV-R scale of 0–100 with higher scores indicating better functioning

substance-related categories, indicating a significant difference in the rate of diagnosing. In every case, the SCID diagnosed more disorders than the Hispanic clinicians.

For SCID versus non-Hispanic clinicians, kappas ranged from −0.06 for mood disorders, to 0.52 for substance use disorders. Difference chi-squares were significant in the anxiety disorders and substance-related categories. McNemar chi-squares were significant for mood disorders, major depressive disorders, panic disorder, and all substance-related categories. In every case, the SCID diagnosed more disorders than non-Hispanic clinicians.

Finally, we compared Hispanic and non-Hispanic clinicians' ratings of patients' symptom severity and their ratings of global assessment of functioning while controlling for language of interview and live vs. video condition (Table 3). On the BPRS, non-Hispanic clinicians rated patients as showing more severe symptoms ( $P < 0.0003$ ). On global assessment of functioning, non-Hispanic clinicians assessed patients as significantly lower in current functioning ( $P < 0.0001$ ) and functioning in the past year ( $P = 0.0014$ ). Neither interview language nor condition (live versus video) made a difference in clinician ratings.

## Discussion

Hispanic and non-Hispanic clinicians differed significantly in the diagnoses they assigned Hispanic outpatients and demonstrated low-levels of agreement; sometimes they assigned similar diagnoses but not to the same persons. Clinician agreement with the SCID also varied in pronounced ways. If a three-way agreement (i.e., both clinicians and SCID) more closely approximates an

accurate diagnosis, then this was not supported for diagnostic categories that are very common to community mental health practice (e.g., mood disorders, substance related disorders, alcohol use disorders, substance use disorders and adjustment disorders). Thus, for instance, it was alarming when the SCID diagnosed alcohol abuse in 34 patients but there was not a single case in which both clinicians agreed with the SCID. And when the clinicians did agree with the SCID, their diagnoses were completely divergent: Hispanic clinicians agreed with the SCID on eight patients and non-Hispanics agreed with the SCID on five entirely different patients. Hispanic and non-Hispanic clinicians agreed with the SCID on 46% of patients with substance abuse disorders; non-Hispanic clinicians identified six additional persons with these disorders. Significant disagreement was also evident in clinicians' ratings of patients' symptoms and assessment of patients' functional capacity for both the current period and past year. This is also at odds with past research on Hispanic patients, which shows that Hispanic clinicians rate symptoms of Hispanic patients more severely than do non-Hispanic clinicians (Malgady and Costantino 1998). But sampling differences in the two studies may account for these findings.

These findings point to our concern about diagnostic agreement across clinicians and instruments. They also run contrary to the assumption that consumers should expect accuracy and consistency in their diagnosis regardless of clinician. Such findings raise questions about what information clinicians rely on to reach their diagnostic conclusions (Neighbors et al. 2003; Strakowski et al. 1996; Trierweiler et al. 2000). Our results seem to concur with past findings that clinicians may be influenced by different factors and make attributions of pathology differently, based possibly on clinicians' cultural and social biases, and assessing its magnitude in very distinct ways (Malgady and Costantino 1998; Neighbors et al. 2003; Strakowski et al. 1996; Trierweiler et al. 2000). As Minsky et al. (2003) point out, bias may be present when clinicians apply DSM-IV criteria to patients differently. As evident in previous research and in our study, diagnostic bias (if considered a factor in our results) is very complex and does not easily follow ethnic lines. Bias may intrude because clinicians may not use diagnostic criteria effectively due to cultural variances in characteristic symptom clusters that clinicians use as "a template for assigning a diagnosis" (Minsky et al. 2003, p. 643). These symptom clusters may be confounded by ambiguous signs, body language, and verbal idiosyncrasies that interfere with appropriate applications of DSM-IV criteria. Some researchers (Minsky et al. 2003; Strakowski et al. 1997) have suggested that structured interviews may attenuate race-related diagnostic differences and explain clinician disagreement or error as the result of ambiguities in the patient-clinician interaction. As

Strakowski et al. (1997) show, diagnostic agreement between clinicians and structured research interview occurs less often when the patients are non-White minority persons than White persons.

Limitations restrict the conclusiveness and generalizability of our findings. Both our clinician and patient samples were small and patient refusal rate (36%) was high. Ours was a self-selected, non-representative sample of Hispanics. Without comparison groups it is impossible to know whether disagreement between clinicians of different ethnicities and between clinicians and SCID would have been any better for other ethnic patients. Our clinician sample too may be biased by more experienced Hispanic clinicians, and the small number of clinicians prevented comparisons of agreement levels of same-ethnicity clinicians or by language, discipline, and gender. Imbalances also existed in live and video interview conditions, with Hispanics conducting most of the live interviews (70%). Diagnosing from video, of course, is not customary practice, and clinicians viewing tapes were constrained by the questions that were asked on tape. Additionally it is clear that the use of videos also introduced additional confounds into both findings and accurate interpretation. Finally, because of the pilot study nature and the concomitant restrictions in budget, we used one interpreter for non-Hispanic, English-only clinicians which introduced a level of potential uncontrolled and unexamined bias. The interpreter was also our SCID administrator, potentially introducing systematic bias that could be reduced when using more persons and separating the roles.

This pilot study is, nonetheless, one of a relatively few recent-day examinations of clinical diagnostic practice with Hispanics. It also introduced two objective measures to examine clinician diagnostic judgment. The use of a community sample based within the real-world, real-time context of a community mental health clinic enhanced its ecological validity, albeit with the limitations noted above. Except for having second clinicians assess patients from videotaped interviews, the study followed a relatively naturalistic design that reflected the everyday practice of a busy urban clinic: assigning patients to clinicians based on schedules and staffing patterns and not by segregating patients a priori by diagnosis. Most community settings have neither the time nor the trained personnel to administer structured diagnostic instruments as part of their intake assessments. In this regard, our assessment of clinicians' agreement on diagnoses reflects community psychiatric practice and points to the need for more training in and use of structured diagnostic interviews in community mental health practice.

Adding to a body of diagnostic research on Hispanics, our study points to the complexity of the problem and to challenges to future research in this area that might help

make sense of the conundrum. This study makes the case for looking at how objective diagnostic measures impact agreement and accuracy. Therefore, we need research that uses larger, representative multiethnic samples of both patients and clinicians in sufficient numbers to provide enough statistical power to explore where the highest and lowest levels of agreement exist, and how we can understand the differences. Larger samples can allow for various comparisons by patient and clinician gender and ethnicities, and between clinicians by discipline, gender, years of experience, and so on. Using in vivo diagnostic interviews for all clinicians' ratings overcomes the methodological confound introduced by videotaped interviews. Multiple SCID raters will attenuate the bias that can come from having a single rater. Comparison groups of ethnic minority and non-minority patients will answer a critical question of whether lack of agreement among our clinicians and the SCID was unique to diagnosing Hispanic patients or would have existed at the same level of disagreement regardless of ethnic differences.

Clearly, we can deduce that usual-care diagnoses in a typical community mental health center are very unreliable. The immediate concern becomes what happens to patients when we link diagnosis to treatment decisions. How diagnoses influence the treatment, including pharmacological interventions, instituted by Hispanic and non-Hispanic clinicians assessing Hispanic patients, and the outcomes of these treatment decisions, are of the utmost importance. Evidence-based therapies are, after all, predicated on accurate diagnoses.

## References

- Alegria, M., & McGuire, T. (2003). Rethinking a universal framework in the psychiatric symptom-disorder relationship. *Journal of Health and Social Behavior*, *44*, 257–274.
- American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., Text revision). Washington, DC: Author.
- Arnold, L. M., Jr, Keck, P. E., Collins, J., Wilson, R. M., Fleck, D. E., Corey, K. B., et al. (2004). Ethnicity and first-rank symptoms in patients with psychosis. *Schizophrenia Research*, *67*, 207–212.
- Cabassa, L. J., Lester, R., & Zayas, L. H. (2007). It's like being in a labyrinth: Hispanic immigrants' perceptions of depression and attitudes toward treatments. *Journal of Immigrant and Minority Health*, *9*, 1–16.
- Cabassa, L. J., Zayas, L. H., & Hansen, M. (2006). Latino adults' access to mental health services: A review of epidemiological studies. *Administration and Policy in Mental Health*, *33*, 316–330.
- Del Castillo, J. (1970). The influence of language upon symptomatology in foreign-born patients. *American Journal of Psychiatry*, *127*, 242–244.
- First, M. B., Gibbon, M., Spitzer, R. L., & Williams, J. B. W. (2002). *Structured clinical interview for DSM-IV-TR Axis I disorders*. NY: New York State Psychiatric Institute.



- Hohman, A. A., & Shear, M. K. (2002). Community-based intervention research: Coping with the “noise” of real life in study design. *American Journal of Psychiatry*, *159*, 201–207.
- Kleinman, A. (1996). How is culture important for DSM-IV? In J. E. Mezzich, A. Kleinman, H. Fabrega Jr & D. L. Parron (Eds.), *Culture and psychiatric diagnosis: A DSM-IV perspective* (pp. 15–25). Washington, DC: American Psychiatric Press.
- Lachar, D., Bailey, S. E., Rhoades, H. M., Espadas, A., Aponte, M., Cowan, K. A., et al. (2001). New subscales for an anchored version of the brief psychiatric rating scale: Construction, reliability, and validity in acute psychiatric admissions. *Psychological Assessment*, *13*, 384–395.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.
- Langenbucher, J., Labouvie, E., & Morgenstern, J. (1996). Measuring diagnostic agreement. *Journal of Consulting and Clinical Psychology*, *64*, 1285–1289.
- Malgady, R. G., & Costantino, G. (1998). Symptom severity in bilingual Hispanics as a function of clinician ethnicity and language of interview. *Psychological Assessment*, *10*, 120–127.
- Malgady, R. G., Rogler, L. H., & Costantino, G. (1987). Ethnocultural and linguistic bias in mental health evaluation of Hispanics. *American Psychologist*, *42*, 228–234.
- Marcos, L. R., Alpert, M., Urcuyo, L., & Kesselman, M. (1973). The effect of interview language on the evaluation of psychopathology in Spanish-American schizophrenic patients. *American Journal of Psychiatry*, *130*, 549–553.
- Minsky, S., Vega, W., Miskimen, T., Gara, M., & Escobar, J. (2003). Diagnostic patterns in Latino, African American, and European American psychiatric patients. *Archives of General Psychiatry*, *60*, 637–644.
- Neighbors, H. W., Trierweiler, S. J., Ford, B. C., & Muroff, J. R. (2003). Racial differences in DSM diagnosis using a semi-structured instrument: The importance of clinical judgment in the diagnosis of African Americans. *Journal of Health and Social Behavior*, *44*, 237–256.
- Overall, J. E., & Gorham, D. R. (1988). The brief psychiatric rating scale: Recent development in ascertainment and scaling. *Psychopharmacology Bulletin*, *24*, 97–99.
- Price, C. S., & Cuellar, I. (1981). Effects of language and related variables on the expression of psychopathology in Mexican-American psychiatric patients. *Hispanic Journal of Behavioral Sciences*, *3*, 145–160.
- Spitznagel, E. L., & Helzer, J. E. (1985). A proposed solution to the base rate problem in the Kappa statistic. *Archives of General Psychiatry*, *42*, 725–728.
- Strakowski, S. M., Flaum, M., Amador, X., Bracha, H. S., Pandurangi, A. K., Robinson, D., et al. (1996). Racial differences in the diagnosis of psychosis. *Schizophrenia Research*, *21*, 117–124.
- Strakowski, S. M., Hawkins, J. M., Jr, Keck, P. E., McElroy, S. L., West, S. A., Bourne, M. L., et al. (1997). The effects of race and information variance on disagreement between psychiatric emergency services and research diagnoses in first-episode psychosis. *Journal of Clinical Psychiatry*, *58*, 457–463.
- Strakowski, S. M., Jr, Keck, P. E., & Arnold, L. M. (2003). Ethnicity and diagnosis in patients with affective disorders. *Journal of Clinical Psychiatry*, *64*, 747–754.
- Trierweiler, S. J., Neighbors, H. W., Munday, C., Thompson, E. E., Binion, V. J., & Gomez, J. P. (2000). Clinician attributions associated with the diagnosis of schizophrenia in African American and non-African American patients. *Journal of Consulting and Clinical Psychology*, *68*, 171–175.
- US Census Bureau. (2004). US interim projections by age, sex, race, and Hispanic origin [data file]. <http://www.census.gov>. Accessed 17 June 2004.
- Vega, W. A., & Alegria, M. (2001). Latino mental health and treatment in the United States. In M. Aguirre-Molina, C. W. Molina & R. E. Zambrana (Eds.), *Health issues in the Latino community* (pp. 179–208). San Francisco: Jossey-Bass.
- Vega, W. A., & Lopez, S. R. (2001). Priority issues in Latino mental health services research. *Mental Health Services Research*, *3*, 189–200.
- Zayas, L. H., Cabassa, L. J., & Perez, M. C. (2005). Capacity-to-consent in psychiatric research: Development and preliminary testing of a screening tool. *Research on Social Work Practice*, *15*, 545–556.