

Confiabilidad de la Cédula de Entrevista Diagnóstica (DIS) en pacientes psiquiátricos mexicanos

Catalina F. González Forteza*
Jorge J. Caraveo Anduaga*
Luciana Ramos Lira*
José Juan Sánchez Báez*

Summary

In order to estimate the prevalence of specific psychiatric disorders in epidemiological studies, the National Institute of Mental Health in the United States has developed a highly structured instrument that can be used by lay interviewers. The clinical information is then classified by a computer program according to three different diagnostic systems: The DSM-III, the RDC and the Feighner's criteria.

The objective of the present study was to evaluate the inter-rater reliability with psychiatric patients.

The chosen design for the study was the interviewer-observers method, with 5 lay interviewers and 15 adult psychiatric patients chosen from two different institutions. Highly disturbed acute patients were not included.

For this exercise, the information was handled applying only the DSM-III criteria. The agreement in each diagnostic category was calculated by using the Intraclass Correlation Coefficient (ICC), as well as the descriptive index for IA presence (p) or IA absence (a) in each case.

The results indicated a high degree of inter-rater reliability. Applying the preemptions criteria the mean agreements were: ICC = 0.879, IA (p) = 0.965 and IA (a) = 0.966. Without preemptions the mean agreements were: ICC = 0.889, IA (p) = 0.931 and IA (a) = 0.986.

Three of the patients interviewed did not report enough clinical information in order to integrate a DSM-III diagnosis. The authors consider that the disposition of the patients to answer the questions is a factor that should be carefully kept in mind when using this kind of instruments, as this factor depends on the type of pathology, the acceptance of the illness as well as the sensibility of the population to accept this interview.

Nevertheless, the high reliability obtained by lay interviewers encouraged us to proceed with the study of the validity of the instrument in Mexico.

Resumen

Para estimar la prevalencia de trastornos psiquiátricos específicos en estudios epidemiológicos, el Instituto Nacional de Salud Mental de los Estados Unidos desarrolló un instrumento altamente estructurado que puede ser utilizado por entrevistadores legos; esto es, por personas sin ninguna experiencia en el campo de la clínica psiquiátrica. La información recabada por el instrumento, es clasificada por un programa de cómputo que emite un diagnóstico basado en tres sistemas: el DSM-III, el RDC y el Criterio de Feighner.

El presente trabajo tuvo como objeto evaluar el grado en el que coinciden las opiniones de los entrevistadores legos respecto

* Investigadores de la División de Investigaciones Epidemiológicas y Sociales. Instituto Mexicano de Psiquiatría, Calz. México-Xochimilco 101, Tlalpan, México D.F.

a los sujetos cuyo grado de perturbación mental amerita internamiento psiquiátrico. El diseño utilizado fue el de entrevistador-observador, y lo aplicaron 5 personas sin experiencia clínica. Se entrevistó a 15 pacientes: 14 internos y uno de consulta externa, de dos instituciones. Se excluyeron los pacientes en fase aguda, cuyo grado de perturbación interfiriera con el desarrollo de la entrevista. La información recabada para este estudio se manejó manualmente siguiendo los criterios de clasificación del DSM-III. El acuerdo diagnóstico al que se llegó fue valorado aplicando los criterios de exclusión especificados por el sistema, y sin aplicarlos. Para medir la concordancia de opiniones en cada categoría, se utilizó el Coeficiente de Correlación Intraclass (ICC); y se calcularon los índices de concordancia respecto a la presencia y la ausencia de las categorías.

Los resultados obtenidos indicaron un alto nivel de confiabilidad de los entrevistadores en el manejo del instrumento. Aplicando los criterios de exclusión, las medias obtenidas fueron ICC = 0.879, con un índice de acuerdo respecto a la presencia de categorías, de 0.965 y a la ausencia de categorías, de 0.966. Sin aplicar los criterios de exclusión, las medias de concordancia fueron: ICC = 0.889 con un índice de acuerdo en presencia de categorías, de 0.931 y en ausencia de categorías, de 0.986. El promedio de categoría diagnóstica por paciente, sin aplicar los criterios de exclusión, fue de 2.9, reduciéndose a 1.9 al aplicar dichos criterios.

Dado que 3 de los pacientes no dieron la información clínica necesaria para integrar algún diagnóstico DSM-III sin aplicar los criterios de exclusión, los autores mencionan que un aspecto importante que debe considerarse en estos casos, es la disposición de los entrevistados para responder; situación que depende tanto de su patología, o sea, del grado en el que están conscientes de su enfermedad, como del grado de sensibilidad de la población para aceptar este tipo de entrevista.

Los autores concluyen que el DIS permite asegurar un alto nivel de confiabilidad de los entrevistadores sin experiencia, y consideran útil el presente estudio para proseguir investigando la validez del DIS en México.

Introducción

La importancia de la confiabilidad en el campo de la salud mental es un aspecto metodológico fundamental, ya que difícilmente se puede alcanzar el progreso científico sin variables psiquiátricas confiablemente evaluadas (2). De manera sencilla, la confiabilidad ha sido definida como la constancia con la que se evalúa un rasgo.

En una amplia revisión bibliográfica llevada a cabo por Caetano (5), acerca de las causas de la baja confiabilidad del diagnóstico en psiquiatría, se señala que las fuentes de variación son: 1) las técnicas de entrevista; 2) la percepción de la patología; 3) la importancia atribuida a los síntomas y 4) la clasificación usada. Kendell (16), a su vez,

puntualiza que la principal fuente de variación en la confiabilidad del diagnóstico, es la formación clínica del entrevistador, pues ésta puede influir en su percepción de los signos y en su captación de los síntomas.

Con el fin de controlar estas fuentes de variación, para mejorar la confiabilidad de la información recabada, este mismo autor (15) señala que se han elaborado tres tipos de instrumentos: 1) las entrevistas estructuradas; 2) las escalas de síntomas y 3) los cuestionarios autoaplicables. Al respecto, Caraveo (6) menciona que una de las tendencias actuales en la epidemiología psiquiátrica es usar instrumentos estructurados, con conceptualizaciones precisas de los signos y síntomas que se van a detectar, y el desarrollo de criterios operacionales para la definición de las categorías diagnósticas.

Este aspecto está considerado y plasmado en el DSM-III (*Diagnostic and Statistical Manual of Mental Disorders*; 3a ed.) (1), cuyos criterios están ampliamente definidos y en constante revisión.

La Cédula de Entrevista Diagnóstica-DIS (*Diagnostic Interview Schedule*) (19), se elaboró basándose en los ejes I y II del DSM-III, que corresponden a trastornos psiquiátricos y a trastornos de la personalidad y del desarrollo, respectivamente. Del eje I se incluyen varios diagnósticos y del eje II sólo el trastorno de personalidad antisocial. El DIS es un instrumento altamente estructurado que permite llevar un riguroso control sobre las fuentes de variación, ya que la técnica de la entrevista la indican su mismo diseño y las instrucciones al entrevistador. La percepción de la patología y la importancia atribuida a los síntomas no influyen en la calificación ni en el desarrollo de la entrevista, puesto que para su evaluación no interviene el juicio clínico del entrevistador, sino la educación de las preguntas y respuestas, obtenidas mediante el seguimiento de un riguroso esquema de sondeo. Asimismo, la clasificación diagnóstica se hace de acuerdo con los criterios del DSM-III (1), el Criterio de Feighner (8), y el RDC (*Research Diagnostic Criteria*) (21), mediante un programa de cómputo que analiza la información obtenida.

De este modo, la principal fuente de variación señalada por Kendell, no influye en la elaboración del diagnóstico. Esto permite que el DIS pueda ser aplicado por personal no especializado, o sea, por entrevistadores legos, aspecto que responde al planteamiento hecho por Cabildo (4), acerca de la necesidad de desarrollar instrumentos que no tengan que ser aplicados por especialistas, sino por cualquier trabajador de la salud.

En el Instituto Mexicano de Psiquiatría se llevó a cabo el presente estudio de confiabilidad del DIS, para evaluar el grado en el que se pusieron de acuerdo los entrevistadores legos respecto a los sujetos cuyo grado de perturbación ameritó el internamiento psiquiátrico.

En estudios anteriores (19, 20) se ha informado acerca de los resultados de confiabilidad y validez en el empleo del DIS-II en inglés, en pacientes psiquiátricos hospitalizados, de consulta externa, expacientes y sujetos normales. El grado de concordancia en el estadístico Kappa fue de 0.40 o más, en todas las categorías diagnósticas consideradas por el instrumento. Respecto a la confiabilidad del DIS traducido al español (3) y aplicado a pacientes bilingües (anglo e hispano parlantes) y monolingües (que sólo hablan español) de la consulta externa psiquiátrica, el grado en el que estuvieron de acuerdo, medido por Kappa, mos-

tró un rango de 0.44 a 0.92, considerándose, por lo tanto, aceptables los niveles de confiabilidad y equivalencia con el DIS en inglés, en la mayoría de los diagnósticos DSM-III que fueron evaluados.

Método

En el presente estudio de confiabilidad se empleó la versión en castellano del DIS-II que el NIMH (*National Institute of Mental Health*) ha utilizado en San Antonio, Texas, para estudiar a la población hispanoamericana. Para poder aplicarlo en nuestro medio, fue necesario traducir al español las instrucciones dirigidas al entrevistador y llevar a cabo un estudio piloto para valorar el instrumento en su conjunto.

Los resultados de este estudio fueron favorables respecto a la comprensión del DIS por los pacientes psiquiátricos de consulta externa y los pacientes hospitalizados. Sin embargo, no sucedió lo mismo con los pacientes psicóticos en estado agudo, ya que no estaban en condiciones de concentrar su atención en la entrevista, ni de responder en forma adecuada a las preguntas. Esta experiencia y sus resultados se describen ampliamente en otro artículo de Ramos y cols (18).

Posteriormente se elaboró el instructivo y el manual, para iniciar la fase de entrenamiento con un curso de capacitación a los entrevistadores, acerca del manejo y la aplicación del DIS. Dicho curso duró dos semanas, y se empleó un abordaje metodológico teórico-práctico que se desarrolló en cuatro etapas:

1. Conocer el DIS en todas sus secciones; aclarar dudas e inquietudes; comprender los términos y la estructura.
2. Manejar el esquema de sondeo para aplicar la entrevista y calificar las respuestas.
3. Hacer ejercicios de calificación mediante la filmación de entrevistas con el DIS a pacientes hispano-parlantes de EU.

Estos videos fueron proporcionados por el *Epidemiologic Catchment Area Program* (ECA) del *National Institute of Mental Health* (NIMH), de Los Angeles, California, lugar en donde se han llevado a cabo diversos estudios con el DIS, debido a que allí viven muchas personas de origen hispano-americano (27% según el censo de 1980) (3).

4. Ejercicios de entrevistas y puntuación mediante *role-playing*, ejemplificando posibles situaciones a las que se puedan enfrentar, y alternando los papeles de entrevistador y observador. En las 2 últimas etapas, se le dio mucha importancia al manejo de la entrevista y al registro de las respuestas, aclarando las dudas y analizando los acuerdos y desacuerdos que hubieran surgido durante los ejercicios.

La muestra del estudio estuvo integrada por 15 pacientes (14 internos y 1 de consulta externa) de dos instituciones de atención psiquiátrica de la Ciudad de México (el Hospital Fray Bernardino Alvarez y la Clínica San Rafael). Para seleccionar a los pacientes se consideraron como criterios de inclusión que éstos fueran mayores de 18 años y que los entrevistadores y los observadores desconocieran el diagnóstico emitido por los médicos. Se excluyeron los pacientes en fase aguda, cuyo grado de perturbación interfiriera con el desarrollo de la entrevista por las razones antes expuestas en el estudio piloto.

Con el fin de evaluar la confiabilidad en el manejo del instrumento, se seleccionó el método entrevistador/observador, en el cual varios puntuadores califican en forma independiente a un mismo sujeto durante la entrevista. De esta forma fue posible obtener un índice de confiabilidad en relación con la calificación y la habilidad de los puntuadores para detectar los aspectos psicopatológicos que considera el DIS. Los resultados de este diseño son más acertados a medida que aumenta el número de puntuadores, razón por la cual se decidió que calificaran el instrumento 5 personas, para que el grado de concordancia que se alcanzara fuera más riguroso y permitiera evaluar con mayor exactitud la confiabilidad de las categorías diagnósticas.

El grupo de entrevistadores quedó constituido por 3 estudiantes del último año de psicología y 2 psicólogos sociales sin entrenamiento clínico específico. De estos 5 puntuadores, uno aplicaba el DIS y cuatro fungían como observadores independientes. Los papeles se alternaron a lo largo del estudio, de modo que todos pudieron entrevistar y observar a varios pacientes en diferentes sesiones.

Dado que no se aplicó el programa de cómputo, la información clínica recabada se analizó mediante procedimientos manuales, agrupando los síntomas para determinar las categorías diagnósticas de los pacientes, de acuerdo con los criterios consignados en el DSM-III. Esta información se evaluó de dos formas: con y sin los criterios de exclusión que estipula este sistema para arribar a las categorías diagnósticas. Por ejemplo, al aplicar los criterios de exclusión, sólo se podría considerar el diagnóstico de esquizofrenia en aquellos pacientes que dieran la información necesaria para cubrir los criterios clínicos del trastorno de ansiedad generalizada y de esquizofrenia, ya que en el trastorno de ansiedad la condición psicótica es una razón de exclusión. Por otro lado, cuando no se aplican los criterios de exclusión en estos pacientes, se consignan ambas categorías diagnósticas. Por lo tanto, un sujeto entrevistado con el DIS, puede presentar uno o varios diagnósticos, según la información que haya dado y el tipo de agrupación que se haya empleado. Para ilustrar mejor esta situación presentamos la tabla I en la que se muestran las categorías diagnósticas consideradas en el DIS, que presentan criterios de exclusión en relación con otros trastornos, según el sistema de clasificación diagnóstica del DSM-III. Cabe señalar que los trastornos de *anorexia nervosa*, abuso o dependencia de drogas o alcohol, síndrome orgánico cerebral y somatización, no están sujetos a estas reglas.

Como al aplicar los criterios de exclusión disminuye el número de diagnósticos captados, se decidió analizar las dos posibilidades de agrupación para estimar la confiabilidad en ambas situaciones y evaluar los efectos significativos, si los hubiera, en los grados de concordancia al aplicar o no los criterios de exclusión DSM-III, tal como lo han señalado Robins y cols (20).

El grado en el que coincidieron los puntuadores se evaluó mediante el Coeficiente de Correlación Intraclase (ICC) (2), que es una medida de confiabilidad altamente recomendada cuando más de dos puntuadores califican simultáneamente una entrevista. Al respecto, Fleiss (9) afirma que esta medida estadística es más acertada a medida que aumenta el número de puntuadores, por lo que recomienda que intervengan por lo menos cinco.

El ICC se calcula de la siguiente manera:

$$ICC = 1 - \left[\frac{NR}{T} \frac{(TR)-i \frac{N}{R-1} (S_i^2)}{(NR-T) (R-1)} \right]$$

Siendo la prueba de significancia:

$$N [1 + (R-1) ICC] > X^2 (N-1, \alpha)$$

Donde N = número de sujetos y R = número de puntuadores.

El siguiente paso del análisis estadístico es proceder a estimar el análisis descriptivo sugerido por Wing (22):

Índice de concordancia cuando el síntoma está ausente:

$$IA(a) = \frac{00}{(00) + (0+) + (+0)}$$

Índice de concordancia cuando el síntoma está presente:

$$IA(p) = \frac{++}{(+++) + (0+) + (+0)}$$

En donde 0 = síntomas ausentes

+ = síntomas presentes

Las puntuaciones del entrevistador más experimental se toman como patrón para compararlas con las de los demás entrevistadores.

Se calcularon el ICC y los IA(a) e IA(p) para estimar el grado de concordancia obtenido en cada categoría diagnóstica.

Resultados

La media de edad de la muestra del estudio fue de 31.4, con desviación estándar de 6.88 años, y la distribución por estado civil fue la siguiente: 10 solteros (66.67%), 3 casados (20.00%) y 2 divorciados (13.33%); en total, 14 hombres y 1 mujer.

TABLA I
Diagnósticos con criterios de exclusión del DSM-III

<i>Diagnósticos DSM-III, con criterios de exclusión en función de otros trastornos</i>	<i>Daño orgánico</i>	<i>Esquizofrenia</i>	<i>Depresión</i>	<i>Manía</i>	<i>Obsesión compulsiva</i>	<i>Agorafobia</i>	<i>Somatización</i>
Esquizofrenia	X	—	X	X	—	—	—
Episodio depresivo	X	X	—	—	—	—	—
Episodio maniaco	X	X	—	—	—	—	—
Agorafobia	—	X	X	—	X	—	—
Fobia simple	—	X	—	—	X	—	—
*Fobia social	—	—	X	—	—	—	—
Trastorno de pánico	—	X	X	—	—	X	X
Obsesiones-compulsiones	X	X	X	—	—	—	—
Disfunción psicosexual	X	X	X	—	X	—	X
*Ansiedad generalizada	—	X	X	—	—	—	—
Personalidad antisocial	—	X	—	X	—	—	—
Trastorno distímico	—	X	X	—	—	—	—

Fuente: Robins L N y cols: *Validity of the diagnostic interview schedule, version II: DSM-III diagnoses. Psychological Medicine, 12:855-870, 1982.*

*Categorías agregadas para este estudio

TABLA II
Distribución de las observaciones positivas captadas con el DIS, sin aplicar criterios de exclusión del DSM-III

Paciente	Somatización	Pánico	Fobia		Agorafobia	Ansiedad generalizada	Depresión mayor	Trastorno distímico	Manía	Esquizofrenia	Anorexia nervosa	Alcohol		Obs-comps.	Drogas		Personalidad antisocial	Disfunción psicosexual	Mini mental State	No. de DXS
			Social	Simple								Abuso	Dependencia		Abuso	Dependencia				
1	-	-	-	-	-	-	5	5	-	-	-	5	-	-	-	5	-	-	4	
2	-	-	-	-	-	-	5	-	-	-	-	5	-	-	-	5	-	-	1	
3	-	-	-	-	-	-	5	-	4	4	-	-	-	5	-	-	5	-	5	
4	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-	1	
5	-	-	-	-	-	-	5	-	5	-	-	-	-	-	-	-	-	-	2	
6	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	
7	-	-	5	-	5	-	5	3	-	4	-	-	5	5	5	-	-	-	8	
8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
9	-	5	-	5	-	5	2	4	-	5	-	5	-	2	-	-	-	-	6	
10	-	-	-	-	-	5	5	-	-	-	-	5	5	-	-	-	-	-	3	
11	-	-	-	-	-	5	5	-	-	-	-	-	-	-	-	-	-	-	2	
12	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	1	1	
13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
14	-	-	-	-	-	5	5	-	1	-	-	5	-	5	5	5	-	-	8	
15	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	5	5	-	2	
$\bar{X} = 2.9$																				

N = 15 pacientes, R = 5 entrev., Observaciones = 75 posibles.

En las tablas II y III aparece la frecuencia con la que el grupo de entrevistadores observó las categorías diagnósticas en cada paciente. Para calcular la media de diagnósticos se contaron aquellas categorías en las que se hubieran hecho 3 o más observaciones. Sin aplicar los criterios de ex-

clusión del DSM-III, el promedio obtenido fue de 2.9, el cual se redujo al 1.8 al aplicar dichos criterios.

En la tabla IV se presenta la frecuencia observada por el total de los entrevistadores para cada categoría del DIS, sin aplicar los criterios de exclusión. Las categorías diag-

TABLA III
Distribución de las observaciones positivas captadas con el DIS, aplicando los criterios de exclusión del DSM-III

Paciente	Somatización	Pánico	Fobia		Agorafobia	Ansiedad generalizada	Depresión mayor	Trastorno distímico	Manía	Esquizofrenia	Anorexia nervosa	Alcohol		Obs-comps.	Drogas		Personalidad antisocial	Disfunción psicosexual	Mini mental State	No. de DXS
			Social	Simple								Abuso	Dependencia		Abuso	Dependencia				
1	-	-	-	-	-	-	5	5	-	-	-	5	-	-	-	5	-	-	4	
2	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-	1	
3	-	-	-	-	-	-	-	-	-	4	-	-	-	1	-	-	1	-	1	
4	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-	1	
5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
7	-	-	-	-	-	-	-	-	-	4	-	-	5	1	5	-	-	-	4	
8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
9	-	-	-	-	-	-	-	-	-	5	-	5	-	-	-	-	-	-	2	
10	-	-	-	-	-	-	5	-	-	-	-	5	5	-	-	-	-	-	3	
11	-	-	-	-	-	-	5	-	-	5	-	-	-	-	-	-	-	-	1	
12	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	1	
13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
14	-	-	-	-	-	5	-	-	-	-	-	5	5	5	5	5	-	-	7	
15	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	5	5	-	2	
$\bar{X} = 1.8$																				

N = 15 pacientes, R = 5 entrev., Observaciones = 75 posibles.

nósticas más frecuentes fueron: depresión mayor, abuso de alcohol, ansiedad generalizada, personalidad antisocial y esquizofrenia. Las categorías con menos observaciones fueron: trastorno de pánico, trastornos fóbicos, dependencia a las drogas y disfunciones psicosexuales; no se encontró ninguna observación en somatización y anorexia nervosa.

En la tabla V se presentan los índices de concordancia respecto a la ausencia y presencia de categorías diagnósticas y el coeficiente de correlación intraclase para cada una, sin aplicar los criterios de exclusión DSM-III. Para calcular la media, se consideraron sólo las categorías con observaciones en el total, obteniéndose los siguientes resultados: IA (a) = .986; IA (p) = .931 e ICC = .889, con una proporción de acuerdos totales de 1.0 = 62.5%.

En la tabla VI se presentan los índices de concordancia en cada categoría, aplicando los criterios de exclusión DSM-III. Para calcular las medias se procedió en la misma forma que en la tabla anterior. Las medias obtenidas fueron las siguientes: IA (a) = .996; IA (p) = .965 e ICC = .879, con 75% de acuerdos totales de 1.0.

Comentarios

Los resultados obtenidos indican un alto nivel de confiabilidad de los entrevistadores en el manejo del instrumento. Ahora bien, podríamos preguntarnos a qué se debe el alto grado de acuerdo obtenido. Grove y cols (12) en base a una exhaustiva revisión de estudios de confiabilidad, se han cuestionado si los altos índices de concordancia obtenidos se deben al diseño de las cédulas de entrevista estructuradas, al criterio diagnóstico, o a ambos. Al respecto coincidimos con Caetano (5), en cuanto a que ambos aspectos influyen en la magnitud de los resultados, ya que tanto el tipo de entrevista como el criterio diagnóstico empleado son fuentes potenciales de variación, y su estricto control incide en una buena confiabilidad.

TABLA IV
Frecuencia observada en cada categoría del DIS por todos los entrevistadores, sin aplicar criterios de exclusión del DSM-III

Categoría diagnóstica	p*	%
Depresión mayor	37	49.33
Abuso de alcohol	30	40.00
Ansiedad generalizada	20	26.67
Personalidad antisocial	20	26.67
Esquizofrenia	18	24.00
Obsesiones-compulsiones	17	22.67
Dependencia al alcohol	15	20.00
Trastorno distímico	12	16.00
Manía	10	13.33
Abuso de drogas	10	13.33
Trastorno de pánico	5	6.67
Fobia simple	5	6.67
Fobia social	5	6.67
Agorafobia	5	6.67
Dependencia a las drogas	5	6.67
Disfunciones psicosexuales	1	1.33
Somatización	—	—
Anorexia nervosa	—	—

* Total de observaciones posibles = 75; 15 pacientes, 5 entrevistadores.

TABLA V
Acuerdo en cuanto a la ausencia y presencia de las categorías diagnósticas y el coeficiente de correlación intraclase (ICC), sin aplicar criterios de exclusión del DSM-III

Categoría diagnóstica	IA (a)	IA (p)	ICC
Somatización	1.0	0*	—
Trastorno de pánico	1.0	1.0	1.0
Fobia simple	1.0	1.0	1.0
Fobia social	1.0	1.0	1.0
Agorafobia	1.0	1.0	1.0
Ansiedad generalizada	1.0	1.0	1.0
Depresión mayor	.950	.946	.9199
Trastorno distímico	.900	.417	.7519
Manía	.970	.818	.7992
Esquizofrenia	.965	.900	.8538
Anorexia nervosa	1.0	0*	—
Abuso de alcohol	1.0	1.0	1.0
Dependencia al alcohol	1.0	1.0	1.0
Obsesiones-compulsiones	.967	.882	.8859
Abuso de drogas	1.0	1.0	1.0
Dependencia a las drogas	1.0	1.0	1.0
Personalidad antisocial	1.0	1.0	1.0
Disfunciones psicosexuales	.987	0*	.0135
\bar{X} =	.986	.931	.889

* p = NS

Otro aspecto que plantean Grove y cols (12) cuestiona el que los entrevistadores con poca experiencia clínica puedan obtener altos grados de concordancia mediante un entrenamiento intensivo. Al respecto, los estudios llevados a cabo (7, 23, 24) con entrevistadores que no son psiquiatras capacitados en el empleo del PSE (*Present State Examination*, versión corta) mostraron que, mediante un entrenamiento de una o dos semanas con pacientes psi-

TABLA VI
Concordancia en cuanto a la ausencia y presencia de las categorías diagnósticas y el coeficiente de correlación intraclase (ICC), aplicando criterios de exclusión del DSM-III

Categoría diagnóstica	IA (a)	IA (p)	ICC
Somatización	1.0	0*	—
Trastorno de pánico	1.0	0*	—
Fobia simple	1.0	0*	—
Fobia social	1.0	1.0	1.0
Agorafobia	1.0	0*	—
Ansiedad generalizada	1.0	1.0	1.0
Depresión	1.0	1.0	1.0
Trastorno distímico	1.0	1.0	1.0
Manía	1.0	0*	—
Esquizofrenia	.965	.900	.8538
Anorexia nervosa	1.0	0*	—
Abuso de alcohol	1.0	1.0	1.0
Dependencia al alcohol	1.0	1.0	1.0
Obsesiones-compulsiones	.971	.714	.6848
Abuso de drogas	1.0	1.0	1.0
Dependencia a las drogas	1.0	1.0	1.0
Personalidad antisocial	1.0	1.0	1.0
Disfunciones psicosexuales	.987	0*	.0135
\bar{X} =	.996	.965	.879

* p = NS

quiátricos cuyo grado de perturbación era severo, se han podido alcanzar niveles de confiabilidad satisfactorios. Sin embargo, hay una diferencia, ya que el PSE es el tipo de entrevista en el que interviene el juicio clínico, en tanto que en el DIS, éste no desempeña un papel importante, por estar controlado por medio del esquema de sondeo. Pero en ambos casos un curso de entrenamiento y capacitación en el manejo de los instrumentos, permite que el personal que no es psiquiatra obtenga niveles altos de confiabilidad.

En nuestro estudio, los trastornos psiquiátricos mayores se captaron frecuentemente en toda la muestra; no se encontró información suficiente para integrar las categorías de somatización y anorexia *nervosa*, según lo estipula el DSM-III, y sólo escasas observaciones respecto a disfunciones psicosexuales, dependencia a las drogas, trastorno de pánico, fobia simple, fobia social y agorafobia. Estos resultados son similares a los del estudio llevado a cabo por Burnam y cols (3) en EU en pacientes de consulta externa psiquiátrica, empleando la versión en español del DIS, en el que informaron no haber encontrado ningún caso de anorexia *nervosa* y una mínima frecuencia de manía, dependencia a las drogas y somatización, por lo que no incluyeron estas categorías en los análisis de confiabilidad.

Por otro lado, en este estudio el promedio de diagnósticos por paciente fue de 2.9, similar a lo que informaron Helzer J E y cols (13), quienes encontraron una media de 2.7 al aplicar el *Renard Diagnostic Interview*, que es el antecedente directo del DIS.

Dado que el diseño metodológico de Burnam y cols (3) y el nuestro son diferentes en la prueba y en su repetición, y en entrevistador/observador, respectivamente, la comparación de los resultados es un punto que debe ser objeto de algunas consideraciones. Aunque ambos métodos se usan en la evaluación de confiabilidad entre puntuadores, éstos no son equivalentes y proporcionan diferentes estimaciones de confiabilidad (14). En este sentido, Kendell y cols (17) afirman que en el diseño de entrevistador/observador los índices de concordancia son frecuentemente más altos, ya que las variaciones en la conducta del entrevistador —tales como el estilo de presentar las preguntas y la influencia de la edad y el sexo— son eliminadas al incluir en los análisis las puntuaciones independientes del observador. De aquí que los índices de confiabilidad obtenidos en nuestro estudio sean más altos que los reportados por Burnam y cols (3).

Asimismo, cabe mencionar que en ambos diseños metodológicos se han señalado algunas desventajas. En el diseño de la prueba y su repetición, las variables que influyen en la evaluación de las puntuaciones son: el tiempo que transcurre de una entrevista a otra, los efectos del tratamiento, la disposición del entrevistado y el estilo, así como el juicio de los entrevistadores. En el caso particular del DIS, el juicio de los entrevistadores está controlado y el estilo tan sólo variará de acuerdo con la personalidad de los entrevistadores; quedan, por tanto, como fuentes de variación, los otros aspectos citados, entre los cuales es importante destacar la disposición del entrevistado. Respecto al diseño de entrevistador/observador, es posible cuestionar el grado de independencia de los registros, ya que como informó Fleiss (10), durante el transcurso de la entrevista y por la estructura del instrumento, el entrevistador puede manifestar conductas no verbales que pro-

porcionen indicios a los observadores acerca de cómo está calificando. Es por esto que Fleiss considera que el procedimiento adecuado y óptimo sería la combinación de ambos diseños.

En este estudio seleccionamos el diseño entrevistador/observador en función de los recursos económicos, materiales y humanos disponibles, asumiendo las limitaciones que implica el propio diseño.

Ahora bien, como ya lo hemos mencionado, la disposición del entrevistado para contestar a las preguntas, siempre constituye una fuente de variación en ambos diseños, máxime cuando no interviene el juicio clínico del entrevistador. Esta situación se hizo evidente en el caso del paciente no. 13 (tabla II), quien no dio ninguno de los síntomas requeridos para cubrir algún criterio del DSM-III, argumentando, al final de la entrevista, que dicha información sólo se la proporcionaba a su médico. Este paciente estaba hospitalizado por ser dependiente al alcohol, según el diagnóstico emitido por el médico que lo atiende.

En los pacientes nos. 6 y 8 —también hospitalizados— (tabla II) tampoco se integró ninguna categoría diagnóstica. El paciente no. 6 padecía un trastorno de personalidad antisocial, de acuerdo con el diagnóstico consignado por el médico que lo trataba. Respecto a este paciente podemos considerar que las propias características de dicho trastorno influyeron en la información que proporcionó durante la entrevista. Si bien manifestó algunos síntomas, éstos eran principalmente de la sección de personalidad antisocial, pero no fueron suficientes para integrar su diagnóstico en términos del DSM-III, pues sólo manifestó tener 1 síntoma de los 3 o más requeridos en el criterio B y sólo 2 síntomas de los 4 o más requeridos en el criterio C.

Con el paciente no. 8, cuyo diagnóstico psiquiátrico era el de síndrome orgánico cerebral (SOC), sucedió algo similar, ya que si bien manifestó algunos síntomas, éstos no fueron suficientes para integrar ningún diagnóstico. Dichos síntomas se refieren a alteraciones psicóticas y afectivas, así como a algunos aspectos considerados en el

TABLA VII
Confiabilidad en la prueba y en su repetición del DIS en español (sin criterios de exclusión) en pacientes de habla hispana

<i>Diagnóstico DIS</i>	<i>K</i>
Esquizofrenia	.66
Depresión mayor	.49
Manía	. . .
Trastorno distímico	.16
Ansiedad generalizada	.52
Trastorno de pánico	.45
Obsesiones-compulsiones	.62
Agorafobia	.49
Fobia simple	.25
Fobia social	.74
Abuso y dependencia al alcohol	.63
Abuso y dependencia a las drogas	. . .
Somatización	. . .
Personalidad antisocial	.92

Fuente: Burnam M A y cols.: *The Spanish Diagnostic Interview Schedule. Reliability and comparison with clinical diagnoses. Arch Gen Psychiatry 40:1189-1196, 1983*

trastorno de personalidad antisocial, tales como: bajo rendimiento escolar, problemas de conducta, vagancia y riñas hasta los 18 años, y, posteriormente, sólo inconstancia laboral. De acuerdo con la experiencia clínica, estos hallazgos frecuentemente están asociados con el SOC. Por otra parte, la valoración de esta categoría diagnóstica en el DIS está fundamentalmente orientada a los trastornos demenciales y a los síndromes confusionales actuales, por lo que es reducido el espectro para detectar a los pacientes que sufren alteraciones orgánicas.

Otro de los pacientes que llama la atención es el número 5 —también hospitalizado— en el que se integraron los diagnósticos de depresión mayor y manía cuando no se aplicaron los criterios de exclusión DSM-III. Cuando dichos criterios se aplicaron no conformaron ninguna categoría. Esto se debió a que el paciente manifestó preocupación delirante, incongruente con el estado de ánimo depresivo y maniaco, y según la definición del DSM-III, esto constituye un criterio de exclusión de los diagnósticos antes mencionados.

En resumen, debemos resaltar el aspecto de la disposición de los entrevistados para responder, situación que depende tanto de la patología y de la conciencia de su enfermedad como de algunos aspectos sociales de la población, por lo que respecta al grado de sensibilidad para aceptar una entrevista de este tipo. Podemos concluir que el DIS permite asegurar un alto nivel de confiabilidad entre los entrevistadores, aún cuando su validez está por evaluarse.

Así pues, y de acuerdo con lo que plantean Spitzer y Fleiss (22) acerca de que no hay garantía de que un sistema confiable sea válido, pero que seguramente no podría serlo si no alcanza niveles satisfactorios de concordancia diagnóstica, consideramos útil el presente estudio para proseguir con la estimación de la validez.

Agradecemos la colaboración de Esther Iglesias, Guillermo Pérez y Laura Teyssier, en la aplicación de las entrevistas, así como el apoyo de los Drs. Salvador González y Carlos Salgado, para llevar a cabo este estudio.

REFERENCIAS

1. American Psychiatric Association Committee on Nomenclature and Statistics: Diagnostic and Statistical Manual of Mental Disorders. 3a Ed., Washington, D. C., 1980.
2. BARTKO J J, CARPENTER W T: On the methods and theory of reliability. *The Journal of Nervous and Mental Diseases*, 163(5): 307-317, USA, 1976.
3. BURNAM M A, KARNO M, cols: The spanish diagnostic interview schedule. Reliability and comparison with clinical diagnoses. *Arch Gen Psychiatry*, 40: 1189-1196, 1983.
4. CABILDO H M: Encuesta para detectar trastornos psíquicos. *Higiene*, 18: 188-208, 1967.
5. CAETANO R: Causas de baja confiabilidad del diagnóstico en psiquiatría. *Acta Psiquiátr Psicol Amér Lat*, 24: 115-130, 1978.
6. CARAVEO A J: Análisis crítico y perspectivas futuras de la epidemiología psiquiátrica en México. *Salud Mental*, 5(1): 13-19, México, 1982.
7. COOPER J E, COPELAND J R M, cols: Further studies on interviewer training and interrater reliability of the Present State Examination. *Psychol Med*, 7: 517-523, 1977.
8. FEIGHNER J P, ROBINS E, cols: Diagnostic criteria for use in psychiatric research. *Arch Gen Psychiatry*, 26: 57-63, 1972.
9. FLEISS J L: Estimating the accuracy of dichotomus judgments. *Psychometrika*, 30: 469-479, 1965.
10. FLEISS J L: Estimating the reliability of interview data. *Psychometrika*, 35: 143-162, 1970.
11. FLEISS J L, COHEN J: The equivalence of weighted Kappa and the Intraclass Correlation Coefficient as measures of reliability. *Educ Psychol Measuram*, 33: 613-619, 1973.
12. GROVE W M, ANDREASEN N C, cols: Reliability studies of psychiatric diagnosis. *Arch Gen Psychiatry*, 38: 408-413, 1981.
13. HELZER J E, ROBINS L N, cols: Renard Diagnostic Interview. Its reliability and procedural validity with physicians and lay interviewers. *Arch Gen Psychiatry*, 38: 393-405, 1981.
14. HELZER J E, ROBINS L N, cols: Reliability of psychiatric diagnosis. *Arch Gen Psychiatry*, 34: 129-133, 1977.
15. KENDELL R E: Defining diagnostic criteria for research puposes. En: *Methods of Psychiatric Research*. Sainsbury and Kreitman (eds.). Londres, Oxford University Press, 1975.
16. KENDELL R E: The issues of reliability and validity. En: *The Role of Diagnosis in Psychiatry*. Blackwell Scientific Publications. Londres, Aberdeen University Press, 27-48, 1976.
17. KENDELL R E, EVERETT J E, cols: The reliability of the Present State Examination. *Soc Psychiatry*, 3: 123-129, 1968.
18. RAMOS L L, CARAVEO A J, GONZÁLEZ F C: La Cédula de Entrevista Diagnóstica (DIS) en México. Comunicación preliminar. *Psiquiatría*. V. 3. No. 2: 133-142, 1987.
19. ROBINS L N, HELZER J E, cols: National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics and validity. *Arch Gen Psychiatry*, 38: 381-389, 1981.
20. ROBINS L N, HELZER J E, cols: Validity of the Diagnostic Interview Schedule, version II: DSM-III diagnoses. *Psychol Med*, 12: 855-870, 1982.
21. SPITZER R L, ENDICOTT J, ROBINS E: Research Diagnostic Criteria: Rationale and reliability. *Arch Gen Psychiatry*, 35: 773-782, 1978.
22. SPITZER R L, FLEISS J L: A re-analysis of the reliability of psychiatric diagnosis. *Br J Psychiatry*. 125: 341-347, 1974.
23. STURT E, BEBBINGTON P, cols: The Present State Examination used by interviewers from a survey agency: report from the MRC Camberwell Community Survey. *Psychol Med*, 11: 185-192, 1981.
24. WING J K, NIXON J M, cols: Reliability of the PSE (ninth edition) used in a population study. *Psychol Med*, 7: 505-516, 1977.