

Los ensayos de campo iniciales del DSM-5: nuevas flores y viejas espinas

“Una rosa es una rosa es una rosa” (1). Para el diagnóstico psiquiátrico, todavía interpretamos este verso como hicieron Robins y Guze en sus Criterios Diagnósticos en Investigación (*Research Diagnostic Criteria*): la fiabilidad es la primera prueba de la validez del diagnóstico (2). Para desarrollar una psiquiatría basada en datos científicos, en el DSM-III y el DSM-IV se adoptó la estrategia de Robins y Guze (es decir, criterios validados empíricamente para los signos y síntomas reconocibles de enfermedad). Los resultados iniciales de fiabilidad de los ensayos de campo del DSM-5 se publican ahora en tres artículos de este número (3–5). Al igual que en todas las ediciones previas del DSM, los métodos utilizados a fin de evaluar la fiabilidad reflejan los estándares actuales de la investigación psiquiátrica (3). En los ensayos de campo del DSM se emplean ahora por primera vez entrevistas independientes realizadas por dos médicos diferentes expertos en los diagnósticos, cada uno ayudado por una lista de verificación computarizada, una evaluación de la coincidencia entre diferentes centros académicos y un plan estadístico preestablecido. Al igual que en la mayor parte de los nuevos intentos, los resultados finales son desiguales, con datos tanto positivos como decepcionantes.

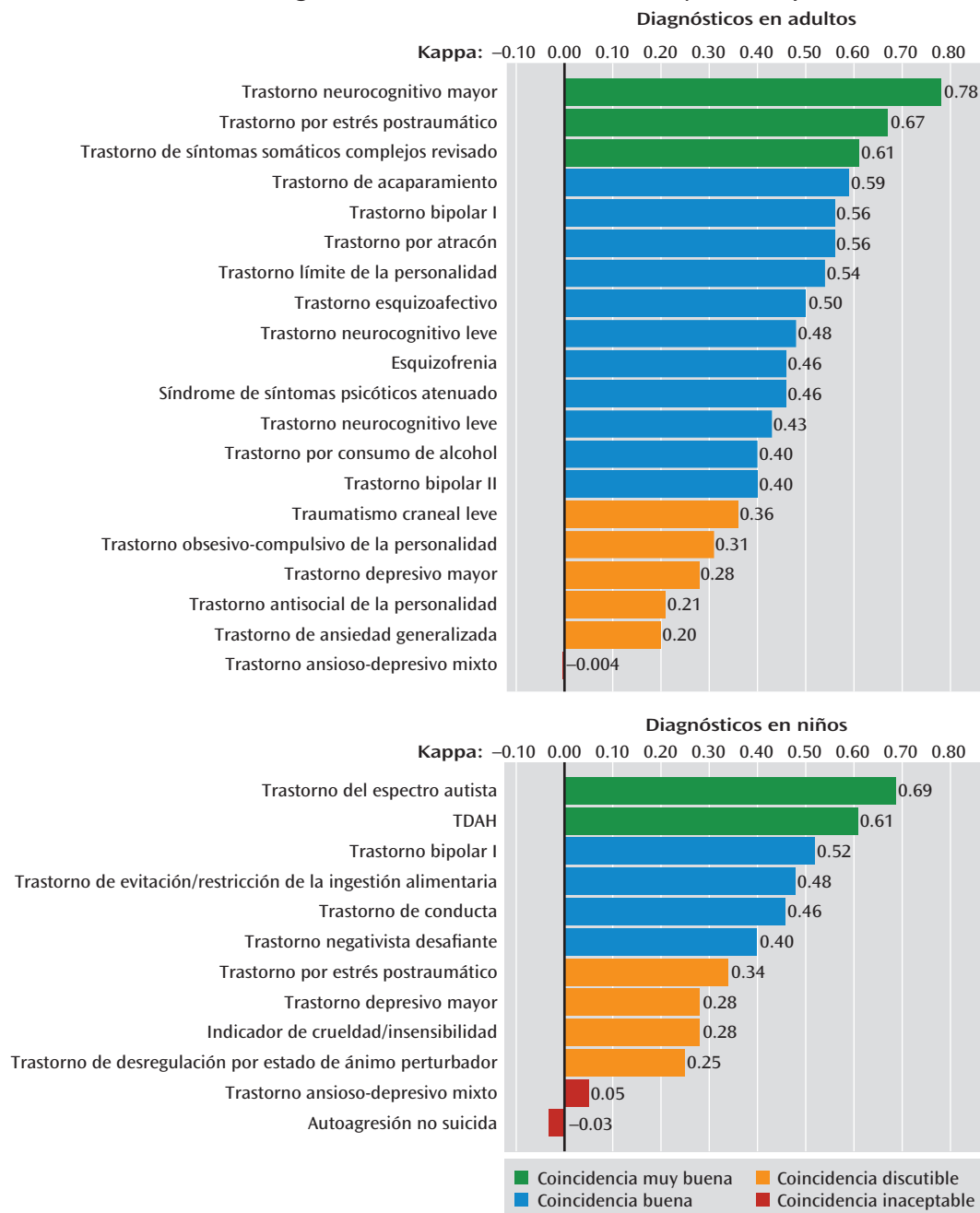
Es posible que el estadístico kappa utilizado para el análisis no resulte familiar para la mayoría de los médicos. A título de ejemplo, si una enfermedad aparece en el 10% de los pacientes de un médico y dos colegas están de acuerdo con su diagnóstico en el 85% de las ocasiones, el estadístico kappa es de 0.46,

similar al estadístico compuesto ponderado para la esquizofrenia en este ensayo de campo del DSM-5 (Figura 1). Los criterios de la esquizofrenia se modificaron radicalmente en el DSM-III y se volvieron a modificar en el DSM-IV debido a las discrepancias mundiales en su diagnóstico. En la actualidad, el problema para diferenciar esquizofrenia, trastorno bipolar y trastorno esquizoafectivo—el *quid* de las discrepancias—se ha resuelto en gran medida y los tres trastornos tienen un buen estadístico kappa.

La discutible fiabilidad del trastorno depresivo mayor, que no se ha modificado con respecto al DSM-IV, constituye obviamente un problema. El trastorno depresivo mayor ha sido siempre problemático porque sus criterios engloban una amplia gama de enfermedades: desde los pacientes melancólicos muy discapacitados hasta muchas personas de la población general que no requieren tratamiento. Mientras que la intensidad de los síntomas en la Escala de valoración de la depresión de Hamilton (*Hamilton Depression Rating Scale*) permite distinguir a aquellos pacientes que responden más específicamente a la farmacoterapia, los criterios del DSM-IV no reflejan esa distinción (6). Un segundo problema no resuelto por los criterios del DSM-IV es la habitual concurrencia de la ansiedad, que disminuye marcadamente los efectos del tratamiento antidepresivo (7). El grupo de trabajo del

Una sola entrevista diagnóstica, con independencia de su fiabilidad, no captura la esencia de lo que está ocurriendo... El diagnóstico exacto debe ser parte de un diálogo clínico continuo con el paciente

FIGURA 1. Fiabilidad de los diagnósticos entre evaluadores de los ensayos de campo iniciales del DSM-5^a



^a Algunos de los estadísticos kappa no superaron el criterio de un error estándar inferior a 0.1. Se incluyen aquí con fines ilustrativos. Para más detalles véanse los informes de los ensayos sobre el terreno (3–5).

DSM-5 decidió no modificar los criterios del trastorno depresivo mayor del DSM-IV y, en su lugar, creó otros diagnósticos para la combinación de ansiedad y depresión. Sin embargo, estos esfuerzos no mejoraron la escasa fiabilidad de la depresión del DSM-IV; “ansiedad y depresión mixtas” tiene un valor kappa de 0. Los médicos a menudo utilizan como un indicador de gravedad la autoevaluación de los pacientes mediante el Inventario de la depresión de Beck (*Beck Depression Inventory*). Los dominios transversales dimensionales en este ensayo de campo se basan asimismo en la autoevaluación (5). Para la depresión existen dos dominios y las correlaciones intraclase—que son similares al estadístico kappa—en los pacientes adultos que

se autoevalúan y repiten la autoevaluación y en los progenitores que evalúan a sus hijos; superan todas ellas el 0.6. En futuras revisiones probablemente será necesario integrar los numerosos factores—autoevaluaciones de los pacientes, sesgos cognitivos, ansiedad concurrente y síntomas vegetativos—que orientan la elección del tratamiento, el pronóstico y la evaluación del riesgo de suicidio.

Los médicos experimentados tienen importantes reservas acerca del esquema diagnóstico de investigación propuesto para el trastorno de la personalidad, cuya aplicabilidad a la práctica clínica todavía está por determinarse (8). La mayor parte de los diagnósticos de trastorno de la personalidad no funcionaron bien en el ensayo de campo. Los trastornos antisocial y obsesivo-compulsivo de la personalidad tuvieron una fiabilidad discutible o no concluyente y otros tipos, como los trastornos narcisista y esquizotípico de la personalidad, no fueron lo bastante frecuentes como para poder evaluarlos. El éxito del trastorno límite de la personalidad constituye, no obstante, un gran avance. En el DSM-III se relegaron la mayor parte de los trastornos de la personalidad al eje II, lo que cortó radicalmente una de las raíces más venerables de la psiquiatría. Pero los médicos reconocieron que la patología del carácter, a pesar de su aparente estabilidad, era a la vez bastante discapacitante y dócil al tratamiento. El trastorno límite de la personalidad surge ahora como un diagnóstico principal por derecho propio, con buena fiabilidad diagnóstica.

El estado de ánimo inestable, una característica primordial del trastorno límite de la personalidad en la edad adulta, es también la característica destacada en la infancia de un nuevo trastorno que se ha denominado *disruptive mood dysregulation disorder* (trastorno de desregulación por estado de ánimo perturbador). Este trastorno tiene un estadístico kappa más pequeño. El trastorno de desregulación por estado de ánimo perturbador se evaluó con más fiabilidad en el ámbito hospitalario, donde se examinó, al igual que el trastorno límite de la personalidad, en sus inicios. Es posible que a medida que aumente la experiencia clínica con este nuevo diagnóstico infantil mejore su rendimiento diagnóstico. La fiabilidad de los diagnósticos de TDAH y de trastorno bipolar en la infancia, que han sido problemáticos sobre todo en presencia de irritabilidad, probablemente se haya beneficiado con la alternativa del trastorno de desregulación por estado de ánimo perturbador; ambos tienen un buen estadístico kappa. El trastorno del espectro autista, recientemente reorganizado y que ha sido también objeto de un gran debate previo, tiene un valor kappa muy bueno, aunque en los ensayos no se incluyó a niños menores de seis años.

El trastorno por estrés postraumático (TEPT) es otro logro histórico, con un valor kappa de 0.67. La serie del DSM se inició porque “era necesario aplicar la ‘etiqueta de psiconeurótico’ a las personas que reaccionan brevemente con síntomas neuróticos ante un estrés considerable; personas que... no eran habitualmente psiconeuróticas” (9). Cuatro ediciones y 60 años más tarde, el TEPT es actualmente un diagnóstico fiable de un trastorno que se podría haber desestimado como comportamiento normal patologizador. Se han introducido otros diagnósticos nuevos o se han redefinido con buena fiabilidad: trastorno neurocognitivo mayor, trastorno de acaparamiento, trastorno de síntomas somáticos complejos y trastorno por atracción, además de los ya comentados.

En los ensayos de campo era necesario llegar a un diagnóstico a partir de una sola entrevista con el paciente y una información colateral mínima. En la práctica psiquiátrica general, los datos de fiabilidad diagnóstica—combinación de la Tabla 1 (muestras ponderadas en un entorno ambulatorio de pacientes adultos, incluyendo el TDAH infantil en la categoría de “otros diagnósticos”) y de las Tablas 2 y 4 (fi-

abilidad de los diagnósticos en adultos y niños [4])—indican que dos tercios de los pacientes tendrán un diagnóstico principal fiable basado en el DSM-5 en la primera visita. Estos diagnósticos habituales fiables son TDAH infantil, TEPT, trastorno límite de la personalidad y trastorno por consumo de alcohol. En el tercio de los pacientes con un traumatismo craneal leve o un trastorno depresivo mayor es posible que no se llegue a un diagnóstico fiable con una sola entrevista. Desde luego, este cálculo será diferente en cada ámbito clínico. Robins y Guze introdujeron una categoría “no diagnosticada” para insistir en que los pacientes fueran examinados de nuevo, más adelante, cuando sus síntomas iniciales no dieran lugar a un diagnóstico inequívoco. En los ensayos de campo del DSM-5 no se examinó el aumento de fiabilidad derivado de que el mismo médico encargado del tratamiento evaluara al paciente a lo largo del tiempo a medida que desarrollara la enfermedad.

“Una rosa es una rosa es una rosa es una rosa” tiene un significado más profundo para Gertrude Stein, en relación no sólo con la clasificación de la flor, sino también con su esencia imperecedera (10). El conocimiento de la evolución natural de una enfermedad, de su respuesta al tratamiento y de su repercusión en la vida de la persona son las razones que aducimos para hacer diagnósticos fiables, pero una sola entrevista diagnóstica, con independencia de lo fiable que sea, no capta la esencia de lo que le está ocurriendo al paciente. De las posibles lecciones para los médicos y para los pacientes y sus familiares que lean estos ensayos de campo, quizá la más importante sea que el diagnóstico exacto debe ser parte del diálogo clínico continuo con el paciente.

La mejoría del diagnóstico también es constante. Las pruebas futuras deben tener en cuenta la utilidad clínica en situaciones terapéuticas reales y la fiabilidad y viabilidad de aplicar los nuevos criterios fuera de centros médicos académicos. Los médicos que trabajan solos y los consultorios de salud mental quizá no posean recursos para el nivel de formación que requieren los ensayos de campo. Uno de los requisitos era que los pacientes hablaran y leyeran inglés, aunque algunos eran bilingües. La fiabilidad quizá no sea la misma en los pacientes con un nivel de estudios más bajo o que no tengan un dominio completo del inglés. Los datos de estos ensayos de campo se utilizarán a fin de introducir nuevas mejoras, por lo que los criterios finales pueden cambiar y requerir nuevas valoraciones después de la publicación del DSM-5. Al igual que sus predecesores, el DSM-5 no consigue todo lo que pretendía, aunque sí marca un progreso continuo para muchos pacientes en los que no se habían cumplido previamente los efectos benéficos de los diagnósticos y del tratamiento.

Bibliografía

1. Goodwin DW: Preface, in *Psychiatric Diagnosis*. By Woodruff RA, Goodwin DW, Guze SB. New York, Oxford University Press, 1974
2. Robins E, Guze SB: Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. *Am J Psychiatry* 1970; 126:983–987
3. Clarke DE, Narrow WE, Regier DA, Kuramoto SJ, Kupfer DJ, Kuhl EA, Greiner L, Kraemer HC: DSM-5 Field Trials in the United States and Canada, part I: study design, sampling strategy, implementation, and analytic approaches. *Am J Psychiatry* 2013; 170:43–58
4. Regier DA, Narrow WE, Clarke DE, Kraemer HC, Kuramoto SJ, Kuhl EA, Kupfer DJ: DSM-5 Field Trials in the United States and Canada, part II: test-retest reliability of selected categorical diagnoses. *Am J Psychiatry* 2013; 170:59–70
5. Narrow WE, Clarke DE, Kuramoto SJ, Kraemer HC, Kupfer DJ, Greiner L, Regier DA: DSM-5 Field Trials in the United States and Canada, part III: development and reliability testing of a cross-cutting symptom assessment for DSM-5. *Am J Psychiatry* 2013; 170:71–82
6. Fournier JC, DeRubeis RJ, Hollon SD, Dimidjian S, Amsterdam JD, Shelton RC, Fawcett J: Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA* 2010; 303:47–53

7. Fava M, Rush AJ, Alpert JE, Balasubramani GK, Wisniewski SR, Carmin CN, Biggs MM, Zisook S, Leuchter A, Howland R, Warden D, Trivedi MH: Difference in treatment outcome in outpatients with anxious versus nonanxious depression: a STAR*D report. *Am J Psychiatry* 2008; 165:342–351
8. Shedler J, Beck A, Fonagy P, Gabbard GO, Gunderson J, Kernberg O, Michels R, Westen D: Personality disorders in DSM-5. *Am J Psychiatry* 2010; 167:1026–1028
9. American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC, American Psychiatric Association, 1952
10. Stein G: "Sacred Emily," in *Geography and Plays*. Madison, University of Wisconsin Press, 1922 (reissued 1993)

ROBERT FREEDMAN, M.D.
DAVID A. LEWIS, M.D.
ROBERT MICHELS, M.D.
DANIEL S. PINE, M.D.
SUSAN K. SCHULTZ, M.D.
CAROL A. TAMMINGA, M.D.
GLEN O. GABBARD, M.D.
SUSAN SHUR-FEN GAU, M.D., PH.D.
DANIEL C. JAVITT, M.D., PH.D.
MARIA A. OQUENDO, M.D., PH.D.
PATRICK E. SHROUT, PH.D.
EDUARD VIETA, M.D., PH.D.
JOEL YAGER, M.D.

De la Oficina de Redacción de The American Journal of Psychiatry. Diríjase la correspondencia al Dr. Freedman (ajp@psych.org). Editorial aceptado para su publicación en septiembre de 2012 (doi: 10.1176/appi.ajp.2012.12091189).

*Los autores son redactores o fueron invitados por estos para colaborar en este editorial. Algunos desempeñan otros cometidos en el proceso del DSM-5. El Dr. Freedman es copresidente del Comité de Revisión Científica, el Dr. Pine es presidente del grupo de trabajo de Trastornos Infantiles, la Dra. Schultz es miembro del grupo de trabajo de Trastornos Geriátricos y redactora del texto y el Dr. Yager es copresidente del Comité de Revisión Clínica y de Salud Pública. Tanto el Dr. Yager como el Dr. Freedman son también miembros del Grupo de Trabajo en la Cumbre, que hace recomendaciones finales al Consejo de Administración de la American Psychiatric Association. El Dr. Gabbard es director jefe de la 5ª edición de *Treatments of Psychiatric Disorders* (Tratamientos de los trastornos psiquiátricos), que será publicado por la American Psychiatric Publishing como el libro de tratamiento inicial del DSM-5. Las declaraciones de intereses económicos del redactor y de los redactores adjuntos se publican todos los años en el número de enero. La Dra. Gau ha recibido honorarios como ponente y fondos para viáticos de Eli Lilly; ha sido investigadora en un ensayo clínico patrocinado por Eli Lilly y ha recibido honorarios como ponente de AstraZeneca y Janssen. El Dr. Javitt ha recibido becas para la investigación de Jazz Pharmaceuticals, Pfizer y Roche; ha actuado como asesor de AstraZeneca, Bristol-Myers Squibb, Cypress, Lilly, Lundbeck, Merck, NPS, Pfizer, Sanofi, Schering-Plough, Sepracor, Solvay, Takeda y Sunovion; participa en el comité asesor de Promentis Pharmaceuticals y posee acciones de Glytech, Inc. La Dra. Oquendo ha recibido becas ilimitadas para la educación u honorarios como ponente de AstraZeneca, Bristol-Myers Squibb, Eli Lilly, Janssen, Otsuka, Pfizer, Sanofi-Aventis y Shire; posee acciones de Bristol-Myers Squibb y recibe pagos por derechos de autor por eC-SSRS de ERT, Inc. El Dr. Vieta se ha desempeñado como asesor y ponente y recibido apoyo para la investigación de AstraZeneca, Bristol-Myers Squibb y Forest y ha trabajado como ponente para GlaxoSmithKline. Los restantes autores refieren no tener relaciones económicas con intereses comerciales.*