

# DSM-5 FIELD TRIALS

## Overview

---

The impact of many changes in medicine, like new treatments and diagnostic guidelines, cannot be fully understood until examined in real-world settings. Scientists often turn to these settings to carry out “field trials” that test the possible interventions in advance of actual implementation.

Such was the case with the researchers behind the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*, the authoritative sourcebook used in this country and much of the world for mental health care. From 2010-2012, the *DSM-5* leaders coordinated and conducted rigorous field trials testing some of the more substantial changes proposed for the upcoming manual.

There is no single prescribed method to any field trial, but certain elements are consistent: planned and controlled testing in environments representative of those in which the product would be used; participants like those who would utilize or benefit from the product; and unbiased assessment of the results.

The *DSM-5* field trials carried out by the American Psychiatric Association (APA) incorporated all three elements. The effort involved mental health professionals and patients in diverse clinical settings nationwide and was aimed at providing a real-world evaluation of certain proposed diagnostic criteria to ensure accurate and useful disorder definitions and criteria.

The field trials proved very important to understanding the effectiveness of proposed revisions and providing guidance as final revisions were made to criteria.

## The Overarching Issues

---

The field trials were a milestone in the *DSM-5* development process. They were conducted after more than a decade of research review, scientific conferences, expert analysis of current and potential disorders and criteria, public comment, and extensive planning for the testing itself. They were informed by the *DSM-III* and *DSM-IV* field trials. The results indicate not only the strength of the process but of the diagnostic advances that have followed earlier *DSMs*.

The field trials looked at how proposed diagnostic criteria performed in clinical settings with what could be thought of as typical clinicians and patients. Clinicians were not provided any special training in the draft criteria and patients often presented a full range of disorders and comorbidities.

Out of more than 200 disorders in *DSM*, about two dozen were chosen for testing, including disorders with high clinical or public health importance, disorders with major proposed changes or newly proposed disorders. One disorder with unchanged criteria

was included as reference disorder to examine the effect of the design of the *DSM-5* trials compared to that of the *DSM-IV* trials two decades earlier.

The trials evaluated the feasibility, clinical utility, reliability, and, to some extent, the validity of individual criteria. Feasibility and clinical utility measured if the proposed definitions and language were understandable and easy to use and if they accurately described patients' symptoms and helped clinicians to make diagnostic decisions. Reliability evaluated how consistently the same diagnostic conclusions were reached by different clinicians using the same criteria to evaluate a patient.

The trials also looked at proposed diagnostic-specific severity measures as well as cross-cutting dimensional measures intended to help identify common symptoms such as depressed mood or anxiety. These assessment tools, including self-administered patient questionnaires, were judged on whether they provided useful information for clinicians and patients and whether they captured changes in symptoms over time that would indicate progress in treatment.

## *The Trials' Design*

---

Two complementary study designs were used to test criteria in the range of settings where *DSM* is commonly used: the first in 11 large academic medical centers and the second in hundreds of smaller, routine clinical practice settings.

About 3,500 child and adult patients took part in both field trials designs—nearly 2,100 at the large academic medical centers and more than 1,400 through routine clinical practice settings.

Both designs incorporated research methods that were not used or, in some instances, not available during past *DSM* field trials. These included test-retest reliability measures in the academic sites (as opposed to inter-rater reliability) and computer-assisted diagnostic checklists to aid routine clinical assessments. Patients were asked to rate current symptoms in 12 psychological domains.

Finally, data were transferred immediately through a secure website to a central server to ensure independence of assessments and data integrity.

## *Field Trials in Large Academic Medical Centers*

---

The first design focused on large academic medical centers, with 11 centers from across the country selected to participate. There were four pediatric sites in Massachusetts, New York, Colorado and California. The seven adult sites were located in California, Texas (three sites), Minnesota, Pennsylvania, and Toronto, Canada. Each location evaluated a different set of two to five target diagnoses commonly seen at that site. The same diagnosis was often evaluated at several different sites to examine the possibility of site differences.

Several significant factors distinguished the *DSM-5* field trials from other field trials, including past *DSM* trials:

- Reliability was measured in the most stringent way possible using test-retest methods. This involved independent interviews by two different clinicians, instead of past models in which multiple clinicians observed and rated the same interview or two clinicians conducted a joint interview.
- All symptomatic patients entering a participating clinic were eligible for the field trials. Past field testing often included only patients likely to have the disorder being studied and excluded patients with comorbidities or other confusing presentations.
- Clinicians were asked to use the clinical interview methods they would normally use to apply *DSM-5* criteria. In past field testing, clinicians often underwent extensive training on the *DSM* draft criteria or used structured research interviews to diagnose patients.
- The study was designed with a sample size adequate to obtain an accurate estimate of the reliability (to within  $\pm 0.1$ ).
- Both site differences and comorbidities were acknowledged and addressed.

### ***Results and Moving Forward***

Patient participation across the academic center sites was large enough for researchers to assess 23 adult and child/adolescent diagnoses. Several other diagnoses that were included in field trials were too rare and the sample sizes were not adequate to achieve acceptable accuracy.

Most diagnoses performed well in the field trials and could be reliably diagnosed. Results on reliability were reported in terms of a disorder's "kappa," a statistical measure indicating how much of the individual differences among diagnoses are not accounted for by random error. Based on this measurement, the reliability of the disorders tested was judged:

- **Very good:** Five diagnoses were evaluated with very good reliability, including major neurocognitive disorder; autism spectrum disorder; posttraumatic stress disorder, or PTSD; somatic symptom disorder; and attention deficit/hyperactivity disorder.
- **Good:** Nine disorders ranked with good reliability, including schizophrenia; schizoaffective disorders; bipolar I disorder; binge eating disorder; alcohol use disorder; mild neurocognitive disorder; borderline personality disorder; avoidant/restrictive food intake disorder; and oppositional defiant disorder.
- **Questionable:** Six diagnoses were ranked with questionable reliability.
- **Unacceptable:** Three diagnoses had unacceptable reliability.

Of the disorders judged as having unacceptable reliability, all subsequently underwent substantial revisions or were dropped from potential inclusion in *DSM-5*.

### ***Depression and Anxiety***

Two disorders that achieved questionable reliability using the field trials' rigorous testing model were major depressive disorder and generalized anxiety disorder. Major depressive disorder's criteria had not been changed from *DSM-IV*, and only minimal

changes were made to the criteria for generalized anxiety disorder. *DSM-IV* field trials had shown good reliability for diagnosing both conditions.

The different results here reflect the fact that in *DSM-5* field trials, patients with confusing presentations were not excluded from the testing and participating clinicians were not given special training in the use of the proposed criteria. In addition, those clinicians were allowed to perform evaluations as they usually would in their offices rather than structured research instruments, as had been the case during the *DSM-IV* testing.

After reviewing the data in more detail, it seemed clear that the questionable reliability—signaling acceptable but relatively low reliability—stemmed in part from how often these conditions are comorbid with other conditions, as well as each other, and the challenges that can present for clinicians in reaching a diagnosis.

For example, PTSD often is accompanied by major depressive disorder, but many clinicians focus first on the posttraumatic stress and consider the depression a lesser priority. Consequently, they may begin to treat the PTSD and ignore the major depressive disorder. Moreover, there is usually considerable hour-to-hour or day-to-day fluctuation in how symptoms of these disorders are expressed, and this too can decrease reliability.

The results of the field trials were published in the January issue of the *American Journal of Psychiatry*<sup>i,ii,iii</sup>.

### *Field Trials in Routine Clinical Practice*

---

The second field trials design addressed smaller group practices or solo practitioners' offices and involved about 600 volunteer mental health professionals representing different disciplines, years of practice, and racial and ethnic backgrounds.

Psychologists, therapists, social workers, psychiatric nurses, counselors, psychiatrists, and other mental health clinicians participated in the routine clinical practice field trials. As clinicians who will be using *DSM-5* for direct care, their involvement was one of these field trials' particular strengths.

The aim of this design was to evaluate the feasibility and clinical utility of proposed criteria, as well as the feasibility of the proposed dimensional measures using patient surveys and clinician-administered diagnostic interviews, completion of diagnostic checklists, severity measures, and clinical utility questionnaires.

The results from this field trial will be published in 2013.

### *Looking Ahead*

---

The field trials were only one of many tools used by APA and *DSM-5* leaders to determine the next manual's final criteria. However, their results were immediately useful for providing further evaluation and direction for revisions to criteria of key

disorders, and they were extremely valuable for understanding the feasibility, clinical utility, and reliability of select disorders.

The APA expects the *DSM-5* field trials to have lasting impact and relevancy for years to come, benefiting researchers, clinicians, and, above all, patients.

*DSM* is the manual used by clinicians and researchers to diagnose and classify mental disorders. The American Psychiatric Association (APA) will publish *DSM-5* in 2013, culminating a 14-year revision process. For more information, go to [www.DSM5.org](http://www.DSM5.org).

APA is a national medical specialty society whose more than 36,000 physician members specialize in the diagnosis, treatment, prevention and research of mental illnesses, including substance use disorders. Visit the APA at [www.psychiatry.org](http://www.psychiatry.org) and [www.healthyminds.org](http://www.healthyminds.org).

---

<sup>i</sup> Am J Psychiatry 2013;170:43-58. 10.1176/appi.ajp.2012.12070998

<sup>ii</sup> Am J Psychiatry 2013;170:59-70. 10.1176/appi.ajp.2012.12070999

<sup>iii</sup> Am J Psychiatry 2013;170:71-82. 10.1176/appi.ajp.2012.12071000