VIEWPOINT

Research Domain Criteria (RDoC) and the *DSM*—Two Methodological Approaches to Mental Health Diagnosis

Helena Chmura Kraemer, PhD
Department of Psychiatry and Behavioral Sciences (Emerita), Stanford University, Palo Alto, California; and Statistical Editor, *JAMA Psychiatry*.

Viewpoints
pages 1159 and 1161
and **Editorial**
page 1165

Corresponding Author: Helena Chmura Kraemer, PhD, Department of Psychiatry and Behavioral Sciences (Emerita), Stanford University, 1116 Forest Ave, Palo Alto, CA 94301-3032 (hckhome@pacbell.net).

The distinction between the approaches represented by the *DSM* and Research Domain Criteria (RDoC) does *not* lie in the choice between categorical and dimensional diagnosis.[1] Every dimensional diagnosis can be converted to a corresponding categorical one by judiciously applying some dichotomization rule. Every categorical diagnosis can be converted to a corresponding dimensional one, for example, by requiring multiple assessments and using the percentage positive.[2]

Generally, dimensional diagnoses will yield greater power in testing hypotheses and precision in estimation of parameters,[3] since power and precision depend on the sensitivity of measures to individual differences among patients. Consequently, dimensional diagnoses are preferable in clinical research. However, in determining eligibility for a clinical trial, clinical researchers also need a categorical diagnosis. In clinical decision making, yes/no decisions, such as whether or not a patient should be given a particular medication, require categorical diagnoses. However, in tracking changes in a patient undergoing treatment, clinicians also need dimensional diagnoses. Clinical researchers and clinicians should be able to choose either a categorical or a dimensional form for the *same* diagnosis, whichever is optimal for a particular purpose. What matters is that medical decision making, to be evidence-based rather than subjective or arbitrary, must be based on clinical research. It is impractical and unwise to propose that one approach (eg, RDoC) be used for research and another (eg, the *DSM*) for clinical decision making. That would be a recipe for disaster for patients with mental health problems.

With that in mind, any methodological approach must be judged by its results, not by its promises. Because the *DSM* in its current form has a long history, we have 35 years of research and clinical results (since the *DSM-III* published in 1980) by which to judge it. RDoC approaches too have been used for many years. In 1987, I did a statistical review of the more than 200 published articles published to date on the Dexamethasone Suppression Test (DST) for psychiatric diagnosis,[4] pointing out the sources of inconsistencies and nonreproducibility of the findings, concluding that "The crucial issue is not what is now known, but what we will know in the future, not only about the DST, but also about other biological signals of psychiatric disorders."[4(p425)] Now, almost 30 years later, the situation is much as it was then. RDoC remains a set of assumptions expressed in a conceptual framework. Every time RDoC is described using terms like *will be*, *may be*, *will undoubtedly inform*, all promises, instead of statements such as "we have consistently shown that…," or another nonreproducible finding is published, concerns increase.

The *DSM* is flawed. The *DSM-5* includes dimensional diagnoses only in section 3. While some *DSM-5* diagnoses incorporate biomarkers (eg, polysomnography, hypocretin deficiency for sleep disorders), and most discuss biomarkers not yet shown to be diagnostic, the *DSM-5* diagnoses focus on observable signs and self-described symptoms. Diagnostic rules (eg, cutpoints) have often been arbitrarily set. The *DSM* process is slow, too influenced by commercial, not scientific factors. The comorbidity among *DSM* diagnoses suggests that some diagnoses have been lumped when they should have been split and others split when they should have been lumped. Progress in identifying the causes, courses, and cures for mental health disorders is slow.

An unwarranted criticism of the *DSM* is that it lacks validity. What the *DSM* lacks is proof of validity, a problem that will also affect all future RDoC-based diagnoses. However, the first requirement for validity is test-retest reliability, and the test-retest reliabilities of *DSM-5* diagnoses are comparable with those of objective medical tests for physical disorders.[5] That means that these diagnoses correspond to patterns that clinicians commonly see in practice, and that patients commonly complain of, a fact not to be ignored.

RDoC *concepts* are right—understanding the biological bases of mental disorders is crucial. However, RDoC *approaches* remain flawed. Whether the result of a brain scan or an assay result would be the same if independently done in the same patients a day later, or at a different laboratory, is seldom documented. In absence of test-retest reliability and consistency over sites, RDoC-based diagnoses cannot be valid.

RDoC studies are often based on muddling samples from different laboratories (eg, data registries) in the belief that "big data," no matter its quality, will answer questions. However, different laboratories have access to different populations, sampled in different ways, are staffed by different researchers who operate differently. If the conclusions based on analyses of data sets from individual laboratories disagree, muddling the samples from discordant studies only exacerbates the situation.

Multiple testing (full-genome scans, hundreds of brain areas) plagues RDoC studies. Statistical adjustment for *P* values does not solve this problem. *P* values are uninformative, reflecting sample sizes, distributions, and unreliability more than they do clinical importance. Requiring interpretable effect sizes and their 95% confidence intervals with every *P* value, as in recent clini-

cal research, has yet to permeate RDoC studies. Effect sizes, not *P* values, require replication.

Multiple-testing studies are better regarded as hypothesis generating rather than hypothesis testing.[6] Such a study might be used to identify the source of the strongest effect. Then 1 or more *independent* studies might be done, validly designed and adequately powered to evaluate *only* that effect. Only when independently confirmed by the consensus of such studies would any result be reported as a finding.

Finally, there may be no genetic or brain-parameter *cause* of any mental health disorder. It may be that a set of genes determines the susceptibility of an individual to environmental influences that, in turn, cause changes in brain structure or function or gene expression that are then expressed as the emotional, behavioral, and cognitive problems identified by a psychiatric diagnosis. In the absence of identification of the environmental influence (not in the RDoC matrix), without providing some detection of the disorder (ie, a diagnosis—contrary to the RDoC approach), and with the interactive effects obscured by the RDoC matrix, such a path is unlikely to be found in RDoC studies.

In 2004, there was consternation among basic behavioral researchers because the National Institute of Mental Health (NIMH) lowered the priority of basic cognitive or behavioral research unless it had a strong disease component.[7] Now, there is consternation among clinicians, clinical researchers, patients and their advocates, because the NIMH has lowered the priority of clinical research with a strong disease component unless it has a strong neuroscience component. The logic underlying the 2004 NIMH decision was right: Just as the charge of the National Cancer Institute is to reduce the burden of cancer and that of the National Heart, Lung, and Blood Institute to reduce the burden of heart and lung disease, the charge of the NIMH is to reduce the burden of mental health disorders on patients and on society. That charge has little to do with the issue of categorical vs dimensional diagnoses, or RDoC vs the *DSM*. Both approaches are right; both are flawed; they should be complementary, not antagonistic. That charge has everything to do with (1) producing *consistently replicable* research results that *benefit patients* with mental health problems: *not promises, but results*, and (2) identification and elimination of flawed methods (eg, unreliable measures, unexplained site differences, ignoring complex interactions in analysis) that yield inconsistent, nonreproducible results.

Ultimately, a diagnosis is valid if it predicts future outcome so as to facilitate reducing the burden of mental health problems. All validity criteria, in one way or another, reflect this. If a risk factor (biological, behavioral, environmental, or a combination of these) can be found that identifies those likely to develop a mental health diagnosis, a factor that can be manipulated so as to reduce incidence (categorical) or level (dimensional), that provides evidence of the validity of that diagnosis. If a diagnosis can be used to identify patients whose mental health problems will be eliminated (categorical) or substantially reduced (dimensional) by one treatment rather than another, that provides evidence of the validity of that diagnosis. The more evidence for validity of a diagnosis accumulates, the more valid the diagnosis. Valid diagnosis of mental health disorders, like that of physical health disorders, is a developing process based on accumulating evidence, not a fixed goal, and should use all resources available: dimensional and categorical, *DSM* and RDoC.

ARTICLE INFORMATION

Published Online: November 11, 2015. doi:10.1001/jamapsychiatry.2015.2134.

Conflict of Interest Disclosures: None reported.

REFERENCES

1. Kraemer HC, Noda A, O'Hara R. Categorical versus dimensional approaches to diagnosis: methodological challenges. *J Psychiatr Res*. 2004;38(1):17-25.

2. Kraemer HC. Measurement of reliability for categorical data in medical research. *Stat Methods Med Res*. 1992;1(2):183-199.

3. Kraemer HC, Blasey C. *How Many Subjects? Statistical Power Analysis in Research.* 2nd ed. Los Angeles, CA: Sage Publications; 2015.

4. Kraemer HC. The methodological and statistical evaluation of medical tests: the dexamethasone suppression test in psychiatry. *Psychoneuroendocrinology*. 1987;12(6):411-427.

5. Kraemer HC, Kupfer DJ, Clarke DE, Narrow WE, Regier DA. DSM-5: how reliable is reliable enough? *Am J Psychiatry*. 2012;169(1):13-15.

6. Ioannidis JPA. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124.

7. Holden C. Behavioral science: NIMH takes a new tack, upsetting behavioral researchers. *Science*. 2004;306(5696):602.