Should We Expect “Neural Signatures” for DSM Diagnoses?

Seth J. Gillihan, PhD, and Erik Pares, PhD

Contemporary researchers in psychiatry have sought to develop a nosology based on empirical observation, in line with the principles spelled out by Drs Eli Robins and Samuel B. Guze in 1970. For more than 2 decades, psychiatrists using neuroimaging have aspired to provide one form of “laboratory study” that Robins and Guze said would have to be in place for a psychiatric diagnosis to be valid: researchers have sought “neural signatures” of psychiatric disorders. Our objective was to examine the feasibility of this endeavor. To this end, we examine whether current psychiatric nosology as defined in the Diagnostic and Statistical Manual of Mental Disorders (DSM) lends itself to the identification of neural signatures for psychiatric diagnoses.

Because neuroimaging largely is used only to detect average activation or structural differences between groups of individuals with the same diagnosis and groups of individuals with no diagnosis, it is unlikely that it will be possible to use neuroimaging technologies to determine which psychiatric diagnosis a given individual warrants. In addition, the heterogeneity of psychiatric disorder categories as defined in the DSM reveals that these diagnoses do not reflect neurologically discrete phenomena. Finally, neural correlates of psychopathology generally are not unique to specific diagnoses.

Although it is unrealistic to hope that neuroimaging will be used to make psychiatric diagnoses as they are currently conceived, neuroimaging is already being used to make headway in 2 other arenas of psychiatric investigation that we briefly review.

Submitted: June 16, 2010; accepted August 11, 2010.


Corresponding author: Seth J. Gillihan, PhD, Center for the Treatment and Study of Anxiety, University of Pennsylvania, 3335 Market St, 6th Floor, Philadelphia, PA 19104 (gillihan@mail.med.upenn.edu).

A t least since the publication of the germinal article by Robins and Guze in 1970, psychiatrists have sought to develop a system of disease classification based on science, as opposed to what Robins and Guze called “prior principles.” At least since the late 1980s, psychiatrists using neuroimaging have aspired to provide one form of “laboratory study” that Robins and Guze said would have to be in place for a psychiatric diagnosis to be valid: researchers have sought “neural signatures” of psychiatric disorders. That is, they have sought to discover neural activation patterns or structural differences that ultimately could indicate that a patient had a disorder and which disorder he or she had.

To their credit, the researchers who have shared that aspiration have articulated it carefully. Schwartz et al, for example, wrote that findings from positron emission tomography studies “may have value . . . as a tool for the differential diagnosis [italics added]” of bipolar and unipolar depression. That is, neuroimaging might help to distinguish between disorders with overlapping presentations, such as with mood disturbance. Similarly, an early study by O’Connell et al noted that single photon emission computed tomography “is a promising technique that appears to have potential in differential diagnosis [italics added].”

More recently, in the scientific literature and scientific press, there has been eloquent caution against, if not pessimism about, the near-term prospects of using neuroimaging to make the sorts of diagnoses found in DSM-IV or DSM-5. In spite of such calls to caution, optimism has remained strong in some quarters—although, once again, the optimists articulate their aspiration carefully. In their recent review, Agarwal et al state that “a logical extension of the work performed to date would be to relate imaging findings to specific symptom patterns, such as psychosis, depression, mania, panic attacks, and attention deficits. Ideally, we would like to find specific [neuroimaging] markers for each psychiatric disease.”

Georgopoulos and colleagues conclude, on the basis of their study of synchronous neural interactions in posttraumatic stress disorder, that “the excellent results obtained offer major promise for the usefulness of the [synchronous neural interactions] test for differential diagnosis.” Brotman et al also suggest that neuroimaging will help us to distinguish between different disorders with similar presentations. “Determining the neural circuitry engaged in processing neural faces may assist in the differential diagnosis of disorders with overlapping clinical features.” Similarly, Etkin and Wager concluded in their meta-analysis of brain imaging in anxiety disorders that “identification of a neural signature common to anxiety disorders may be useful in terms of both diagnosis and nosology.”

Enthusiasm for this enterprise grows from the admirable goal of developing a laboratory test for a psychiatric condition that might permit the diagnosis of that condition more precisely than is possible when relying only on a clinician’s behavioral observations and interpretations of a patient’s self-report.

To explain why it is unlikely that it will be possible to use neuroimaging technologies to determine which DSM diagnosis a given individual warrants, we will begin with a brief description of what neuroimaging can achieve today. Specifically, we will observe that today neuroimaging is used largely only to detect average functional or structural differences between groups of individuals with the same diagnosis and groups of individuals with no diagnosis. Then we will identify 2 problems associated with the current use of neuroimaging technologies to investigate psychiatric disorders. Those 2
problems will make it abundantly clear why neuroimaging is unlikely to help identify neural signatures for psychiatric disorders—as those disorders are conceived in *DSM-IV* (and *DSM-5*). Finally, we will say why, even if it is unrealistic to hope that neuroimaging will be used to make psychiatric diagnoses as they are currently conceived, it is already being used to make headway in 2 other arenas of psychiatric investigation. We should note that for a very limited subset of *DSM* categories in which the disorder is defined by structural changes in the brain (eg, dementia due to brain tumor), as well as for neurologic disorders with known neural effects (eg, Alzheimer’s disease), diagnosis based on brain imaging is a more realistic prospect.

**TYPICAL NEUROIMAGING FINDINGS TODAY ARE ABOUT GROUPS, NOT INDIVIDUALS**

Typical neuroimaging studies contrast neural activity (either at rest or in response to some stimulus) or neural structure between individuals with a given psychiatric diagnosis and healthy controls. The controls are chosen to represent a psychiatrically “clean” group of individuals who are free from present (and often past) psychopathology. Some studies even exclude from the healthy control group individuals who report any psychopathology in their immediate family (eg, reference 11) in order to produce maximal separation between groups on the relevant psychiatric variable(s). Resulting differences between groups are attributed to the disease status of the group with the psychiatric diagnosis.

It is one thing to provide information about neural differences between a group of controls and a group of individuals with a *DSM* diagnosis, and it is quite another to provide information that makes it possible to reliably classify individuals as either belonging or not belonging in a particular diagnostic category. To provide the latter sort of information, neuroimaging findings would have to reveal relatively nonoverlapping distributions between individuals with and without a specific disorder; hypothetical results of this kind are shown in Figure 1A.

By way of analogy, consider the fact that men tend to perform better than do women on tests of spatial ability such as mental rotation; for example, a meta-analysis of the results from 286 studies found that men scored higher on tests of spatial ability than did women, and this difference was highly statistically significant. However, the magnitude of this effect was moderate at $d = 0.37$, or about one-third of a standard deviation. In other words, the average difference between men's and women's scores is roughly one-third the size of the average amount that a randomly selected individual's score differs from the mean. Put simply, knowing someone's sex tells us very little about how he or she is going to score on a test of spatial ability.

In the same way, knowing an individual’s psychiatric diagnosis does not allow us at present to make strong predictions about what patterns of brain activity that individual will produce in a neuroimaging scanner, either at rest or during a symptom provocation task. Some individuals with the disorder will produce patterns of brain activity that are indistinguishable from the brain activity of individuals without the disorder—that is, the scan will lack sensitivity with respect to *DSM* diagnoses, yielding false negatives. Also, some individuals without the disorder will produce patterns of brain activity that are indistinguishable from the brain activity of individuals with the disorder—in other words, the scan will lack specificity with respect to *DSM* diagnoses.

---

**Figure 1. Nonoverlapping (A) Versus Overlapping (B) Hypothetical Score Distributions**

*The bars in both A and B represent means (SE) of the hypothetical distributions of scores.*

*All individuals with a specific disorder score lower on some outcomes (eg, prefrontal cortex perfusion) than do all control individuals. In this scenario, the outcome shows strong sensitivity and specificity such that scores reliably classify individuals into the control or disorder category.*

*Most neuroimaging studies produce overall group differences with substantial overlap between the 2 group distributions of scores. In this scenario, it is difficult to classify an individual with certainty on the basis of score, given that a score of 3, for example, could represent an average-scoring person with the disorder or a lower-scoring person without the disorder.*
Should We Expect Neural Signatures for DSM Diagnoses?

HETEROGENEITY WITHIN CURRENT DIAGNOSTIC CATEGORIES

For it to be possible to identify neural signatures of psychiatric disorders, psychiatric diagnoses would have to be directly linked to relatively fixed or discrete neurobiological phenomena; that is, psychiatric disorders would have to be "natural kinds." In other areas of medicine, it often is possible to identify something like natural kinds, a situation in which a diagnostician can say categorically that a person either does or does not have a given disease. For example, a person either is or is not positive for the human immunodeficiency virus (HIV), and a blood test will reveal the person's HIV status with reasonable certainty. Similarly, a single genetic polymorphism determines whether a person will develop Huntington's chorea, and a genetic test can reveal the individual's disease-relevant genotype. In both of these cases, and in many others, medical categories appear to be "carving nature at its joints" to reveal natural kinds.

However, as is widely recognized, the diagnostic categories described in the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR) do not demarcate homogenous natural kinds. For one thing, as psychiatrists ever more keenly recognize, psychiatric disorders are not categorical, but dimensional; once we recall that disorders are expressed along continua, we notice how heterogeneous and indistinctly demarcated their appearances can be. For example, a diagnosis of major depressive disorder can be assigned to an individual who meets the minimal criteria for major depressive disorder for 2 weeks as well as to an individual with a severe 2-year depression. Of course, the symptoms that constitute disorders are also dimensional; moreover, 2 individuals can receive the same diagnosis but have very different symptoms. To glimpse such heterogeneity, it helps to consider a case such as posttraumatic stress disorder (Table 1).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Patient 1</th>
<th>Patient 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trauma</td>
<td>Childhood sexual abuse</td>
<td>Car accident</td>
</tr>
<tr>
<td>Time since trauma</td>
<td>15 years</td>
<td>18 months</td>
</tr>
<tr>
<td>Reexperiencing</td>
<td>Intrusive distressing recollections</td>
<td>Nightmares</td>
</tr>
<tr>
<td>Numbing/avoidance</td>
<td>Avoiding trauma memory</td>
<td>Avoiding trauma reminders</td>
</tr>
<tr>
<td></td>
<td>Psychogenic amnesia</td>
<td>Lack of interest</td>
</tr>
<tr>
<td></td>
<td>Sense of foreshortened future</td>
<td>Feelings of detachment</td>
</tr>
<tr>
<td>Hyperarousal</td>
<td>Difficulty staying asleep</td>
<td>Irritability</td>
</tr>
<tr>
<td></td>
<td>Difficulty concentrating</td>
<td>Hypervigilance</td>
</tr>
</tbody>
</table>

Two individuals who meet the minimum posttraumatic stress disorder diagnostic criteria (1 reexperiencing symptom, 3 numbing/avoidance symptoms, 2 hyperarousal symptoms) have no symptoms in common, including having different criterion A traumatic events. They also differ in length of time since the trauma by an order of magnitude.

A large number of DSM-based diagnoses share this feature of being highly heterogeneous. Depression, for example, requires either depressed mood or anhedonia plus 4 (of 8) additional symptoms. Several of the symptoms can involve opposite presentations—including increased or decreased appetite, insomnia or hypersomnia, and psychomotor agitation or slowness—again leading to very different clinical presentations being given the same diagnosis. The issue of phenomenological heterogeneity is a general concern with DSM diagnosis, but it may figure especially prominently in neuroimaging studies that try to associate a clinical syndrome with activation in specific brain regions. That is, in order to develop a valid and reliable predictor (neuroimaging results) of a criterion (psychiatric diagnosis), the criterion must be relatively consistent and fixed. The pervasive heterogeneity allowed for in the DSM suggests that current diagnostic entities are not the sort of homogeneous or natural kinds that would have to be in place for neuroimaging to identify anything such as their neural signatures.

While each edition of the DSM features substantial heterogeneity within diagnostic categories, changes in the DSM across editions introduce additional inconsistency in diagnosis. Returning to the example of posttraumatic stress disorder, the proposed revisions for DSM-5 (outlined in reference 14) include additional posttraumatic stress disorder criteria such as "persistent distorted blame of self or others about the cause or consequences of the traumatic event(s)" and "reckless or self-destructive behavior." Additionally, a posttraumatic stress disorder diagnosis under the proposed DSM-5 criteria would require 8 symptoms instead of 6 as in DSM-IV-TR. Although these changes may not in fact be made for the next edition of DSM, their proposal underscores the truism that psychiatric categories as conceived in DSM are not natural kinds. In addition to variation in our conception of those categories over time and across individuals, there is of course also variation from place to place—as is made obvious by comparing, for example, DSM-IV and ICD-10, with their slightly different understandings of "the same" diagnosis.
These apparent limitations notwithstanding, perhaps the diagnostic categories as defined in the DSM capture something essential that exists beneath the superficial differences in symptom presentation across individuals diagnosed with the same disorder (and across different systems from different times and places). That is, let us assume that there are many symptoms for each psychiatric disorder just as there are many symptoms of nonpsychiatric diseases (eg, influenza) and that individuals may experience and report different symptoms of the same underlying pathology. We then may ask whether neuroimaging has revealed patterns of neural activity or structure associated with specific psychiatric diagnoses.

**OVERLAP BETWEEN DIAGNOSTIC CATEGORIES**

The sort of neuroimaging research described above, in which a group with a disorder is compared to a “clean” group, is well suited to isolate differences between a group with a psychiatric diagnosis and a group without a psychiatric diagnosis. Above, we emphasized that the data from such studies are not yet robust enough to reliably distinguish between individuals who have and individuals who do not have the disorder of interest. Here we emphasize that the data from such studies cannot yet distinguish between individuals with disorder A and individuals with disorder B. Indeed, there is considerable evidence to suggest that the results from such studies reveal patterns of neural activity and morphometry that are shared across diagnoses.

For example, a meta-analysis of neuroimaging studies of anxiety disorders reported common areas of activation (amygdala, insula) across posttraumatic stress disorder, social phobia, and specific phobia—areas that were also found to be activated in studies of human fear conditioning. The commonalities found across these very different anxiety disorders suggest that neuroimaging has yet to reveal patterns of neural activity that are unique to specific anxiety disorders.

We might expect that anxiety disorders would hang together in this way, given the common denominator of exaggerated fear across these diagnostic categories. However, abnormalities in the amygdala also have been reported consistently in neuroimaging studies of depression. For example, Gotlib and Hamilton reviewed the literature on neuroimaging of depression and concluded that “most consistently, the amygdala and subgenual [anterior cingulate cortex] appear to be overactive in [major depressive disorder], and the [dorsolateral prefrontal cortex] underactive.” A similar pattern of results has been reported in bipolar disorder; Keener and Phillips summarized the relevant neuroimaging results as showing increased activity in emotion processing regions (including the amygdala) and decreased activity in executive regions (eg, dorsolateral prefrontal cortex).

Again, it can be argued that the identification of similar neural areas of hypoactivity or hyperactivity across anxiety and depression makes sense in light of the high comorbidity between these disorders. Perhaps the overactive limbic regions, including the amygdala, represent runaway negative affect, either anxious or depressed, while the hypoactive prefrontal regions are markers of insufficient regulation of limbic areas.

However, differential morphometry and/or function of limbic and prefrontal regions appear to be associated with a wide variety of psychiatric disorders. In schizophrenia—a disorder primarily of thought rather than of mood—a meta-analysis of structural imaging studies showed smaller medial temporal lobe volume compared to controls. Individuals with schizophrenia also have been found to have hyperactive amygdalas during viewing of emotional faces, as well as decreased prefrontal cortical activity during a theory-of-mind task. Psychopathy (which shares features with the DSM diagnosis of antisocial personality disorder) has been associated with similar neural patterns; in their recent review, Wahlund and Kristiansson stated that “a dysfunctional amygdala has been suggested as one of the core neural correlates of psychopathy. . . . Aside from the amygdala, frontal lobe dysfunction has been suggested in psychopaths.” Overactive emotional centers, including the amygdala, and underactive prefrontal regions have been found even among individuals classified as spouse abusers (a DSM-IV-TR V-code), as well as among healthy individuals following sleep deprivation.

Recently, there have been attempts to use multivariate approaches and more complex algorithms to identify psychiatric neural signatures (eg, reference 25). However, it is unclear whether these approaches will succeed in overcoming the limitations spelled out above. In short, the substantial similarities in neural findings across diagnostic categories are grounds to be skeptical about the prospect of using neuroimaging to determine which particular form of DSM-defined psychopathology an individual has. Instead, neuroimaging technologies seem to be identifying neural correlates of general psychopathology, a point that Insel and Wang made cogently in their recent article.

As Andreasen and others have discussed, the evolution of the DSM since 1980 has involved a preference for diagnostic reliability at the expense of validity. The result is that the boundaries of DSM-defined disorders are “boundaries of convenience that permit reliable definition, not boundaries with any inherent biological meaning.” Given that the disorders were defined without respect to biomarkers, there is little reason to expect that studies of neural structure and function should validate the existing categories. Thus, it should not be surprising that neuroimaging approaches that are based on the current diagnostic system have failed to reveal disorder-specific neural correlates.

In light of the limitations noted herein with regard to identifying neural signatures for current diagnostic categories, the National Institute of Mental Health has initiated the Research Domain Criteria (RDoC) Project. This approach is aligned with earlier recommendations that brain imaging should focus on the component processes of psychopathology. The project is based on a reconceptualization of psychiatric...
problems as disorders of brain circuitry and aims to integrate findings from cognitive neuroscience, genetics, and experimental laboratory studies to produce diagnostic categories that align with pathophysiology. As Insel and colleagues state explicitly, "We are a long way from knowing if this approach will succeed." Nevertheless, this project underscores the need for changes in approach if we are going to better understand the brain bases of mental disorders.

DISTINGUISHING AMONG GOALS OF PSYCHIATRIC NEUROIMAGING

To say that brain imaging has not yet been able to reliably diagnose psychiatric disorders does not gainsay the possibility that advances in new combinations of tasks and technologies will someday make it possible to tell depression from anxiety from schizophrenia from psychopathy (and so forth). Nor does it gainsay the possibility that, in combination with other types of information, "including clinical data, genetic information and cognitive testing," neuroimaging will help to make such distinctions. Indeed, neuroimaging is already being used to make headway in at least 2 other areas of psychiatric investigation. The first seeks to establish associations between genes and the building blocks (or components or endophenotypes) of complex phenotypes such as psychiatric disorders. The second seeks to produce novel treatments for DSM disorders.

Imaging Genetics

One of the most promising areas of research in psychiatric neuroimaging has been the investigation of the effects of psychiatric disorder–related genetic variability on neural activity. One of the advantages of this approach over trying to associate brain activity with actual psychiatric diagnoses is that the genetic alleles of interest represent natural kinds. As such, every individual with a given genotype will have the same allele at that locus, assuming perfectly reliable genotyping methods. Thus, the fixed target of investigation that is missing from psychiatric diagnosis is present in imaging genetics. As a corollary, this approach may identify effects on the brain of genetic risk factors that are shared across disorders, which could help to elucidate the mechanisms that contribute to the pervasive comorbidity in mental illness. Imaging genetics also has advantages over traditional behavioral genetics, given that “the penetrance of gene effects will be greater at the level of the brain versus at the level of behavior.” Additional work in this area has shown that the 5-HTTLPR genotype accounts for approximately 10% of the variance in amygdala reactivity ($d = 0.63$). Behavioral effects of 5-HTTLPR tend to be substantially less, which supports the stronger penetrance of genetic effects at the level of the brain versus at the level of behavior; for example, a meta-analysis of the effects of 5-HTTLPR on a measure of neuroticism revealed an effect size of $d = 0.18$. Additional work in this area has shown that the 5-HTTLPR genotype strongly influences the degree to which the cingulate cortex regulates the amygdala, a crucial finding in light of the evidence for underregulation of limbic areas by prefrontal regions across many psychiatric diagnoses. Given the association between 5-HTTLPR genotype and complex psychiatric illnesses like major depressive disorder, particularly under conditions of significant life stress, these results begin to reveal the biological mechanisms that confer risk for depression.

Similar work in the context of schizophrenia has begun to reveal genetic variability related to disease risk that influences neural systems. This work has aimed to elucidate the pathways from genetic variability to behavior outcomes. It is well established that gene-environment interactions account for substantially more variability than do genes alone. The challenge, as with depression and 5-HTTLPR variability, has been to identify the more proximal factors that are affected by schizophrenia-related genes, including effects on neural outcomes. Seminal work in this area has revealed some of the neural endophenotypes associated with schizophrenia-related genes. For example, Huguelet and colleagues reported that KCNH2, a gene that confers vulnerability to schizophrenia, significantly affected neural activity in the hippocampus during a memory task and in the prefrontal cortex during a task of working memory. These results begin to link genetic risk factors with processes that are known to be affected in schizophrenia. Additionally, functional connectivity studies similar to that of Pezawas et al have begun to reveal putative neural circuits that are disrupted in schizophrenia. Esslinger and colleagues reported that a single nucleotide polymorphism in ZNF804A that is associated with risk for psychosis significantly predicted the degree of functional connectivity between dorsolateral prefrontal cortex and the hippocampal formation. This series of studies has begun to establish neural circuits that may underlie some of the classes of major mental disorders such as depression and schizophrenia. Nevertheless, this promising line of research is in its very early stages; thus, a considerable amount of work remains to be done before we can know how much light this approach will shed on the biological mechanisms of psychiatric disorders.

Treatment Development

Psychiatric neuroimaging also has contributed to the development of novel treatments for psychiatric illness. One such treatment is electrical deep brain stimulation for intractable depression. On the basis of findings from functional and structural neuroimaging studies, Mayberg and colleagues developed a neural model of major depressive
disorder that included hyperactivity in the subgenual cingulate cortex. The initial test of this model in a treatment setting used electrodes implanted in this region of the brain in 6 individuals with treatment-refractory depression.44 Four of the 6 patients in this group achieved remission of their depression, which is impressive given that all of these patients had failed to respond to multiple medication trials, talk therapy, and electroconvulsive therapy. Subsequent work45 confirmed the efficacy of deep brain stimulation for treatment-resistant depression; Lozano and colleagues45 found that among 20 treatment-resistant patients who underwent deep brain stimulation for depression, 12 were significantly improved and 7 achieved full recovery (for a review, see reference 46). Thus, while it has not been possible to establish sensitive and specific neural markers of depression, researchers have succeeded in using neural models of depression to develop promising new treatments. Additional work is needed in this area in light of the small number of patients who have been treated to date in randomized controlled trials.

CONCLUSIONS

Existing neuroimaging research of psychiatric disorders has identified patterns of neural structure and activation that differentiate groups of affected individuals from groups of healthy controls. To date, however, this line of research has failed to identify neural signatures that distinguish between individuals with and without a given psychiatric disorder—much less between individuals with different psychiatric disorders. The lack of neuroimaging sensitivity and specificity no doubt reflect in part the diagnostic heterogeneity of DSM-defined disorders. Indeed, the shortcomings outlined here are not unique to neuroimaging, but rather have accompanied a series of new methodologies (eg, studies of neurotransmitter levels; the human genome project), each of which has brought its own disappointments with respect to diagnostic value in psychiatry. In short, it appears that currently available behavioral methods of diagnosing the presence of mental disorders are more accurate and substantially less expensive than are neuroimaging approaches. While the observations presented here suggest that psychiatric neuroimaging is unlikely to become a useful tool for DSM-based diagnosis, review of 2 active research programs provides a more encouraging picture of the potential uses of neuroimaging research in psychiatry. Additional work in these and related areas may continue to reveal the biological underpinnings of psychiatric disorders and related pathology.

Author affiliations: Department of Psychiatry, University of Pennsylvania, Philadelphia (Dr Gillihan), and The Hastings Center, Garrison, New York (Dr Parens).

Potential conflicts of interest: None reported.

Funding/support: The writing of this article was supported by a grant from The Dana Foundation, New York, New York, to Dr Parens (Interpreting Neuroimages: Interdisciplinary Engagement With the Complexities).

Previous presentation: An earlier version of this work was presented at a workshop in Philadelphia, Pennsylvania, February 25–26, 2010, that was funded by The Dana Foundation and organized by The Hastings Center.

Acknowledgment: The authors thank Martha J. Farah, PhD, of the University of Pennsylvania, for providing helpful suggestions regarding content. Dr. Farah has no potential conflicts of interest relative to the subject of this article.

REFERENCES

Should We Expect Neural Signatures for DSM Diagnoses?


