EDITORIAL

The Challenges and Opportunities of Small Effects
The New Normal in Academic Psychiatry

Martin P. Paulus, MD; Wesley K. Thompson, PhD

Explanations and accurate predictions are the fundamental deliverables for a mechanistic or pragmatic approach that academic psychiatric research can provide to stakeholders. Starting with this issue, we are publishing a series of Viewpoints describing the research boundaries and challenges to progress in our field. In this issue, Simon[1] raises the need for better explanatory model using data from electronic health records. This Viewpoint acknowledges an important issue: variables or constructs that are used to help explain the current state of individuals or to generate predictions need to account for a substantial proportion of the variance of the dependent variable or outcome measure to be clinically useful. However, similar to findings from genetics literature, systems neuroscience approaches using brain imaging are beginning to show that variability in structural and functional brain imaging only accounts for a small percentage of the explained variance when considering a variety of clinical phenotypes, especially in large population-representative samples.[2] For example, in a 2016 analysis of UK Biobank data,[3] the functional activation related to a face processing task, which activated the fusiform gyrus and amygdala, accounted for a maximum of 1.8% of the variance of 1100 nonimaging variables. These findings are in line with emerging results from the Adolescent Brain Cognitive Development study[4] focused on the association between screen media behavior and structural MRI characteristics. Importantly, these large-scale studies have used robust and reliable estimators to reduce false-positive discoveries. Thus, similar to genetics literature, it appears that individual processing differences as measured by neuroimaging account for little symptomatic or behavioral variance.

Viewpoint page 349

There is evidence that the association between individual variation on self-assessed symptoms and behavioral performance on neurocognitive tasks is weak.[5,6] Moreover, many behavioral tasks show limited test-retest reliability and little agreement between task conceptualization and actual agreement with emerging latent variables of these tasks. Therefore, it is unlikely that we will find strong relationships between what individuals are reporting about themselves and how they objectively behave. It seems that the individual experience of a person with a mental health condition, which has been proposed to be an important end point for explanatory approaches,[7] is not well approximated by the behavioral probes that are currently available.

These and other findings have profound implications for our theoretical understanding of psychiatric diseases. Specifically, small effect sizes make it unlikely that psychiatric disorders can be explained by unicausal or oligocausal theories. In other words, there is not going to be a unifying glutamatergic or inflammatory disease model of mood disorders. What's more, even if there is a relationship between markers for these disease processes and the state of a psychiatric disorder, as currently conceived, it may not be sufficiently strong to be used by itself to make useful person-level predictions. This is not to say that these processes are not contributing to the etiology or pathophysiology of the disorder but rather that their impact is likely to be small so as to not be individually useful in helping patients and other stakeholders explain their current disease state. As a consequence, there is a low probability of a generic disease process for a group of psychiatric disorders or a final common pathway for a disease.

One possible reason for the lack of a strong relationship between units of analyses, ie, between brain circuits and behavior or behavior and symptoms, is many-to-one and one-to-many mapping. In other words, the brain has many ways of producing the same symptoms, and very similar brain dysfunctions can produce a number of different clinical symptoms. An example of one-to-many mapping is the phenotypic heterogeneity of Huntington disease, which, as an autosomal dominant disorder, has a simple genetic basis but enormous clinical variability via the modulation of multiple biochemical pathways.[8] In comparison, the clinical homogeneity of motor neuron disease is betrayed by a significant genetic variability, leading to similar symptoms.[9] Therefore, it is quite possible that phenotypically similar groups result from different processes and phenotypically heterogeneous individuals actually share broadly similar underlying pathophysiology.

These many-to-one and one-to-many mappings put a profound strain on case-control studies, ie, comparing individuals diagnosed with a particular psychiatric disease with controls that are matched on a limited number of variables. Case-control designs have very limited explanatory depth and are fundamentally uninformative of the disease process because they are correlational, provide little specificity and questionable sensitivity, and have questionable generalizability to populations.[10] Single-case designs together with hierarchical inferential procedures might provide a reasonable alternative.[11] Single-case designs use individuals as their own control, can use controlled interventions to examine causality, and are well suited to uncover individual differences across phenotypically similar participants. However, care must be taken not to subdivide studies so finely that defects of small sample sizes,

including elevated rates of type I and II errors, become problematic even for large epidemiologically informed samples.

Latent variable approaches, such as principal components or factor analyses, can be useful unsupervised statistical methods to uncover relationships between variables within and across units of analyses. However, the underlying assumption is that these latent variables reflect common relationships among all individuals. Instead, it is more likely that relationships differ across individuals and may even differ across states within an individual. Recent approaches to this problem use both latent variable and mixture approaches to differentiate different subgroups of individuals with depression.[12] Others have used deviation from normative regression models to identify heterogeneity in schizophrenia and bipolar disorder.[13] Both sets of approaches support the hypothesis that there are no generic depressive, bipolar, or schizophrenia diseases. At the other extreme, considering that psychiatric diseases emerge from causal factors that vary across units of analyses ranging from molecular to social,[7] one might hypothesize that each individual patient with a mental health condition is an exemplar of a rare disease model. In this case, no generalizable model might be possible, and useful individual-level predictions would be elusive.

Thus, we are facing the classical problem of variance-bias trade-off,[14] which has been examined in great detail in the statistical literature. Specifically, how do we arbitrate between generating a few generic models with useful explanatory or predictive values vs multiple models that may tend to overexplain and overfit individual patient's disease etiology, pathophysiology, and clinical course? This decision cannot be arbitrated solely on statistical grounds but will need to judiciously incorporate expert knowledge about the disease and candidate processes on different units of analyses because the permutational complexity of the variables to be considered is so large that even data sets with thousands of individuals may not provide a sufficient sample size to approach this using exploratory techniques resistant to overfitting.

At this time, we are standing at a precipice: our explanatory disease models are woefully insufficient, and our predictive approaches have not yielded robust individual-level predictions that can be used by clinicians. Yet there is room for hope. Larger data sets will be widely available, multilevel data sets that span assessments from genes to social factors are being released, new statistical tools are being developed, within-subject statistical designs are being rediscovered, and attempts to include expert knowledge into latent variable approaches might help arbitrating the variance-bias trade-off. Fundamentally, academic psychiatry cannot continue to move forward with small *n* case-control studies to provide tangible results to stakeholders.

ARTICLE INFORMATION

Author Affiliations: Laureate Institute for Brain Research, Tulsa, Oklahoma (Paulus); Deputy Editor, *JAMA Psychiatry* (Paulus); Family Medicine and Public Health, University of California, San Diego, La Jolla (Thompson).

Corresponding Author: Martin P. Paulus, MD, Laureate Institute for Brain Research, 6655 S Yale Ave, Tulsa, OK 74136-3326 (mpaulus@laureateinstitute.org).

Published Online: February 27, 2019. doi:10.1001/jamapsychiatry.2018.4540

Conflict of Interest Disclosures: Dr Paulus is supported by grant P20GM121312 from the National Institute of General Medical Services and grant U01DA041089 from the National Institute for Drug Abuse and has received personal fees for scientific writing from UpToDate. No other disclosures were reported.

REFERENCES

1. Simon GE. Big data from health records in mental health care: hardly clairvoyant but already useful [published online February 27, 2019]. *JAMA Psychiatry*. doi:10.1001/jamapsychiatry.2018.4510

2. Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. *Cell*. 2017;169(7):1177-1186. doi:10.1016/j.cell.2017.05.038

3. Miller KL, Alfaro-Almagro F, Bangerter NK, et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat Neurosci*. 2016;19(11):1523-1536. doi:10.1038/nn.4393

4. Paulus MP, Squeglia LM, Bagot K, et al. Screen media activity and brain structure in youth: evidence for diverse structural correlation networks from the ABCD study. *Neuroimage*. 2019;185:140-153. doi:10.1016/j.neuroimage.2018.10.040

5. Eisenberg IW, Bissett PG, Enkavi AZ, et al. Uncovering mental structure through data-driven ontology discovery [published online December 12, 2018]. *PsyArXiv*. doi:10.31234/osf.io/fvqej

6. Thompson WK, Barch DM, Bjork JM, et al. The structure of cognition in 9 and 10 year-old children and associations with problem behaviors: findings from the ABCD study's baseline neurocognitive battery [published online December 13, 2018]. *Dev Cogn Neurosci*. doi:10.1016/j.dcn.2018.12.004

7. Kendler KS. Levels of explanation in psychiatric and substance use disorders: implications for the development of an etiologically based nosology. *Mol Psychiatry*. 2012;17(1):11-21. doi:10.1038/mp.2011.70

8. Ross CA, Aylward EH, Wild EJ, et al. Huntington disease: natural history, biomarkers and prospects for therapeutics. *Nat Rev Neurol*. 2014;10(4):204-216. doi:10.1038/nrneurol.2014.24

9. Dion PA, Daoud H, Rouleau GA. Genetics of motor neuron disorders: new insights into pathogenic mechanisms. *Nat Rev Genet*. 2009;10(11):769-782. doi:10.1038/nrg2680

10. Sedgwick P. Case-control studies: advantages and disadvantages. *BMJ*. 2014;348:f7707. doi:10.1136/bmj.f7707

11. Smith JD. Single-case experimental designs: a systematic review of published research and current standards. *Psychol Methods*. 2012;17(4):510-550. doi:10.1037/a0029312

12. Drysdale AT, Grosenick L, Downar J, et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med*. 2017;23(1):28-38. doi:10.1038/nm.4246

13. Wolfers T, Doan NT, Kaufmann T, et al. Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. *JAMA Psychiatry*. 2018;75(11):1146-1155. doi:10.1001/jamapsychiatry.2018.2467

14. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning With Applications in R*. New York, NY: Springer-Verlag New York; 2013.