Letter

Getting Serious about Variation: Lessons for Clinical Neuroscience (A Commentary on ‘The Myth of Optimality in Clinical Neuroscience’)

Alexander J. Shackman1,2,3,6,* and Andrew S. Fox4,5,6,*

Psychiatric illness poses a significant burden on global public health and the economy. Recent decades have witnessed the development of powerful new tools for quantifying variation in the genome and brain, leading to initial optimism that psychopathology might soon be defined and diagnosed on the basis of objective biological assays. Yet, major breakthroughs have proven elusive. In psychiatric genetics, the first wave of small-scale studies proved difficult to replicate, leading some to question whether gene hunting should be abandoned altogether [1]. The resulting period of crisis and reflection ultimately motivated the widespread adoption of more rigorous analytic approaches and the development of mega-cohorts and consortia with the tens of thousands of subjects needed to reliably detect subtle gene–disease associations [1]. This second-wave research demonstrates that the amount of disease-relevant information captured by the vast majority of individual genetic variants (loci) is vanishingly small <1% [1]. In short, while genomic approaches have proven valuable for discovering new molecular targets and validating existing therapeutics, they are not useful for routine screening or diagnosis.

Whither neuroscience? In their thoughtful review, Holmes and Patrick (H&P) make it clear that that neuroimaging research faces a similar kind of crisis, and that there is no universal, unconditionally optimal pattern of brain structure or function [2]. These insights have enormous implications for a range of stakeholders, from researchers and clinicians to funding agencies and policy makers. Here, we highlight a few of their most important conclusions and expand on their recommendations for increasing the impact of clinical neuroscience research.

Echoing other recent commentators [3–6], H&P remind us that, as neuroimaging samples have grown larger, providing more stable and precise estimates of brain–disease associations, it has become painfully clear that popular candidate neuromarkers of psychiatric disease and risk – amygdala activity and connectivity, striatal reactivity to reward, hippocampal volume, and so on – explain a statistically significant, but small amount of disease-relevant information (e.g., risk, status, treatment response, course). Like common genetic markers, they are far too small to be useful in the clinic; they simply lack the degree of sensitivity and specificity required to benefit individual patients [5].

So, where do we go from here? H&P make several useful suggestions, emphasizing that ‘it is unlikely that we will achieve a breakthrough in our understanding of how the brain’s intricate functions give rise to psychiatric illness by investigating a handful of candidate biomarkers at a time’. They highlight the utility of developing more complex multivariate approaches and the necessity of very large data sets (see Box 1 for detailed recommendations).

H&P also stress the need go beyond data that are simply ‘Big’. Marshaling a range of evidence, they emphasize that particular behavioral, psychological, and neurobiological traits (e.g., amygdala hyper-reactivity) are typically neither good nor bad. Most are associated with a combination of costs and benefits, with the ultimate consequences for health and disease conditional on the larger environment (e.g., presence of danger, exposure to stress or adversity) and neural context (e.g., integrity of countervailing regulatory systems; see also [7,8]). Making sense of this complexity, H&P tell us, requires data sets that are ‘Broad’, encompassing measures of ‘brain structure and function as well as diverse clinical, demographic, behavioral, genetic phenotypes’. Doing so promises to accelerate the development of evidence-based risk calculators and stratified treatment strategies that respect the diversity of human neurobiology [9].

One point not addressed in detail by H&P is the relationship between biomarker development and mechanistic research. Consider the case of genomics. Second-wave genetics research demonstrates that the small-but-reliable associations uncovered by large-scale studies can pinpoint biological pathways with substantial ‘phenotypic’ consequences when manipulated with drugs or other direct perturbations. This raises the possibility that while small-but-reliable associations between neuromarkers and disease are not clinically useful, they could still prove helpful for discovering and prioritizing targets for mechanistic work or for assessing the degree to which discoveries made in animal models translate to diverse human phenotypes and contexts [10]. Such work is likely to be especially valuable when it goes beyond ‘the usual suspects’ and identifies previously unknown or underexplored neuromarkers that are amenable to mechanistic exploration.

With new tools for understanding biological variation comes an opportunity and a responsibility to use them to improve the lives of patients. Developing better
Reliability is typically assessed in small, nonrepresentative samples, precluding de

6. Reliability. In contrast to clinical measures, the test–retest reliability of most neuromarkers is unclear. Reliability is typically assessed in small, nonrepresentative samples, precluding definitive conclusions. We urge investigators to assess and report reliability in larger, more diverse samples.

Biomarkers is important for a variety of reasons, from screening and diagnosis to treatment stratification and monitoring [3]. Developing neuromarkers that more seriously reckon with the complex interactions of human brains, contexts, and outcomes would accelerate the development of new therapies and the repurposing of existing ones [11]. H&P remind us that this is a major challenge, one that will require substantial time and resources, new kinds of multidisciplinary collaborations and training models, and a sober assessment of what particular kinds of neuromarkers really can and cannot do for patients and clinicians.

Acknowledgments
We acknowledge financial support from the California National Primate Research Center, National Institutes of Health (DA040717 and MH107444); University of California, Davis; and University of Maryland, College Park. We thank B. Roberts and R. Podlrack for insightful discussions and L. Friedman and K. DeYoung for assistance.

1Department of Psychology, University of Maryland, College Park, MD 20742, USA
2Neuroscience and Cognitive Science Program, University of Maryland, College Park, MD 20742, USA
3Maryland Neuroimaging Center, University of Maryland, College Park, MD 20742 USA
4Department of Psychology, University of California, Davis, CA 95616, USA
5California National Primate Research Center, University of California, Davis, CA 95616, USA
6Both authors contributed equally to this work; the order of authors was determined by a coin flip

*Correspondence: shackman@umd.edu (A.J. Shackman) and dfox@ucdavis.edu (A.S. Fox).
https://doi.org/10.1016/j.tics.2018.02.009

References
Supplementary References

Alexander J. Shackman*1,2,3 and Andrew S. Fox*4,5

1Department of Psychology, 2Neuroscience and Cognitive Science Program, and 3Maryland Neuroimaging Center, University of Maryland, College Park, MD 20742 USA. 4Department of Psychology and 5California National Primate Research Center, University of California, Davis, CA 95616 USA

* A.J.S. and A.S.F contributed equally to this work; the order of authors was determined by a coin flip

Please Address Manuscript Correspondence To:
Dr. Alexander J. Shackman (shackman@umd.edu)
Department of Psychology, University of Maryland, College Park MD 20742 USA
Dr. Andrew S. Fox (fox.drew@gmail.com)
Department of Psychology, University of California, Davis, CA 95616 USA
ANOTATED SUPPLEMENTARY BIBLIOGRAPHY

- Psychiatric illness is a leading burden on global public health and the economy (DiLuca & Olesen, 2014; Global Burden of Disease Collaborators, 2016).

- Initial optimism that the miseries of psychopathology might soon be defined and diagnosed on the basis of objective biological assays (Casey et al., 2013; Charney et al., 2002; Hyman, 2007).

- The second-wave of psychiatric genetics research (Bogdan, Baranger, & Agrawal, in press; Chabris, Lee, Cesarini, Benjamin, & Laibson, 2015; Plomin & von Stumm, 2018; Sullivan et al., 2018; Timpson, Greenwood, Soranzo, Lawson, & Richards, 2018).

- GWAS remains a critical tool for validating existing therapeutics and for discovering and prioritizing disease-relevant molecular pathways. This reflects the fact that small-but-reliable genetic associations can be used to uncover biological pathways that have large 'phenotypic' consequences when manipulated with drugs or other mechanistic approaches (Gandal, Leppa, Won, Parikshak, & Geschwind, 2016; Pankevich, Altevogt, Dunlop, Gage, & Hyman, 2014; Sullivan et al., 2018; Timpson et al., 2018).

- Popular candidate neuromarkers of psychiatric disease and risk explain a significant, but small amount of disease-relevant information (e.g., risk, status, treatment response, course)—far too small to be useful in the clinic (Prata, Mechelli, & Kapur, 2014; Savitz, Rauch, & Drevets, 2013). For example:
  - *Amygdala function* (Fox et al., 2015; Kaczkurkin et al., 2016; Sjouwerman, Scharfenort, & Lonsdorf, 2017; Stout, Shackman, Pedersen, Miskovich, & Larson, 2017; Swartz, Knodt, Radtke, & Hariri, 2015)
  - *Striatal reactivity to reward* (Stringaris et al., 2015)
  - *Hippocampal volume* (Logue et al., 2018)

- Paper-and-pencil measures of personality and actuarial approaches that leverage readily available demographic and clinical information often outperform candidate neuromarkers (e.g., Heinrich et al., 2016; Whelan et al., 2014), with some interesting exceptions (Jollans & Whelan, 2016; Poldrack et al., 2018).

- Best practices’ are likely to enhance the scientific and clinical value of psychiatric neuroimaging research (Gabrieli, Ghosh, & Whitfield-Gabrieli, 2015; Holmes & Yeo, 2015; Jollans & Whelan, 2016; Kapur, Phillips, & Insel, 2012; Poldrack et al., 2018; Rosenberg, Casey, & Holmes, in press; Savitz et al., 2013; Woo, Chang, Lindquist, & Wager, 2017; Yarkoni & Westfall, 2017).
• **Dimensional Phenotypes.** DSM-5 diagnoses present significant barriers to the development of clinically useful neuromarkers, including substantial symptom heterogeneity, frequent comorbidity, and low inter-rater reliability (i.e., uncertain ‘ground truth’) (Chmielewski, Clark, Bagby, & Watson, 2015; Fried, 2017; Galatzer-Levy & Bryant, 2013; Hasin et al., 2015; Kotov et al., 2017; Krueger et al., *in press*; Regier et al., 2013; Zald & Lahey, 2017).

• **Big Data.** Mega-cohorts provide increased power and more precise and realistic estimates of brain-disease associations. They afford important opportunities for assessing generalizability (i.e., they guard against model ‘over-fitting’); and for investigating and controlling confounding demographic and imaging variables (e.g., motion) (LeWinn, Sheridan, Keyes, Hamilton, & McLaughlin, 2017; Miller et al., 2016; Schnack & Kahn, 2016; Smith & Nichols, 2018; Varoquaux, *in press*).

• **Reliability.** In contrast to clinical measures, the reliability (e.g., test-retest stability) of most neuromarkers remains unclear (Lilienfeld & Treadway, 2016; Tomarken, 1995). Often reliability is assessed using small (\(n<30\)), non-representative samples (e.g., Cannon, Cao, Mathalon, Gee, & consortium, 2018), precluding definitive conclusions about clinical utility.

• **Rigor and Reproducibility.** Understanding the neurobiological underpinnings of mental illness demands rigorous, cumulative research. Yet, the actual reproducibility of biomedical research is alarmingly low (Fox, Lapate, Davidson, & Shackman, 2018; Munafò et al., 2017). Pre-registering study hypotheses and sharing data are easy ways to enhance reproducibility, increase power, and accelerate progress.
REFERENCES


*Biological psychiatry, 80*, 775-785.


Varoquaux, G. (*in press*). Cross-validation failure: Small sample sizes lead to large error bars. *NeuroImage*.


