Depressive disorders are a leading cause of global disability, afflicting approximately 280 million individuals worldwide each year (1). In the United States, more than one in five individuals will experience a lifetime depressive disorder, diagnoses and service utilization are surging, and direct health care costs exceed $68 billion annually (2–6). These unfortunate observations underscore the need to develop a better understanding of the neural systems that underlie depression.

Major depressive disorder (MDD) is a heterogeneous phenotype that typically emerges in late adolescence or early adulthood (7, 8). Clinical presentation can be transient or recurrent, with periods of waxing and waning impairment and distress. Comorbidity with anxiety disorders, substance misuse, and other illnesses is common (9, 10). Given this phenotypic, developmental, and—in all likelihood—etiological complexity, it is unsurprising that neuroimaging studies of MDD have implicated a diverse array of brain regions, including the amygdala, ventral striatum, thalamus, and cingulate (11, 12).

Among the regions linked to depression, the amygdala has received some of the most intense empirical scrutiny. This body of research has led many to conclude that amygdala hyperreactivity is a robust marker of MDD (11, 12). Ancillary analyses have suggested that these effects, which were largely driven by studies of negative faces and scenes (20, 21), were only evident at a liberal threshold, and only when pooling studies of MDD and anxiety (21). Li and Wang reported significant hyperreactivity in the left amygdala to emotional faces and scenes in individuals with current depressive disorders, but this was only found when aggregating positive and negative stimuli (11). In the most comprehensive analysis, McTeague and colleagues observed significant amygdala reactivity to negative faces and aversive challenges, consistent with a causal role (19). Third, three recent coordinate-based meta-analyses (CBMAs)—all adhering to methodological best practices and collectively encompassing dozens of studies and thousands of participants—provide convergent evidence of left amygdala hyperreactivity in individuals with MDD (11, 20, 21).

Despite this progress, it is clear that most of the work necessary to understand the nature and degree of the amygdala’s contribution to depression remains undone. Consider the CBMA evidence. To ensure an adequate number of studies, all of the meta-analytic teams were forced to engage in substantial “lumping,” and their results reflect a mixture of adults and youth, medicated and unmedicated cases, and a panoply of emotional and cognitive tasks. Janiri and colleagues found evidence of left amygdala hyperreactivity, but this was only evident at a liberal threshold, and only when pooling studies of MDD and anxiety (21). Li and Wang reported significant hyperreactivity in the left amygdala to emotional faces and scenes in individuals with current depressive disorders, but this was only found when aggregating positive and negative stimuli (11). In the most comprehensive analysis, McTeague and colleagues observed significant hyperreactivity in the left amygdala to emotional stimuli in individuals with interview-verified MDD or anxiety diagnoses (20). Ancillary analyses suggested that these effects were largely driven by studies of negative faces and scenes (20, 21). While these results clearly show that left amygdala reactivity to negative stimuli is elevated, on average at least, among individuals with MDD, it remains unclear whether this association reflects differences in the perception of negative faces, the generation of negative affect to aversive stimuli (e.g., unpleasant scenes, threat of shock), or some combination of the two (22).
But the most significant and often overlooked limitation of the CBMA evidence is the raw input, the gist for the meta-analytic mill. While all of the CBMAs have impressively large pooled samples, the size of the constituent imaging studies is worrysome small. In the most recent CBMA (N = 2,383 [11]), the median sample size was just 39 participants—19 cases and 20 controls—far too small to provide stable conclusions, even under the most generous (and frankly unrealistic) assumptions (23). For a benchmark “large” effect (d = 0.80 or R² = 14%) and a liberal whole-brain corrected threshold (α_{one-tailed} = 0.01, Z_{critical} = 2.33), the power to detect case/control differences in activation is just above chance (53.1%). In the absence of publication, confirmation, or other biases favoring particular outcomes, CBMAs derived from underpowered studies are vulnerable to false negatives (24). But in the presence of such biases, underpowered studies will tend to capitalize on chance sampling variation and questionable research practices in ways that optimistically bias meta-analytic results—an outcome clearly demonstrated in the candidate gene literature (25–28).

From this perspective, the new report from Tamm and colleagues (29) in this issue is a welcome addition to the literature. Leveraging data acquired from >20,000 older UK Biobank participants (median age = 64 years), the authors estimated associations between three depression phenotypes and amygdala reactivity to negative faces with an unprecedented degree of statistical precision (27). Depression phenotypes included an ad hoc four-item self-report scale of acute (past 2 weeks) depressive symptoms, self-reported lifetime depression diagnosis, and probable lifetime major depression based on a diagnostic questionnaire. None of the assessments employed trained interviewers, and only the last used formal diagnostic criteria (for a detailed critique of depression phenotyping in the UK Biobank, see Cai et al. [30]). Individual differences in amygdala reactivity were quantified in an unbiased manner using a bilateral amygdala region-of-interest. Notably, both the data and code are publicly available, facilitating future use by other investigators.

Tamm and colleagues’ analyses revealed null associations between amygdala reactivity to negative faces and self-reported symptoms and lifetime diagnoses. Relations between amygdala reactivity and the much stricter diagnostic questionnaire were numerically stronger and statistically significant (p < 0.01). Nonetheless, the magnitude of this association was vanishingly small (d = 0.03, R² = 0.03%) and nonsignificant in models that included demographic covariates (p = 0.13). The authors conclude by noting that “an association between depression and amygdala responses to negative faces is not likely to be as large as previously suggested...[and] that amygdala responses to negative facial expressions should not be considered an important feature/biomarker of depressive symptoms, at least not in the general population.”

Tamm and colleagues’ observations add to a growing body of psychiatric imaging research demonstrating that amygdala hyperreactivity and other popular candidate biomarkers explain statistically significant but quantitatively negligible amounts of disease-relevant information—risk, status, treatment response, course, and so on—in large samples. This pessimistic conclusion is hardly specific to the amygdala. A recent meta-analysis demonstrated that dampened ventral striatum reactivity to reward is significantly (p = 0.007) and consistently (nine studies; median N = 91) associated with the future emergence of depression, consistent with a causal role (12). Yet the strength of this small-but-reliable association is far too weak (R² = 1%) to be useful for screening, clinical, or treatment development purposes.

From a conceptual perspective, Tamm and colleagues’ diagnostic-questionnaire results are reasonably well aligned with work in younger populations (reviewed above), suggesting that 1) higher levels of amygdala reactivity to negative faces probabilistically increase the likelihood of anhedonia and anxiety symptoms among individuals exposed to NLEs; 2) amygdala reactivity is more strongly associated with state-independent risk than acute symptoms; and 3) on average, amygdala reactivity to negative faces and scenes is elevated in groups of individuals with verified acute MDD. Individuals with MDD show a wide variety of clinical presentations, and this body of evidence is consistent with the possibility that amygdala hyperreactivity is only etiologically relevant for a subset of patients and symptoms. Determining whether this is true or simply wishful thinking is a key challenge for the future. From a mechanistic perspective, the small-but-reliable “hits” uncovered by Big Data studies—including Tamm and colleagues’ diagnostic-questionnaire results—do not preclude much larger effects with targeted biological interventions (12, 31). Indeed, work in animals demonstrates that focal perturbations of specific amygdala cell types can have dramatic, complex, and even opposing consequences for reward- (“wanting”) and anxiety-related behaviors (32, 33).

In sum, work conducted over the past decade has yielded steady advances in our understanding of depression. Yet the underlying neurobiological mechanisms remain elusive, actionable biomarkers remain out of reach, existing treatments are far from curative, and relapse and recurrence are common (34–36). Tamm and colleagues’ report serves as a sober reminder that simple box-and-arrow neurobiological explanations—which equate amygdala hyperreactivity with depression independent of clinical presentation, severity, disease stage, developmental period, adversity exposure, imaging technique, and fMRI paradigm—are no longer tenable.

Addressing these challenges will require an increased investment in psychiatric research, one commensurate with the staggering burden that depression and anxiety impose on global public health. UK Biobank and other Big Data studies (e.g., ABCD, All of Us) clearly have an important role to play in overcoming these challenges, but to be maximally useful the next generation of biobank and large-scale psychiatric studies will need to overcome the significant limitations of existing ones. This will require the recruitment of demographically
representative samples and adequate representation of severe psychopathology, rigorous psychiatric phenotyping, and reliable imaging approaches—three notable limitations of the Tamm study (10, 30, 37–40). To really move the needle on our understanding—and ultimately on clinical practice—we will need to move beyond negative-face paradigms and other kinds of tried-and-true experimental challenges (31). Even if the amygdala is mechanistically involved in the development of maladaptive anhedonia or anxiety—as suggested by prior work in humans and animals—then conventional negative-face paradigms are fundamentally the wrong experimental assay. Some of these challenges can be overcome by appropriately focused “Medium Data” projects (N=200–2,000; e.g., Tulsa 1000) or by pooling data via existing consortia (e.g., ENIGMA). It is also worth reminding ourselves that the amygdala is a heterogeneous collection of nuclei linked by a network of microcircuits (41). Fully understanding the amygdala’s relevance to depression and other illnesses requires that future studies more fully embrace this neuroanatomical complexity. From the perspective of prediction, it is clear that cross-validated multivariate machine-learning approaches and related techniques—which quantitatively synthesize multiple sources of imaging and nonimaging information at the population or patient levels—are more likely to yield clinically useful tools than studies focused on isolated “hot spots” of brain function or structure (31). A greater emphasis on reliable dimensional phenotypes (e.g., anhedonia) and the development of integrative cross-species models promises to further accelerate efforts to alleviate the suffering caused by depression (12, 31).

AUTHOR AND ARTICLE INFORMATION
Department of Psychology (Grogans, Shackman), Neuroscience and Cognitive Science Program (Shackman), and Maryland Neuroimaging Center (Shackman), University of Maryland, College Park; Department of Psychology and California National Primate Research Center, University of California, Davis (Fox).

Send correspondence to Dr. Shackman (shackman@umd.edu).

Supported in part by the California National Primate Center; National Institutes of Health (MH121409, MH121735, MH128336, MH129851, OD011017); ASAP Foundation; University of California, Davis; and University of Maryland.

The authors acknowledge assistance from K. DeYoung, L. Friedman, and J. Smith and critical feedback from A. Etkin and N. Kalin.

Dr. Shackman reports serving on a scientific advisory board for Hoffmann-La Roche AG, with a focus on the development of novel anxiolytic compounds. The other authors report no conflict of interests with commercial interests.

Accepted May 6, 2022.


REFERENCES


3. SAMHSA: Key substance use and mental health indicators in the United States: results from the 2018 National Survey on Drug Use and Health. Rockville, MD, Center for Behavioral Health Statistics and Quality, 2019


