The critical importance of basic animal research for neuropsychiatric disorders

Tracy L. Bale[1], Ted Abel[2], Huda Akil[3], William A. Carlezon[4], Bita Moghaddam[5], Eric J. Nestler[6],

Kerry J. Ressler[7] and Scott M. Thompson[8]

1 Center for Epigenetic Research in Child Health and Brain Development, Departments of
Pharmacology and Psychiatry, University of Maryland School of Medicine, Baltimore, MD,
tbale@som.umaryland.edu

2 Department of Pharmacology and Iowa Neuroscience Institute, Carver College of Medicine,
University of Iowa, Iowa City, IA, ted-abel@uiowa.edu

3 Molecular and Behavioral Neuroscience Institute, University of Michigan, Ann Arbor, MI,
akil@med.umich.edu

4 McLean Hospital, Department of Psychiatry, Harvard School of Medicine, Belmont, MA,
bcarlezon@mclean.harvard.edu

5 Department of Behavioral Neuroscience, Oregon Health and Science University, Portland,
OR, bita@ohsu.edu

6 Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY,
eric.nestler@mssm.edu

7 McLean Hospital, Department of Psychiatry, Harvard School of Medicine, Belmont, MA,
kressler@mclean.harvard.edu

8 Department of Physiology, University of Maryland School of Medicine, Baltimore, MD,
sthompson@som.umaryland.edu

Contact information:

Tracy L. Bale, Ph.D.
Professor and Director, Center for Epigenetic Research in Child Health and Brain Development
Departments of Pharmacology and Psychiatry
University of Maryland School of Medicine, Baltimore, MD 21201
tbale@som.umaryland.edu
ph 410-706-5816

The challenges and critical importance of keeping our thinking about neuropsychiatric disorders mechanisms and classifications up-to-date have prompted a dynamic discourse as to the value, appropriateness, and reliability of the animal models we use and the outcomes we measure. At a time when most major pharmaceutical companies are disbanding their research and development in neuroscience and mental health programs, academia is the remaining bastion for the identification and validation of novel drug targets and the development of novel therapeutic approaches. The 2018 ACNP meeting provided the ideal audience for a discussion centered on this topic, offering a professional setting in which diverse viewpoints could be presented, and allowing voices from all arenas of basic and translational neuroscience to be heard. At its core, there is a growing and healthy discussion about the degree of confidence we have that our preclinical animal models and their behavioral endpoints can predict the success of new drugs in clinical trials. While it is clear that we cannot completely model complex neuropsychiatric and neurodevelopmental disorders in animals, there is certainly agreement across domain criteria and endophenotypes of disease that have proven informative, particularly when multiple levels of outcome measures, including genomics, epigenomics, cellular properties, circuits, network dynamics, and behaviors are included. This issue is now of critical importance with the arrival of big 'omics data sets that provide novel and informative insights about genetic risk and molecular pathophysiology. How can we best examine these insights into disease risk and resilience using preclinical measures? Similarly, while genome-wide association studies (GWAS) and genome sequencing studies provide valuable clues into the genes and pathways that are associated with disease risk, how can we determine how these loci interact with each other as well as with the environment to assess their importance across developmental and life stages? Most importantly, we seek to identify ways to develop novel therapeutic approaches to improve the lives of those with neuropsychiatric disorders including those of neurodevelopmental origins, a challenge that remains even after decades of intensive research.

How do we go from results focused on genes, cells, and behaviors in rodents and other animal models to much more complex and heterogeneous syndromes in humans? How do we control for the variables across labs and between species and mouse strains that make interpretation or translation difficult? At the cellular level, utilization of human inducible pluripotent stem cells (iPSC), induced neuron-like cells, and organoids for phenotypic characterization including migration assays and electrophysiology provide an important compliment to animal models (Quadrato et al. 2016). For example, genetic networks may be well preserved, facilitating the study of the epistasis of multiple risk factor genes, each conferring low effect on disease etiology. How closely these cells recreate aspects of human disease conditions is a critical question remaining to be answered, however. Furthermore, it is important to acknowledge that such cells are necessarily studied under highly artificial conditions, often lacking well-established critical disease risk factors, and where the relevant environmental and cellular interactions are not known. Certainly, the outcomes that can be measured with these approaches will remain far from the symptoms of human neuropsychiatric diseases. It is inconceivable that a pharmaceutical company would move forward with a clinical trial for a molecule that was only validated in a cell model without some efficacious signal obtained from animal models.

In our ACNP panel presentations and the resulting discussion, the following points were made and emphasized as critical to our abilities to move forward with reliable studies important to the field. It was clear from audience discussion that there is unanimous appreciation for the importance of animal studies relevant to mental health disorders as a necessary and invaluable aspect of biomedical research in drug discovery, development, and validation. The panel focused on several key areas including validation of animal models important to studying depression by Eric Nestler, an in-depth discussion as to what is being measured in animal

behavioral studies and their relevance to disease by Josh Gordon, an appreciation for the value in understanding basic cellular processes and 'normal' brain function in order to apply a framework to disease risk from Huda Akil, how the development of antidepressant drugs and novel drug targets for depression provide evidence as to the importance of animal models by Scott Thompson, and the importance of model systems and conserved circuits to broadly understand disease from Kerry Ressler. Additional insight and discussion on the importance of basic research in animal behaviors was discussed by Bita Moghaddam and Ted Abel. Bill Carlezon also discussed the importance of using endpoints that can be objectively defined and studied in both humans and laboratory animals, with an emphasis on those that are becoming increasingly available in humans via smart phones and wearable devices, as a potentially transformative opportunity to better align psychiatry and neuroscience.

While the overarching focus of the panel discussion remains pertinent to the broad field of neuropsychiatric disorders, the topics discussed below largely utilize affective disorders as examples of the strengths and challenges and future directions required for the effective and informative use of animal models. Further, while the field encompasses a valuable breadth of species in such studies, examples provided below rely on rodents as the predominant preclinical model organism.

Challenges in modeling the complex human brain
The panel presentations and discussion started by acknowledging the complexities of neuropsychiatric diseases and the challenges of modeling and testing these complexities in rodents (Nestler and Hyman, 2010). Using depression as an example, the panel highlighted that it is impossible to model the entire syndrome of depression in an animal model. First, depression is highly heterogeneous, which means that no single model could possibly capture the distinct etiologies and pathophysiological mechanisms involved. Second, depression is only

~35% heritable, with likely many hundreds of genes comprising that risk, each with only a miniscule of independent impact that differs among depressed individuals; it is therefore impossible to capture even that 35% heritability in an animal model.

By contrast, the best-established risk factor for depression is a history of life adversity or experience of stress or trauma. Accordingly, there is validity in focusing on chronic stress procedures (involving stress exposure at points across the lifespan, including stress experienced before, during and after pregnancy and birth) as models for aspects of human depression risk. It is essential when using these models for the investigator to distinguish between maladaptive vs. adaptive responses to such stress, namely, to differentiate mechanisms of stress *susceptibility* from mechanisms of stress *resilience*.

Focus on Stress Biology: Fundamental Research, Animal Models and Relevance to Human Affective Disorders

It may be useful to step back and emphasize why studying stress neurobiology in animal models is relevant to human mood disorders. First, psychosocial stress represents the major antecedent of depression and other affective disorders. *Therefore, stress is a definable trigger of the presentation of these diseases in vulnerable individuals*. However, the relevance of stress neurobiology is more fundamental. At their core, affective disorders are disorders of stress coping. These disorders arise when the strategies that the organism relies upon to meet environmental challenges are insufficient, dysregulated, or otherwise maladaptive, and the affective correlates of this dysregulation can include anxiety, persistently negative mood, or in the case of traumatic experiences, post-traumatic stress disorder (Selye, 1955). Therefore, understanding the fundamental biology of stress is essential to uncovering the basis of vulnerability (or resilience) to these diseases, and for designing new classes of drugs for their

treatment. This is akin to how our understanding of the biology of cellular signaling, cell growth, and cell death was essential to lay the groundwork for novel treatments developed for cancer.

Thus, in the case of affective disorders, the dysregulated system represents a central function (stress responsiveness and coping) that has a clear parallel in rodents and other mammals. Indeed, the general feature of stress neurobiology, such as the existence of a hypothalamic – pituitary – adrenal axis, the role of the hippocampus, amygdala, reward circuitry, and frontal cortex, the identification of key molecules, and the broad organizational features of stress circuitry are solidly established in rodents and remarkably translatable to humans (Nemeroff and Vale, 2005). We, of course, currently lack a full understanding of when and how the stress system becomes dysfunctional, details that basic research utilizing state of the art technologies can provide. This is the area where bi-directional translation is critical; where human findings can drive questions in animal models, and animal models can provide some answers that can be tested in humans. Clearly, the readouts in animals - the choice of appropriate and meaningful behavioral tests (preferably a sophisticated range that are relevant to human behaviors), as well as the neural, genomic and other correlates of the manipulations - are critical. But the idea is not to pretend that we are fully modeling the human illness or even specific dimensions of the illness. *Rather, the goal is to achieve a better understanding of an essential biological function that is key to the illness, and to ensure that our level of understanding is actionable for translation.* Thus, we view the goal of stress research in two ways: 1) to achieve a better understanding of the biology of stress and its dysregulation that will shed light on the pathophysiology of affective illnesses risk and resilience, and, 2) to provide sophisticated behavioral models for discovering new molecular drug targets and for testing novel treatments. Of course, to make progress, it is critical to identify strategies for validating the usefulness of these animal models and their relevance to human disease.

Chronic stress rodent models arguably have the best validity of any currently well characterized model for depression-related disorders. For instance, several chronic stress models have recently recapitulated many of the abnormal gene expression patterns found in the depressed human brain examined postmortem, and changes in expression have then been correlated with stress susceptibility (Akil et al., 2018). This finding illustrates methods of validating an animal model beyond simply examining homologous behaviors detected in the models and in depressed humans, which can be difficult. Demonstrating shared circuit abnormalities in animal models and depressed humans offers another possible avenue to establish face validity. Thus, while the field should be challenged to establish better behavioral assays – with greater translational potential and interpretability – as a consequence of chronic stress in rodents, it is factually wrong to disparage the validity and importance of chronic stress models. It is clearly irresponsible to suggest that animal research is not useful or that human-derived cell models could be substituted. This is a *reductio ad absurdum* argument, as noted above, as no pharmaceutical company would or should embark on a clinical trial without some preclinical evidence and efficacy signal in an animal study. It is nevertheless important to keep our focus on what animal research is useful for.

Using model systems to understand human neuroscience

The problems facing those wishing to understand the neurobiology of psychiatric disorders are vast, with complexity at all levels of analysis, from molecular to cellular to systems underlying complex individual and social behaviors. While it is clear that animal models cannot fully represent the complexity of human neuropsychiatric disorders, failing to utilize the opportunities that do exist to elucidate mechanisms resembles 'throwing the baby out with the bathwater.' In several areas of neuroscience there is tremendous evidence for conservation across species, from mouse to human, for basic behaviors and underlying circuit-, cellular-, and molecular-based mechanisms. Examples include fear and startle circuits, and their regulation of appetitive

and aversive behaviors that are relevant to anxiety disorders, post-traumatic stress disorder, and addiction. The fear and startle circuits originally identified by basic research involve brain stem and midbrain regions including amygdala that are highly conserved in form and function across mammals; advances that are being used to improve translation to therapeutics (LeDoux and Pine, 2016). Decades of ensuing basic research work have established very clear circuitry for a variety of behavioral domains that has largely held up across human imaging and physiology combined with up-to-date rodent circuitry tools such as optogenetics, chemogenetics, calcium imaging, and electrical-based physiology tools. While much more needs to be established, powerful approaches such as single cell RNA sequencing across regions and species, and large scale genetic tools combined with transcriptomics and digital phenotyping across species are enabling truly novel and powerful translational approaches – in which we are not modeling disorders per se, but instead the component parts, from molecules to cells to circuits to aspects of behavioral syntax. Rather than dismissing our use of disease model systems, we must take advantage of the power of today's tools, across species, to achieve a much greater understanding of the mammalian brain and how its function is disrupted in disease.

Antidepressants as an example

So, why then have so few novel, neuroscience-based, rationally identified antidepressants reached the clinic - do our animal studies fail to identify promising new targets? Actually, no. Many argue that preclinical research has identified an excess of targets. The disparity then between what preclinical 'models' have suggested as potential targets and the paucity of efficacious new drugs developed in the last 50 years strongly supports a need for more stringent criteria and more robust evidence to validate drug targets, prior to advancing them toward translation to humans. Moreover, it has become increasingly difficult over the past two decades to test molecules with novel mechanisms of action in humans, where challenges with clinical

trials including patient recruitment, trial design, the placebo response, and regulatory obstacles have dramatically hindered the field's ability to test novel mechanisms flowing through the preclinical research pipeline.

One clear hurdle for this field is that human neuropsychiatric disorders are diagnosed based on clinical criteria that do not necessarily relate to specific brain mechanisms or behaviors. Further, these clinical criteria are often not quantitative or objective. From this perspective, outcomes would benefit from more laboratory-based studies of human subjects to identify quantitative phenotypes (perhaps those that could be measured with 'digital' phenotyping) that ideally are responsive to treatment. Preclinical models would then focus on these quantitative phenotypes. It is also important to note that the field has focused largely on studies of the disease process once it has already developed. What processes give rise to neuropsychiatric disorders and what are their antecedents? We've made dramatic inroads into heart disease by focusing on elevations in blood pressure and cholesterol that occur years before a heart attack. What are the analogous processes for neuropsychiatric disease? For depression, we have a hypothesis - that it is stress or trauma experienced during specific windows of brain development or maturation, and an atypical response to stress, that sets the stage for disease. Hypotheses such as this can be tested in a clinical laboratory setting to further inform the preclinical work.

Precision is important
Perhaps the most efficient place to increase rigor begins with cleaning up our own language. As noted above, depression and other neuropsychiatric disorders are uniquely human diseases and, speaking precisely, nonhuman animals cannot therefore be 'depressed' or 'schizophrenic' or 'autistic'. We risk damaging our credibility with funding agencies and the public when we speak of 'depressed mice.' The Director of the National Institute of Mental Health, Josh Gordon, articulated, for example, the important difference between a model *of* depression versus a

model *useful for* the study of depression, the latter being the more accurate and objectifiable approach. Similarly, it follows that a compound tested only in nonhuman animals cannot truly be an 'antidepressant' or exert an 'antidepressant action.' The best we can say about such compounds, and only after they have been tested and proven in depressed patients, is that they are antidepressant-like or produce antidepressant-relevant responses.

With this in mind, what preclinical approaches then offer the best utility at this time for identifying promising and druggable targets for neuropsychiatric disorders such as depression? One clear point that provoked a very strong response from the ACNP audience was related to the use of behavioral endpoints in rodents that are not obviously relevant to the diseases they are intended to study. Many human antidepressants produce behavioral changes in nonhuman animals that can be detected with the tail suspension test (TST) or forced swim test (FST), suggesting that they may have some predictive validity. These tests are probably not sufficiently stringent for identification of novel antidepressant candidates, however, because responses can be elicited in control, stress-naïve animals, whereas the mood improvement responses of non-depressed human subjects to known antidepressants have not been documented. In addition, responses in these tests in rodents are elicited by acute drug administration of monoamine-based antidepressants (including selective serotonin reuptake inhibitors or SSRIs), whereas the human antidepressant response to these medications occurs only with chronic (several weeks to months) administration, suggesting that the mechanism(s) by which they alter behaviors in the TST or FST are distinct from their antidepressant action in humans (Conn and Roth, 2008). Furthermore, these behaviors have no obvious relation to any of the symptoms of human depression, i.e. they lack face validity. Nevertheless, the FST and TST may be useful for dose-finding studies, establishing target engagement, or measuring responses to an acute stress. Emphasizing what was discussed above, it is important to recognize that the ability for a drug to

impede what would be considered a 'normal' behavioral response to a novel stressful provocation in an otherwise healthy animal should be interpreted with caution.

An overreliance on tests such as TST or FST may account for the high rate of false positives and "me too" drugs in antidepressant drug discovery. In contrast, models based on chronic stress have construct validity because they are characterized by changes in behaviors that do have parallels with symptoms of human depression, including behaviors related to the hedonic or reward-seeking state of experimental animals, such as in sucrose preference, female urine sniffing, or social interaction tests. Chronic stress reduces the preference for, or response of, animals to these naturally rewarding stimuli, and SSRIs or other monoamine-based antidepressants restore normal behaviors when administered chronically, but not when given acutely (Wilmer, 2016). Further, in parallel with the human antidepressant response, single doses of ketamine in rodents restore normal reward behaviors in a rapid and more persistent manner. These behaviors offer particularly attractive insight into brain function because we know much more about the key brain regions and pathways that mediate these behaviors in rodents, as well as their homologous circuitry in humans. For example, decreases in function in these circuits are prominent in preclinical research and are generally mirrored by changes in functional connectivity in human fMRI studies of depressed patients (Drysdale et al., 2017), raising the prospect of translational functional biomarkers of depressive states and their response to antidepressants. False positives may also derive from overly simplistic thinking of the genetic underpinnings of psychiatric disease. It is unlikely that changes in the expression of single gene products, as revealed in response to chronic stress in animal studies or from examination of human tissue, can by themselves explain complex neuropsychiatric diseases, especially when the human genetics suggest complex polygenic risk. In interpreting such results, greater skepticism, more rigor, and consideration of resilience and susceptibility, rather than a causal interpretation, is warranted.

Although the majority of the most promising novel and rationally identified antidepressants have not yet reached the clinic given the obstacles for clinical trials stated above, it is important to note that there are currently 32 antidepressant compounds or formulations in phase I, II, or III trials or pending FDA approval, increased from 20 in 2015 (https://mentalhealthdaily.com/2018/02/13/new-antidepressants-2018-drugs-in-clinical-trials/). These compounds are directed against up to 14 novel molecular targets for which there are no current therapeutics. All this suggesting that we are, in fact, headed in the right direction.

Summary

It is true that the field of psychiatry has failed to rapidly advance many medications with new mechanisms of action, such as novel antidepressants, for over five decades. The only non-monoamine-based, mechanistically *new* antidepressant drugs are ketamine and brexanalone — with very recent FDA approval for human use. While the discovery of ketamine's antidepressant effect was somewhat serendipitous, basic research and animal models were imperative in that discovery (Trullas and Skolnick, 1990), and remain critical for establishing ketamine's mechanism of action on which additional studies can direct attention toward development of additional drug targets. Specifically, the antidepressant effects of ketamine were observed when this drug was being used to model aspects of schizophrenia in humans. This clinical 'model' was explicitly guided by basic research establishing an N-methyl-D-aspartate (NMDA) receptor and glutamatergic mechanistic link between ketamine and other pro-psychotic hallucinogens. Concurrent with these clinical studies, animal research using a variety of stress models reported that stress exquisitely influenced glutamate neurotransmission and NMDA-receptor dependent synaptic plasticity, thus validating the hypothesis that an NMDA antagonist such as ketamine may have antidepressant-like effects. Further preclinical work with ketamine showed that its behavioral and corticolimbic effects may be produced by increased glutamate

neurotransmission mediated by non-NMDA receptors, leading to the discovery of multiple new molecular mechanisms that are currently thought to mediate the antidepressant effects of ketamine. Thus, without the preclinical work in multiple areas of research and model systems, identification of a new class of antidepressant would not have been possible.

The relative lack of novel drug development cannot be laid at the feet of animal models alone, but rather represents the total collective failure of all approaches used in psychiatry—animal genetics or cellular models, circuits and behavioral studies, human brain imaging, human genetics, and clinical trials. Indeed, the failure largely reflects the unique and incredible complexity of the human brain and its disorders compared with all other organ systems. Moving forward, the best path to break this impasse in drug discovery is to better integrate these various levels of analyses with a sharper focus on advancing basic discoveries into the clinic (**see Take Home Points in Text Box 1**). But let's not kid ourselves*, clinical translation is difficult*; it has taken more than four decades to advance precision medicine approaches in cancer and immunology, and the ultimate translational successes have relied upon very basic, fundamental research. Therefore, it is clear that we have a great deal more to learn about the brain before a renaissance of CNS drug discovery becomes possible; we cannot predict now from where these discoveries will arise. Investment in investigator-initiated basic research has been the guiding principle of research funded by the National Institutes of Health for more than half a century, and it should be continued in order to stimulate the most provocative and creative research necessary for understanding the brain and its complex disorders required for eventual disease prevention and therapeutic successes.

References

1. Akil H, Gordon J, Hen R, Javitch J, Mayberg H, McEwen B, Meaney MJ, Nestler EJ. Treatment resistant depression: A multi-scale, systems biology approach. Neurosci Biobehav Rev. 84:272-288, 2018.

2. Conn PJ, Roth BL. Opportunities and challenges of psychiatric drug discovery: roles for scientists in academic, industry, and government settings. Neuropsychopharmacol. 33: 2048-2060, 2008.

3. Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, Meng Y, Fetcho RN, Zebley B, Oathes DJ, Etkin A, Schatzberg AF, Sudheimer K, Keller J, Mayberg HS, Gunning FM, Alexopoulos GS, Fox MD, Pascual-Leone A, Voss HU, Casey BJ, Dubin MJ, Liston C. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nat Med. 23: 28-38, 2017.

4. LeDoux JE, Pine DS. Using neuroscience to help understand fear and anxiety: a two-system framework. Am J Psychiatry. 173: 1083-1093, 2016.

5. Nemeroff CB, Vale WW. The neurobiology of depression: inroads to treatment and new drug discovery. J Clin Psychiatry. 66:5-13, 2005.

6. Nestler EJ and Hyman SE. Animal models of neuropsychiatric disorders. Nat Neurosci. 13: 1161–1169, 2010.

7. Quadrato G, Brown J, Arlotta P. The promises and challenges of human brain organoids as models of neuropsychiatric disease. Nat Med. 22: 1220-1228, 2016.

8. Selye H. Stress and disease. Science. 122:625-631, 1955.

9. Trullas R and Skolnick P. Functional antagonists at the NMDA receptor complex exhibit antidepressant actions. Eur. J. Pharmacol. 185: 1–10, 1990.

10. Willner P. The chronic mild stress (CMS) model of depression: History, evaluation and usage. Neurobiol Stress. 6: 78-93, 2016.

Text box: Needs for the field

• ***In vivo experiments***. The use of laboratory animals to study human disease states, including neuropsychiatric diseases, continues to have significant merit. It remains inconceivable that new treatments could be developed and successfully advanced to testing in humans without some degree of validation in laboratory animals.

• ***Precision***. The concept of 'an animal model of a neuropsychiatric disease' is not precise language and needs to be refined. As a tool to bring about change, a helpful mnemonic is to change the routine terminology from 'animal models of condition X' to 'animal models *useful in the study of* condition X'. Precision in words matters and keeps us from overinterpreting, overpromising, and progating misconceptions.

• ***Integration and complexity***. We need to prioritize the identification and study of domains with cross-species homologies, such as molecular and cellular processes, neuroanatomy, circuit function, physiology, and objectively-defined behaviors.

• ***Innovation***. The general lack of innovation in developing animal models for the study of neuropsychiatric diseases is sobering when compared to advances seen in the capabilities, pace, and efficiency of molecular biology. One example of a new direction is strategic utilization of endpoints in laboratory animals that are homologous to those that can be derived from digital devices (e.g. smart phones, wearables), since these devices are increasingly prevalent and bring the promise of enabling predictions that will assist in diagnosing, treating, and preventing neuropsychiatric disease.

• ***Leadership setting the tone***. The degree to which leaders in the field are emulated—in their approaches, terminologies, and interpretations—is enormous and unquestionable. Imprecise language, sloppy study design, data overinterpretation, and inappropriate framing of behavioral outcomes are examples of scientific behaviors that are propagated and repeated, and that set precedents that are difficult to counter. We need to challenge ourselves and our colleagues to a higher bar. Thoughtful science is not only necessary, it is mandatory.