SHARE

Improving the Utility and Translation of Animal Models for Nervous System Disorders: Workshop Summary

DETAILS

BUY THIS BOOK

FIND RELATED TITLES

AUTHORS

Diana E. Pankevich, Theresa M. Wizemann, and Bruce M. Altevogt, Rapporteurs; Forum on Neuroscience and Nervous System Disorders; Board on Health Sciences Policy; Institute of Medicine
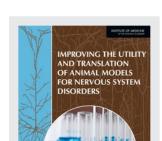
Visit the National Academies Press at NAP.edu and login or register to get:

– Access to free PDF downloads of thousands of scientific reports

– 10% off the price of print titles

– Email or social media notifications of new titles related to your interests

– Special offers and discounts

IMPROVING THE UTILITY AND TRANSLATION OF ANIMAL MODELS FOR NERVOUS SYSTEM DISORDERS

WORKSHOP SUMMARY

Diana E. Pankevich, Theresa M. Wizemann,
and Bruce M. Altevogt, *Rapporteurs*

**Forum on Neuroscience and
Nervous System Disorders**

Board on Health Sciences Policy

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

The serpent has been a symbol of long life, healing, and knowledge among almost all cultures and religions since the beginning of recorded history. The serpent adopted as a logotype by the Institute of Medicine is a relief carving from ancient Greece, now held by the Staatliche Museen in Berlin.

Suggested citation: IOM (Institute of Medicine). 2013. *Improving the utility and translation of animal models for nervous system disorders: Workshop summary.* Washington, DC: The National Academies Press.

"Knowing is not enough; we must apply.
Willing is not enough; we must do."
—Goethe

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

Advising the Nation. Improving Health.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

PLANNING COMMITTEE ON IMPROVING TRANSLATION OF ANIMAL MODELS FOR NERVOUS SYSTEM DISRODERS[1]

RICHARD J. HODES (*Co-Chair*), National Institute on Aging
STEVEN M. PAUL (*Co-Chair*), Weill Cornell Medical College
TIMOTHY COETZEE, National Multiple Sclerosis Society
MARK A. GEYER, University of California, San Diego
WALTER J. KOROSHETZ, National Institute for Neurological Disorders and Stroke
ALAN I. LESHNER, American Association for the Advancement of Science
GERARD J. MAREK, Abbott Laboratories
WILLIAM Z. POTTER, Foundation for the National Institutes of Health
SHARON ROSENZWEIG-LIPSON, IVS Pharma Consulting
TODD SHERER, The Michael J. Fox Foundation for Parkinson's Research
DAVID SHURTLEFF, National Institute of Drug Abuse
JUDY SIUCIAK, Foundation for the National Institutes of Health
STEVIN ZORN, Lundbeck USA

IOM Staff

BRUCE M. ALTEVOGT, Project Director
DIANA E. PANKEVICH, Program Officer
ELIZABETH K. THOMAS, Senior Program Assistant (*until November 2012*)

---

[1]Institute of Medicine planning committees are solely responsible for organizing the workshop, identifying topics, and choosing speakers. The responsibility for the published workshop summary rests with the workshop rapporteurs and the institution.

v

FORUM ON NEUROSCIENCE AND NERVOUS SYSTEM DISORDERS[1]

STEVEN HYMAN (*Chair*), The Broad Institute
SUSAN AMARA, Society for Neuroscience
MARC BARLOW, GE Healthcare, Inc.
MARK BEAR, Massachusetts Institute of Technology
KATJA BROSE, Neuron
DANIEL BURCH, PPDi
C. THOMAS CASKEY, Baylor College of Medicine
TIMOTHY COETZEE, Fast Forward, LLC
EMMELINE EDWARDS, NIH Neuroscience Blueprint
MARTHA FARAH, University of Pennsylvania
RICHARD FRANK, GE Healthcare, Inc.
DANIEL GESCHWIND, University of California, Los Angeles
HANK GREELY, Stanford University
MYRON GUTMANN, National Science Foundation
RICHARD HODES, National Institute on Aging
STUART HOFFMAN, U.S. Department of Veterans Affairs
THOMAS INSEL, National Institute of Mental Health
PHILLIP IREDALE, Pfizer Global Research and Development
DANIEL JAVITT, Nathan S. Kline Institute for Psychiatric Research
FRANCES JENSEN, University of Pennsylvania School of Medicine
STORY LANDIS, National Institute of Neurological Disorders and Stroke
ALAN LESHNER, American Association for the Advancement of Science
HUSSEINI MANJI, Johnson & Johnson Pharmaceutical Research and Development, Inc.
DAVID MICHELSON, Merck Research Laboratories
RICHARD MOHS, Lilly Research Laboratories
JONATHAN MORENO, University of Pennsylvania
ATUL PANDE, GlaxoSmithKline, Inc.
STEVEN PAUL, Weill Cornell Medical College
TODD SHERER, The Michael J. Fox Foundation for Parkinson's Research
PAUL SIEVING, National Eye Institute

---

[1]Institute of Medicine forums and roundtables do not issue, review, or approve individual documents. The responsibility for the published workshop summary rests with the workshop rapporteurs and the institution.

vii

JUDITH SIUCIAK, Foundation for the National Institutes of Health
MARC TESSIER-LAVIGNE, The Rockefeller University
WILLIAM THIES, Alzheimer's Association
NORA VOLKOW, National Institute on Drug Abuse
KENNETH WARREN, National Institute on Alcohol Abuse and
  Alcoholism
JOHN WILLIAMS, Wellcome Trust
STEVIN ZORN, Lundbeck USA
CHARLES ZORUMSKI, Washington University School of Medicine

IOM Staff

BRUCE M. ALTEVOGT, Forum Director
DIANA E. PANKEVICH, Program Officer
ANDREW POPE, Director, Board on Health Sciences Policy

Reviewers

This workshop summary has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published workshop summary as sound as possible and to ensure that the workshop summary meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the process. We wish to thank the following individuals for their review of this workshop summary:

Nathalie Breysee, Lundbeck Research, USA
Malcolm MacLeod, University of Edinburgh
David Shurtleff, National Institute on Drug Abuse
Rae Silver, Columbia University
Mark Tricklebank, Eli Lilly and Company
Bart van der Worp, University Medical Center Utrecht

Although the reviewers listed above have provided many constructive comments and suggestions, they did not see the final draft of the workshop summary before its release. The review of this workshop summary was overseen by **Floyd E. Bloom,** The Scripps Research Institute. Appointed by the Institute of Medicine, he was responsible for making certain that an independent examination of this workshop summary was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final

ix

content of this workshop summary rests entirely with the authors and the institution.

x

Contents

xi

1

Introduction and Overview[1]

Nervous system diseases and disorders are highly prevalent and substantially contribute to the overall disease burden. Despite significant information provided by the use of animal models in the understanding of the biology of nervous system disorders and the development of therapeutics; limitations have also been identified. Treatment options that are high in efficacy and low in side effects are still lacking for many diseases and, in some cases are nonexistent. A particular problem in drug development is the high rate of attrition in Phase II and III clinical trials. Why do many therapeutics show promise in preclinical animal models but then fail to elicit predicted effects when tested in humans?

On March 28 and 29, 2012, the Institute of Medicine Forum on Neuroscience and Nervous System Disorders convened the workshop "Improving Translation of Animal Models for Nervous System Disorders" to discuss potential opportunities for maximizing the translation of new therapies from animal models to clinical practice. The primary focus of the workshop was to examine mechanisms for increasing the efficiency of translational neuroscience research through discussions about how and when to use animal models most effectively and then best approaches for

---

[1]The planning committee's role was limited to planning the workshop, and the workshop summary has been prepared by the workshop rapporteurs as a factual summary of what occurred at the workshop. Statements, recommendations, and opinions expressed are those of individual presenters and participants, and are not necessarily endorsed or verified by the Institute of Medicine, and they should not be construed as reflecting any group consensus.

1

the interpretation of the data collected. Specifically, the workshop objectives were to

- Discuss key issues that contribute to poor translation of animal models in nervous system disorders.
  - Examine case studies that highlight successes and failures in the development and application of animal models.
- Consider strategies to increase the scientific rigor of preclinical efficacy testing.
  - Explore the benefits and challenges to developing standardized animal and behavioral models.
  - Identify methods to facilitate development of corresponding animal and clinical endpoints.
- Identify methods that would maximize bidirectional translation between basic and clinical research.
- Determine the next steps that will be critical for improvement of the development and testing of animal models of disorders of the nervous system.

ORGANIZATION OF THE WORKSHOP AND SUMMARY

The first session of the workshop reviewed the current state of animal models for the study of nervous system disorders (Chapter 2). This was followed by concurrent breakout group discussions of the application of animal models to a number of specific diseases and conditions, including neurodegeneration, Alzheimer's disease, stroke, schizophrenia, addiction, and pain. Upon reconvening the full workshop, the moderators for each breakout group summarized discussion highlights (Chapter 3). The next three workshop sessions expanded on the issues raised in the breakout group discussions regarding the utility of standardization of animal models (Chapter 4), corresponding animal and clinical endpoints (Chapter 5), and the translational gap between preclinical and clinical studies (Chapter 6). In the closing session, the main themes from the presentations and discussions were summarized by the session chairs (Chapter 7).

Topics Highlighted During Presentations and Discussions[2]

Several main issues were discussed across multiple workshop presentations and discussions:

- **Utility of animal models:** Throughout the workshop, many presentations and participant remarks highlighted the utility of animal models in understanding of fundamental pathology and disease pathogenesis. Animal models have also been an important component of the development pathway for nervous system disorder therapeutics. In particular, animal models have aided in the validation of potential therapeutic targets and assessment of pharmacodynamics. For example, these models have demonstrated that functional components of hippocampal circuits are similar between aged rats and humans. Measurements of prepulse inhibition in animals have been used as a predictive model of antipsychotics in humans due to similarities across species in response characteristics.

- **"Animal model" or "model animal"?** During the workshop many participants reiterated that there are really no animal models that fully mimic or recapitulate human diseases or disorders. Rather, models of a particular aspect or specific target of interest for any given nervous system disorder have been generated through genetic, surgical, or pharmacological manipulation or through other means in whole animals. While these models cannot replicate every aspect of a complex human disease, they are useful for dissecting out particular mechanisms, confirming hypotheses and developing therapeutics. In this respect, nomenclature was raised as an issue. Several participants noted that improved stratification of models, to match specific components of disease mechanisms or phenotypes, might lead to a greater ability to answer well-defined questions and the eventual development of better therapeutics. Following on this point, there was some discussion about not labeling models as a model of a whole disease or disorder (e.g., an animal model of schizophrenia), but rather to label them for the hypotheses or mechanism they are testing. Some participants noted that the challenge of this approach is that symptoms are usually addressed in the context of

---

[2]Rapporteurs' summary based on the presentations and the summaries presented by the meeting and session chairs in the final workshop session.

the disease and that research and drug development requires a specific labeled indication.

- **Standardization versus best practices:** Standardization in the context of the workshop was discussed both in terms of development of similar animal testing paradigms across laboratories or the development of a single "standardized" model for a particular aspect of a disease or disorder. In the discussions about standardization, a range of opinions were expressed. Many in the breakout group on addiction, for example, supported the use of multiple models that employ different approaches, so that converging evidence could be gathered, rather than promoting one or two select models. In contrast, some participants in the neurodegeneration breakout group emphasized the need for standardization of animal models of neurodegenerative diseases. Across the different research areas a variety of concerns were raised, one concern was that premature standardization might not be helpful and could stifle innovation by constraining research to specific areas or models. Some participants believe that over-standardization might artificially inflate sensitivity and reduce generalizability, resulting in clinically irrelevant findings.

- **Reproducibility of research:** Challenges associated with the reproducibility of animal model research were discussed. The standardization session began with a discussion about the increasing difficulty of reproducing published studies and highlighted a recent article that explored this issue (Prinz et al., 2011). Many participants agreed that improving experimental protocols and statistically analyzing experiments appropriately might help in moving the field forward and increasing reproducibility. A few participants noted that increasing the level of information detailed in publications, including the explicit reporting of replication of experimental results, would also be beneficial. Several participants emphasized the importance of reproducibility and the scientific validity of data as decisions are made about when to move from the laboratory into clinical trials.

- **External and internal validity:** The focus of the workshop was on the utility of animal models in the translation of basic research to human diseases and disorders. Workshop participants spent the majority of the workshop discussing the *external validity* of animal models or the extent of which research that uses animal models can be correctly generalized to human diseases

(van der Worp et al., 2010). However, as evidenced by the numerous comments by speakers and participants, external validity can only be discussed in combination with *internal validity* or the extent to which the design and conduct of a research experiment eliminates the possibility of bias (van der Worp et al., 2010). Many participants discussed the importance of randomization of animals to treatment groups, blinding of treatment assignments, sample size calculation, and outcome measures; all these are factors that might negatively impact the internal validity of the experimental results.

- **Bidirectional translation:** Many participants believed that animal models and human clinical research inform each other through bidirectional translation, from the bench to bedside and back again. Several participants discussed the importance of cross-validation of animal model endpoints with clinical measures and of clearly establishing what the corresponding endpoint is intended to predict. It was suggested that failures in clinical trials might be due to a mismatch between endpoints used in preclinical animal studies and those used in trials. A few participants expressed the view that there might be particular utility in the use of imaging and biomarkers in both animal models and human clinical research. However, it was noted that in some diseases, biomarkers assayed in humans do not have parallels in animals. Some participants suggested that data collected from negative outcomes, either in animal or preclinical studies, might also help to improve translation.

- **Collaboration:** Collaboration among sectors was discussed as a critical component in efforts to bridge the gap between preclinical and clinical research and to facilitate bidirectional translation. Several examples of precompetitive alliances[3] and cross-sector collaborations were discussed, both lateral efforts (across institutions and companies) and vertical (from preclinical through to clinical). Examples included a government-funded consensus building initiative; a government-facilitated, precompetitive public–private partnership designed to address specific issues in drug development; a for-profit consultancy bringing validated models from academia to drug developers; and the use of quanti-

---

[3]Precompetitive alliances are collaborations among competitors to achieve goals that can be more effectively accomplished by a group effort and have the potential to benefit all groups involved (IOM, 2010).

tative systems pharmacology as a translation tool, applying mathematical model–based decisions to support drug development. Many participants said that the inclusion of regulators in early conversations around validation and qualification of endpoints might also help to facilitate translation and the eventual development of new therapeutics.

- **Communication:** Many participants emphasized the need for similar terminology across basic and clinical research to facilitate translation of success in animal models to success in patients. It was noted that cross-disciplinary collaborations already require an understanding of each other's vocabulary. There was also discussion on the publication of animal studies, particularly concerns about the impact of bias were discussed (e.g., failure to publish negative results, publication of poorly designed or executed studies). Several participants mentioned the need for honesty in discussions about the predictive value of any given model. Several participants noted that in some cases it is not the models that need to be improved, but the dialogue about the models needs to be forthcoming. Of particular note, was the need to establish realistic expectations about what animal models predict.

- **Training:** Throughout the workshop, issues related to experimental design and statistical analysis of results were discussed as they related to the ability to translate data produced from animal models into potential therapeutics. Many participants agreed that these issues might be resolved with greater emphasis of these topics during training of graduate students and postdoctoral fellows. Several participants expressed concern over the deficiencies that current trainees have in basic research design and statistical analysis which are prerequisites for conducting any experiment. Many participants highlighted specific examples of areas where knowledge of and focused training on might be beneficial, including identification of primary experimental hypotheses and outcome measures; the importance of predefining analyses plans and inclusion and exclusion criteria; completion of sample size calculations; and identification of dependent and independent variables.

- **Animal model alternatives:** The workshop closed with a spirited discussion about the overall value of animal models as part of the therapeutic development pathway. The decreasing confidence in the ability of animal models to predict efficacy of drugs with

novel mechanisms of action was emphasized by several workshop participants. Concerns were raised about the potential that animal models may, in fact, be screening out potentially effective compounds. In particular, one participant noted that if the mouse models for polygenic psychiatric disorders are really as poor as described, perhaps it is time to ask under what circumstances it would be both worthwhile and ethical to go straight into human clinical trials after establishing safety. Should research continue with the current emphasis on animal models, or should more focus be placed on cellular models and/or Phase 0[4] human clinical trials? No answer was provided to this question, but many participants agreed that exploring methods to speed therapeutic development to first-in-human trials would require changes in the way current animal models are used. A few participants cautioned that, at minimum, this approach would not be successful without validated targets with the potential for efficacy and clinical benefit. Several participants also suggested that using animal models in conjunction with new emerging tools and technologies might help in the creation of a complete picture of disease pathophysiology and mechanisms, which would better aid in the creation of new pathways for the development of therapeutics.

---

[4]An exploratory investigation new drug study, or Phase 0 study, is conducted very early in Phase I and involves microdosing in a very limited number of human participants. See *Guidance for Industry, Investigators, and Reviewers: Exploratory IND Studies*, http://www.fda.gov/downloads/drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM078933.pdf.

2

Evaluation of Current Animal Models

Session chair Stevin Zorn, executive vice president of neuroscience research at Lundbeck Research USA, opened the session with the following questions: What leads to poor translation of animal models of nervous system disorders to clinical practice? Is it the models themselves, researcher expectations of models, how models are used to make predictions and decisions, or perhaps the level of knowledge about underlying pathophysiology for any given disease? Finally, Zorn asked, what is the impact of generalizing animal model capabilities, of poor study design, or of publication bias?

To set the stage for discussion, Steven Paul provided background on the challenges of drug discovery for nervous system disorders and why animal models are useful. Mark Tricklebank followed with a discussion about validation of animal models for drug discovery and how translation of preclinical research can be enhanced through skillful study design, planning, and proper statistical analysis; these points were echoed by other speakers and many participants throughout the workshop. Next, Katrina Kelner described three forms of publication bias that can impact the success of animal models: (1) the tendency to publish positive findings; (2) the publication of poorly designed and executed animal studies that could contribute to incorrect conclusions; (3) and assumptions about what constitutes "good science."

EXPECTATIONS FOR ANIMAL MODELS IN DRUG DEVELOPMENT

The past decade has been challenging for the biopharmaceutical industry. These challenges include the development and regulatory approval of innovative new medicines, said Steven Paul, director of the Helen and Robert Appel Alzheimer's Disease Research Institute at Weill Cornell Medical College and former president of Lilly Research Laboratories. It has been a particularly difficult time for neuroscience research and development. There have been therapeutic successes in select areas, for example, multiple sclerosis and epilepsy. However, in other areas of research, such as schizophrenia and depression, new drugs are not significantly better than those developed 50 years ago. For the most part, there are still no disease-modifying drugs for diseases like Alzheimer's and Parkinson's. Many pharmaceutical companies are restructuring and/or deemphasizing research on nervous system diseases and disorders, Paul noted, while some have completely left the field.

A major problem in bringing a new therapy to market is attrition, Paul said. When a new drug target is discovered and validated, lead candidate molecules are then identified and optimized. Preclinical testing evaluates the pharmacology and toxicology of the lead compound in animal models and Phase I clinical studies establish safety and dosing in humans. The problem of attrition occurs in Phases II and III, during the testing of whether the drug "works" as a treatment for the particular disease.

A 2010 study by Paul and colleagues found that roughly 66 percent of compounds that entered Phase II development did not advance to Phase III and about 30 percent of those that did enter Phase III failed to move on to regulatory submission (Paul et al., 2010). A more recent analysis suggests that, from 2008 through 2010, 82 percent of compounds failed to advance from Phase II clinical trials, with more than half failing due to issues of efficacy (Figure 2-1). Equally concerning, Paul said, is data from 2007 through 2010 show that 50 percent of compounds failed in Phase III, 66 percent of which were due to efficacy concerns (Arrowsmith, 2011b). By Phase III, Paul suggested, there should be confidence in efficacy and compounds should really only fail infrequently or due to rare adverse events. The approval rate for new drug applications (NDAs) by the U.S. Food and Drug Administration (FDA) is currently about 70 percent. In other words, 30 percent of new drugs that make it through Phase III to regulatory filing will not gain regulatory approval.

FIGURE 2-1 Failure of drugs in Phase II according to reason for failure (a) and therapeutic area (b).
SOURCE: Paul and Tricklebank presentations, March 28, 2012, from Arrowsmith, 2011a.

Animal models have long had an important role in the drug development process. As outlined by Paul, current expectations are that animal models can help researchers to

- Better understand the fundamental pathology and pathogenesis of a disease. An example is how genes and mutations result in observed disease phenotypes. This knowledge of underlying disease biology can aid in selecting and validating drug targets and defining pathways of intervention;
- Test drugs and treatments that could be effective for a specific disease or a particular symptom of a complex disease;
- Ascertain the safety of a drug or treatment, for example, toxicity or adverse events; and
- Establish the therapeutic index, or dose, that produces the desired effect compared to the dose that is toxic or lethal, for a given drug or treatment prior to testing in humans.

Paul concluded that although, animal models cannot solve all of the challenges of drug development, they are an important part of the solution. In particular, improvements in the ability of animal models to predict efficacy can help to reduce issues of attrition.

CHOICE AND VALIDATION OF ANIMAL MODELS FOR CENTRAL NERVOUS SYSTEM DRUG DISCOVERY

Mark Tricklebank, director of the Lilly Center for Cognitive Neuroscience at Eli Lilly and Company in the United Kingdom, said there are an unprecedented number of potential molecular targets in the nervous system, but there is little direct, clinically-based evidence to support a rational method for choosing one target over another. Selection of targets at the start of a drug discovery program would benefit from a solid biological hypothesis. This might be accomplished through manipulation of the target in vitro and development of a strong hypothesis of the expected in vivo results (Sarter and Tricklebank, 2012). Once a therapeutic target is selected, advances in in vitro screening technologies have made identification of potential drug molecules relatively easy. Tricklebank concurred with Paul that the system breaks down at the target validation stage, with too many false positives or false negatives.

Working Backward from Failure in the Clinic

Investigational new therapeutics are dismissed on the basis of the clinical result, most commonly, for lack of efficacy. However, Tricklebank cited a recent analysis suggesting that in many cases where the conclusion is lack of efficacy, the fact is that the hypothesis was not tested adequately and the results were not definitive (Morgan et al., 2012).

The way in which animal models are used may also contribute to failure in clinical trials, Tricklebank offered several possible reasons:

- Preclinical evidence supporting the hypothesis is given unrealistic weight in comparison to evidence against the hypothesis. Pressure on researchers to focus on positive results and ignore information that might need to be addressed and understood before deciding to move forward.
- Insufficient attention is paid to the design and analysis of experiments so that false positives or false negatives incorrectly influence decisions.

- Preclinical data collected to support the hypothesis are irrelevant to the mechanisms underlying the disease of interest. For example, not much is known about the pathophysiology of psychiatric diseases, and therefore it is difficult to make rational drug target choices.
- The preclinical data accurately conveys the drug responses in a very specific, genetically circumscribed population of animals maintained in highly controlled environments, but this does not necessarily hold true in the heterogeneous human clinical population.
- The measurements taken have, at best, only face validity for the variables that the experimenter would like to measure, and they need to be quantified in relation to the disease of interest. There is a need to understand in greater detail what researchers want to investigate, Tricklebank said.
- The measured effects are confounded by competing responses.
- There is a species gap, so the systems being manipulated in experimental animals might never fully predict outcomes in humans.

Expanding on the issue of experimental design and analysis, Tricklebank cited a study of 513 publications in top neuroscience journals that looked for appropriateness of statistical analysis (Nieuwenhuis et al., 2011). The study showed that 79 of these publications used an incorrect statistical approach. In a number of cases, the signal was so large that this error was irrelevant, but in two-thirds of the cases, this incorrect statistical analysis actually influenced the interpretation of the experimental results, Tricklebank explained.

Another issue is validation of the assay and model. Tricklebank outlined six types of validity to be considered in defining animal models (Box 2-1; see also Markou et al., 2009).[1] There is some confusion, he noted, about what an "assay" is and what a "model" is. An assay is a means of quantifying a dependent variable. A model is a theoretical description of the way a system, process, or disease works. An animal

---

[1]The types of validity described in Box 2-1 can be generalized under *external validity* or the "extent to which the results of an animal experiment provide a correct basis of generalizations to the human condition." Not included in this list, but discussed throughout the workshop by participants, is the concept of *internal validity* or the "extent to which the design and conduct of the trial eliminate the possibility of bias" (van der Worp et al., 2010).

BOX 2-1
Types of Validity Relative to Neuroscience Animal Models

Construct validity: Ideally mimicking the molecular and/or structural basis of the disease

Convergent validity: Evidenced by high correlation among performance patterns across cognitive tasks designed to measure the same neurocognitive process

Criterion validity: The ability of performance in one task to predict performance on another, more ecologically valid test

Discriminant validity: Evidenced by low correlation among outcomes across tasks designed to measure distinct neurocognitive constructs

Face validity: Degree of similarity to disease-specific symptoms

Predictive validity: Based on currently available therapy

SOURCE: Tricklebank presentation (March 28, 2012).

---

model induces over- or under-expression of a biological variable which the assay quantifies.

Improving the Probability of Clinical Success Through Validation

Ensuring that preclinical models have some validity for the system of interest starts with a better definition of the animal behaviors to be measured across the domains of sensory-motor function, arousal, affect, motivation, cognition, and social processes. Assays to measure these aspects of animal behavior should be designed to be as similar as possible to the assays used in the clinical situation, Tricklebank said. Also, assays measuring these functions in the clinical study should be as close as possible to those used in animals. Focusing on measuring the right thing and measuring it accurately is important. This includes maximizing signal-to-noise, removing confounds, determining both intra- and inter-laboratory reproducibility, and testing compounds against the most appropriate baseline perturbation.

Consider the manipulation: Is that manipulation clinically relevant; is it engaging the circuitry that is thought to be dysfunctional in the clinical indication of interest? From a practical perspective, Tricklebank suggested that increasing the view of psychiatric disorders as aspects of disturbed brain circuitry will lead to more rational profiling of compounds. Validation of potential compounds delivered via local injection, for example, would require evidence of engagement of the neurocircuitry involved in the disease. Biomarkers can also serve as indicators of the circuitry involved in the measured behaviors. Improving animal models might occur through the use of clinically relevant pharmacological, environmental, neurodevelopmental, and genetic methods to perturb or impair normal function.

In discussion, Sharon Rosenzweig-Lipson of IVS Pharma Consulting added that in addition to establishing the validity of an animal model, it is important to understand how decisions are made based on the model. How much of failure to translate is attributable to novel animal models versus models that are well defined (e.g., standards of care, known mechanisms)?

Collaborative Partnerships and Cross-Disciplinary Research

The design and validation of preclinical assays and models for drug discovery is ideally served by precompetitive, collaborative approaches via industrial, academic, and clinical consortium, Tricklebank said.

The Lilly Center for Cognitive Neuroscience, he explained, has adopted a very detailed approach to profiling a molecule though assays, disease models, tools, targets, and biomarkers as the basis for Phase I clinical studies. To do this, Lilly is leveraging external capacity, capabilities, and innovation all along the development pathway through a consortium of academic and industry scientists.

On a much broader scale, the European Union (EU) through its Innovative Medicines Initiative (IMI) launched the Novel Methods leading to New Medications in Depression and Schizophrenia (NEWMEDS) project, an international consortium of academic and industry scientists.[2] Tricklebank is involved in NEWMEDS Work Package 2 (WP2), which is focused on cognition assays and animal models. A key aspect of WP2 is multisite validation, establishing new assays in one laboratory and then running them in the partner laboratories to gauge reproducibility from

[2]Discussed further by Steckler in Chapter 6.

methodological and pharmacological sensitivity perspectives. One example, Tricklebank mentioned, involves the development of a touchscreen-based translational assay of animal cognition and validation of the pharmacological sensitivity.[3] Multiple laboratories within the consortium are using this touchscreen assay under standardized experimental conditions to compare the rate of task acquisition and response to drugs administered during task acquisition or posttraining.

Tricklebank noted that convergent approaches combining behavior and physiology are critical for the preclinical validation of drug targets. He concluded that cross-disciplinary approaches are essential for the preclinical validation of drug targets, and collaborative precompetitive approaches to verifying findings will pay off in the long run.

THE IMPACT OF PUBLICATION BIAS

The process of peer review and publication in established journals improves the quality of science and helps to filter out invalid or unimportant results. However, scientific literature is still subject to publication bias (Easterbrook et al., 1991; Sena et al., 2010; ter Riet et al., 2012). Katrina Kelner, editor of *Science Translational Medicine*, reviewed three types of publication bias.

Failure to Publish Negative Results

Sociological factors strongly influence what is published in the literature and what is not, Kelner said. A much larger fraction of papers that report positive findings are published than those reporting negative findings. The end result is that scientific literature can be distorted. Kelner illustrated this with a hypothetical scenario (see Box 2-2).

This problem has no easy solution. There have been several calls for the pharmaceutical industry to publish their preclinical data and results from failed clinical trials (Clozel, 2011; Rogawski and Federoff, 2011). Several new journals are dedicated to publishing negative results (e.g., *Journal of Negative Results in Biomedicine*, *Journal of Negative Results—Ecology and Evolutionary Biology*, *Journal of Articles in Support of the Null Hypothesis*). In addition, Kelner noted that *PLoS ONE* will publish studies with negative findings. To this point, one participant

---

[3]Discussed further by Bussey in Chapter 4.

BOX 2-2
Hypothetical Publication Bias Scenario:
Failure to Publish Negative Results

Twenty laboratories tested drug X and only one laboratory found that it lowers blood pressure and published their statistically significant results ($p < 0.05$). The rest of the laboratories did not observe any effect and did not publish their results. Because the p value was set at 0.05, there was a 5 percent chance of getting a positive result by chance alone. If researchers could look at the results of all 20 studies done all over the world together, they would have concluded that it is unlikely that drug X lowers blood pressure.

Therefore, based on the single positive result that was published, drug X appeared to have important clinical potential to control blood pressure. The results were recognized by peer/service user reviewers and endorsed by publication in a top-tier journal. Shortly after, another laboratory that worked on drug X tried to replicate the published study but could not. Its work was published in a lower-tier journal and far fewer people saw it. Three other laboratories that had found negative results decided to publish their work, but no journal was interested.

SOURCE: Kelner presentation (March 28, 2012).

---

suggested that scientific quality and transparency would increase if there were a widespread effort by journal editors to publish a certain percentage of articles each year that contain negative results.

During the discussion, a participant suggested that top-tier journals establish sections for "replications and challenges" where one laboratory's failure to replicate another's work can reach the same level of attention as the original article. This could be a data-driven commentary on articles that have been published. Kelner supported the idea of a site that could publish negative results following review for rigorous experimental design and analysis. A participant from the National Institutes of Health (NIH) noted that the concept of a repository of negative data has been proposed many times. The NIH has, in essence, such a repository of data from its funded studies and perhaps that knowledge of what works and what does not could somehow be tapped to help address publication bias.

One participant noted that another pathway toward publication bias might arise if agencies disproportionately fund labs with publications of positive results in high-impact journals; this might lead to greater fund-

ing of fewer labs that are emphasizing singular approaches. A participant followed up that this could potentially lead to a push for postdoctoral fellows and graduate students to seek out training opportunities in highly recognized labs leading toward further scientific biases.

Poorly Designed and Executed Studies

Another type of publication bias is the publication of studies that are poorly designed, executed, and/or analyzed, in which the conclusions drawn are invalid or not meaningful.[4] This results in proliferation of articles that are basically uninformative. Weeding these studies out of the submission pool can be difficult. Journals rely, by necessity, on the expertise of their peer reviewers and not all reviewers are aware of or appropriately trained in experimental design and data analysis. Kelner noted that medical journals have been ahead of preclinical journals in adopting specific, independent review criteria, for example, the Consolidated Standards of Reporting Trials or CONSORT guidelines (Schultz et al., 2010).

Journals can help address this, Kelner said, by enforcing quality standards for peer review and scientists can help journals by developing and disseminating relevant standards for their field. In addition, she said, funding agencies can require high-quality output from their funded scientists.

Cultural Assumptions About What Good Science Is

The third type of publication bias Kelner described results from cultural assumptions about what constitutes "good science." A common assumption is that the best science increases fundamental knowledge, not practical application. This belief is part of the cultural fabric of the scientific community, including study sections, mentors, and some top-tier journals, she said.

---

[4]The National Institute of Neurological Disorders and Stroke convened major stakeholders in June 2012 to discuss how to improve the methodological reporting of animal studies in grant applications and publications. The main workshop recommendation was that at a minimum studies should report on sample-size estimation, whether and how animals were randomized, whether investigators were blind to the treatment, and the handling of data (Landis et al., 2012).

Some journals, such as *Science Translational Medicine*, are embracing a new assumption, that is, the best science makes progress in improving the lives of human beings. From this new perspective, Kelner suggested, work that was previously thought to be dull becomes exciting, other studies become much less interesting, and the difficulties of doing research in humans might become worth tackling. This is not to say that science should stop doing fundamental research, she stressed, but rather, unexamined assumptions about what makes "good science" can introduce biases in what research is completed and which studies are published.

New Knowledge Calls for New Models

Neuroscience has changed significantly over the past few decades, moving from the study of simple systems, such as the squid axon, to imaging studies of complex systems such as humans. Scientists can conduct sophisticated, non-invasive studies in the living human brain to learn about diseases and the effects of treatments. In conclusion, Kelner challenged workshop participants to think about whether it is time to design a whole new set of animal models based on this knowledge, rather than trying to refine existing models. She suggested that resources be focused on advancing the understanding of neurological and psychiatric disease in people, and then, on the basis of that information, building new in vitro and animal models for drug development.

In response, Paul noted that most psychiatric drugs were discovered in humans, not in animals (Conn and Roth, 2008). Chlorpromazine, for example, was developed as an anesthetic sedative, but had a unique calming effect that was then tested and approved for use in psychotic schizophrenic patients. A participant added that if the mouse models for polygenic psychiatric disorders are really as poor as people think, perhaps it is time to ask under what circumstances it would be both worthwhile and ethical to go straight into human clinical trials after establishing safety.

3

Translation from Animal Models to the Clinic: Case Examples from Neuroscience Research

The first day of the workshop included concurrent breakout groups intended to facilitate in-depth analysis of the translational success of animal models in six areas of neuroscience research. Those areas are Alzheimer's disease, neurodegeneration, stroke, addiction, schizophrenia, and pain. In preparation for the subsequent workshop sessions, breakout groups focused their discussions on three key questions:

1. Would this research area benefit from a new or improved standardized animal model?
2. How well do animal model and human clinical endpoints correlate in this area of research?
3. What is needed to bridge the translational gap between animal models and clinical science in this area?

These smaller breakout groups enabled discussions about the role and effectiveness of animal models in the development of therapies for nervous system disorders. Following the breakout discussions, each group moderator summarized the main points of discussion for all attendees.

ANIMAL MODELS FOR ALZHEIMER'S DISEASE

In discussing current animal models for Alzheimer's disease it is important to think about the human phenotype and what is being modeled in terms of the animal phenotype. The moderator, Bradley Hyman, professor of neurology at Harvard Medical School, said that animal models of

21

Alzheimer's disease, based on the genetics of the disease and the closely related frontotemporal dementia, replicate at least some of the pathology. Researchers have been successful at modeling very specific aspects of Alzheimer's disease in the mouse (e.g., plaques, tangles). Although these are incomplete models of the human disease, they have been well received in the field as potentially relevant models for use in drug discovery.

Patients with Alzheimer's disease will display both amyloidopathy and tauopathy; however, scientists often focus, in a reductionist way, on one or the other in an animal model. A participant added that even though the anatomy in the mouse is different than the human, mutant tau mice are relatively good models in that they recapitulate tau-dependent neurodegeneration. This has led a number of companies to focus on antibodies that block tau-dependent neurodegeneration in these mouse models.

Hyman reiterated that mouse models are partial, or incomplete, models of the overall human phenotype. In an animal model, pathological changes are studied in the context of a unique and isolated event (i.e., lesion) over a relatively short period of time. Alzheimer's disease, as it occurs in humans, is the sum of how lesions occur and evolve over the course of many years or decades. Mapping where in the evolution of the human disease an individual mouse phenotype model fits is an important and often uncertain piece of information. Hyman questioned the hypotheses tested in humans that do not have exact correlates in animal models (e.g., differences in when amyloid deposition occurs between animal models and in human disease).

Several participants also discussed the use of imaging and fluid biomarkers in both animal models and clinical research (e.g., positron emission tomography or PET); ligands that can identify beta-amyloid load in the brains of humans; analysis of beta-amyloid tau and phospho-tau as biomarkers in cerebrospinal fluid. These types of biomarkers are now used for early diagnosis and to monitor disease progression in humans and many participants discussed the need to translate them back into animal models. As new therapeutics are examined, using similar biomarkers in both animal models and humans may allow for better translation of animal findings into humans. It was noted that the Alzheimer's Disease Neuroimaging Initiative (ADNI) is working toward Standardization of Alzheimer's disease biomarkers across the 57 ADNI sites, enhancing quality assurance, quality control, and better analysis of the clinical and imaging data in the ADNI public database.

In summary, Hyman emphasized the following about Alzheimer's disease animal models:

- Some mouse models exist that are close genocopies of inherited early-onset disease and have been well received as potentially relevant models.
- Mouse models are incomplete models of the human phenotype.
- Behavioral results in mouse models are not as robust as biochemical and neuropathological readouts.
- There is a need to better match animal models to the appropriate stage of human clinical disease.

Animal models are limited in terms of how far they can be extrapolated toward the human condition. However, compared to many other types of neurologic diseases, Alzheimer's disease research has some very promising successes that can be built upon.

ANIMAL MODELS FOR NEURODEGENERATION

Neurodegeneration research spans Parkinson's disease, Huntington's disease, ammyotrophic lateral sclerosis (ALS), and multiple sclerosis, for example. Robert Ferrante, professor in the departments of neurological surgery, neurology, and neurobiology at the University of Pittsburgh, noted that much of this breakout group's discussion centered on standardization of models and whether they accurately reflect neurodegenerative diseases. Ferrante suggested that current animal models for Huntington's disease and ALS may accurately reflect not only pathophysiological mechanisms of human disease, but also neuropathology and behavioral phenomena. For other disorders, however, it is much more difficult.

In addition to emphasizing the need for standardization of animal models of neurodegenerative diseases, participants in this breakout also discussed enforcing standards for preclinical studies in animals and raised concerns about the publication of research that cannot be replicated. It was noted by some participants that although the NIH has set standards for conducting animal research[1] and papers have described these standards, widespread adoption of these recommendations has been slow (Kilkenny et al., 2010).

---

[1]See http://www.ninds.nih.gov/funding/transparency_in_reporting_guidance.pdf.

Breakout session participants also discussed reevaluating the scientific approach to drug discovery for neurodegenerative diseases. In general, the target-centric approach to neurodegenerative diseases has failed during the past 50 years and there was discussion of a systems biology approach to disease research, as well as in silico models of disease.

In discussion whether animal models accurately reflect human neurodegenerative disease, the issue was raised as to whether animal model studies might be replaced with more Phase 0 clinical trials in humans. In this regard, there was a call for the identification of pharmacodynamic markers and biomarkers that are clearly reflective of the disease. For example, some noninvasive mechanisms, such as high-definition fiber tracking in traumatic brain injury and other disorders, reflect what is occurring in the brain and could be developed as a biomarker for many neurodegenerative disorders. Several participants in this group also noted the need for correlation between the biomarkers used in patients and in animal models.

Ferrante summarized the main points of discussion in this breakout session as

- There is a need for more standardization of models reflecting neurodegenerative disease in patients.
- The scientific approach to neurodegeneration research (e.g., target-centric versus systems-based, in silico, etc.) may need to be updated.
- Research could benefit from increased focus on Phase 0 human clinical trials.
- There is a need for pharmacodynamic markers and biomarkers that clearly reflect the disease.

ANIMAL MODELS FOR STROKE

The biology of ischemia is different from that of neurodegeneration, explained group moderator Constantino Iadecola, professor of neurology and neuroscience at Weill Cornell Medical College. The stroke process starts with an arterial occlusion, which can be reproduced effectively in animals. Several participants noted that animal models of stroke are generally predictive and adequate. However, the models are not perfect, and Iadecola said it could be argued that the mechanisms whereby a thrombosis or embolus forms in humans may not be mimicked exactly by the

surgical occlusion of an artery in an animal model. Nevertheless, the basic reaction of the tissue to the occlusion is fairly standard between different species.

Despite the relative suitability of the animal model for the study of stroke, clinical trials have not produced effective treatments and many pharmaceutical companies have scaled back or abandoned stroke research programs. A few participants suggested that failure in the clinic is partly because endpoints used in preclinical animal studies are different from those used in clinical trials. For example, animal studies often assess stroke volumes histologically using a chemical marker, which does not really reflect cell death, while clinical trials measure functional outcomes. Fortunately, Iadecola said, advanced technologies, such as diffusion-weighted imaging, can be done in animals and humans, allowing for better correlation between animal and human studies.

One participant noted that there is an ongoing disconnect between animal studies and clinical trials with regard to what happens in human stroke and how animal stroke data are obtained. For example, in some animal studies, the investigational drug is given before the stroke is induced and protection conferred. However, this is not possible in patients as drugs are administered 6, 12, or 24 hours after the stroke, with no effect.

As a result of discussions at a number of symposiums, organizations have been formed to address this issue. Researchers are working toward studying stroke in the animal models such that the studies mirror more closely what happens in the clinic. For example, Iadecola said, researchers are using more animals with risk factors for stroke, such as diabetic animals and aged animals as opposed to the younger "teenage" animals on which classic stroke work has been based. At the same time, clinical trials are taking into account the conditions under which protection has been observed in animal studies (e.g., sex of animals, time window of drug administration).

In summary, Iadecola highlighted the main points from this breakout session:

- Although neuroprotection has been demonstrated in numerous animal studies, treatment of humans has not been effective.
- Adequate animal models of stroke exist, but successful translation of the science from animal models to humans has been limited.

- The discordance between animal and human studies may be due to bias in study design (e.g., different endpoints) or to the failure of animal models to mimic clinical disease adequately.

ANIMAL MODELS FOR ADDICTION

Athina Markou, professor in the department of psychiatry at the University of California, San Diego, described the development of the smoking cessation drug varenicline as an example of successful development of a therapy for addiction. Markou speculated, however, that the success of varenicline in clinical trials is attributable more to the fact that there was a very strong theoretical rationale and less to the translation of preclinical animal studies. The animal models used were valid models, she said, but they were simple.

In contrast, despite the significant amount of research that has focused on dopamine, there are almost no drugs that have made it to the market for the treatment of addiction. Markou noted that there is disagreement as to why this is the case. Some argue that the hypothesis that dopamine mediates dependence and addiction for all drugs of abuse is incorrect. A counterpoint is that the clinical trials of dopaminergic drugs have not been done properly or perhaps have not been done at all. There are potential targets, such as the dopamine D3 receptor, but Markou noted there is currently little interest on dopaminergic targets by pharmaceutical companies. Other potentially good targets have not been explored sufficiently and breakout group participants discussed the need to incentivize industry and to educate pharmaceutical manufacturers that there is a significant market for addiction treatments.

Unlike some other nervous system disorders (e.g., schizophrenia where the etiology is not definitively known), the etiology of drug dependence is known to be excessive exposure to the drug. As such, animals can be similarly exposed to a drug and studied. Several breakout participants noted that models of addiction exist, although there is always room for improvement. For example, it was noted that additional emphasis is needed on more sophisticated models, such as models that examine the switch from drug experimentation to addiction or of relapse. A few participants in this breakout session raised concerns that a standardized animal model of addiction might not be the best approach. Rather, it was suggested that studying a variety of models that employ different approaches could provide converging evidence.

One issue for animal models of addiction is the heterogeneity of the human population with regard to addiction. The vast majority of people who experiment with drugs do not become dependent, suggesting a genetic component to addiction. In fact, genetic studies have provided some potential targets and it was suggested that one way to move forward is to try to over- or under-express these genes in mouse models and to study the heterogeneity of addiction development.

Finally, participants in the breakout session discussed concerns with clinical trials. As in other breakout groups, the need for cross-validation of animal model endpoints with clinical measures was noted. Also, current clinical measures for addiction studies were said to be inadequate in that the primary outcome measure is drug consumption—did the patient stop taking the drug or not? There are many processes that lead to addiction or to lack of abstinence that are not assessed in clinical trials, Markou noted. For example, is drug consumption the result of physical withdrawal or perhaps due to some cue that reminded the patient of the drug? For alcohol dependence, is controlled use of alcohol an acceptable endpoint? Patient compliance is also an issue. In many clinical trials for addiction, it is not clear whether the patients have actually taken the therapeutic drug and failure of the trial may be because the patients do not achieve adequate levels of the therapeutic drug in their system.

In summary, Markou highlighted the following points made by various participants in this breakout group:

- There are good animal models of addiction. Rather than standardization of models, many participants noted that the use of multiple models that employ different approaches could provide converging evidence.
- Genetic animal models may be helpful in understanding the heterogeneity of human addiction.
- The many potential therapeutic targets for addiction that have not been adequately researched and incentives for research in this area may be needed.
- Cross validation among animal models, clinical endpoints, and processes that are assessed in human trials is lacking.

ANIMAL MODELS FOR SCHIZOPHRENIA

In the breakout session on animal models for schizophrenia, much of the discussion focused on processes, explained breakout moderator Holly

Moore, associate professor of clinical neurobiology in psychiatry at Columbia University. Topics included the neurobiological processes that might underlie psychological processes disrupted in schizophrenia; the process of doing research; and the process of dialogue between clinicians and researchers using animal models.

Many participants in this breakout session believed that current animal models for schizophrenia, while informative, are not adequate. It was noted that there are useful assays of behavior and cognition and of the neurocircuitry mediating the cognitive process affected in schizophrenia. However, divergent opinions were expressed on how useful those assays are and whether it is necessary to assay neurocircuitry or whether looking for direct impacts of therapeutics on behavior is sufficient. Animal models are being developed to probe the neurocircuitry underlying cognitive deficits, as well as the basic processes underlying psychosis and negative symptoms in schizophrenia.

Moore noted that some breakout group participants thought that the path forward is to go back to clinical and epidemiological research and ask "what is wrong" in schizophrenia. One simple approach to answer this question would be to examine patient behaviors while imaging their brains. This would allow researchers to determine what is behaviorally and cognitively aberrant and what neurocircuits are activated in correlation with the observed deficits.

Armed with that information, researchers could develop assays in animals that have homology with assays used in the clinic. First, however, there needs to be reliable and objective assays for humans that can predict a clinically significant change such as worsening or improvement in the patient's clinical profile. For animal modelers, clinical outcomes such as reduction in symptoms based on subjective scales are not useful. On the other hand, an objective assay of cognition and behavior in humans without data on the clinical significance of these outcomes is also not helpful.

In some cases, objective assays in humans and the homologous assays in animals may be very similar. For example, prepulse inhibition measurement of sensorimotor gating is similar across animals and humans and is mediated by the same circuits in the brain. In other cases, assays that are guided by similar circuits do not look the same in an animal as they do in a human from a phenomenological point of view.

Moore noted that many breakout session participants thought that homology at the level of neurocircuits might be a useful starting point for dialogue between clinicians and researchers that use animal models about

what mediates symptoms or behavioral pathology. Others suggested that it is not necessary to understand how neurocircuits mediate an aberrant behavior; that it would be possible to have a reliable and validated assay of a problematic behavior from a psychological point of view (e.g., an assay of a sensorimotor deficit). Another concern raised by one participant was that many animal models of schizophrenia focus on primary pathology and not how drugs might act on or become a compensatory mechanism.

Finally, once there are assays in animals that have some homology with the assays used for humans and which have been shown to predict clinically significant outcomes or functional outcomes, are those animal models being fully used? Studying systems in control, or intact, animals is relevant for target validation and pharmacodynamics. Once studies in an intact animal establish that the drug is binding to circuits of interest and modulating both circuit activity and behaviors known to be mediated by that circuit, the question remains whether the drug will work on that same circuit and to the same extent in humans. Several group participants emphasizes that this is where an animal model of disease guided by epidemiology and symptomology is important.

Moore summarized the main points of this breakout session as

- Animal models and human clinical research inform each other.
- Control, or intact, model systems are useful for target validation and pharmacodynamics.
- Animal models of disease would benefit if they were guided by epidemiology and symptomology.
- There is a need for animal assays that are translatable and predict clinical outcomes and for assays to have some homology across species and determinants.

Ideally, once an appropriate animal model is in place, the clinical trials would be designed to ask the same questions that the animal models asked, using the same objective assays in the clinical trials that were chosen for use in the animal studies because, at the very beginning of the process, they were objective assays that had some clinical relevance.

ANIMAL MODELS FOR PAIN

A large number of animal models are used by pain researchers. Participants in this breakout session discussed the adequacy of these models and the appropriateness of the assays used relative to clinical outcome measures.

In some ways, the field of pain research is unique in that it is possible to mimic the initial inciting events, explained A. Vania Apkarian, professor in the Neuroscience Institute at Northwestern University and group moderator. Researchers can cause peripheral neuropathy, for example, and study diabetic neuropathic pain-related behavior in animals. Classically, the outcome measure in pain studies has been nociception, on the assumption that reflexive outcomes (e.g., sensitivity to touch or heat) correlate to some extent with the human condition.

Apkarian relayed that many group participants felt there were many useful animal models of pain and that a standardized model was not needed. Rather, to make the most of existing models, it is important to ask the right questions. As in other sessions, participants also discussed the need for biomarkers that can be assayed in humans and animals alike.

Human brain imaging studies are changing the field of pain research through investigation of chronic pain conditions in humans, Apkarian said. There was discussion about the need to start looking at correlates of chronic pain in animal models. As current models are essentially models of inciting a painful condition, the question has not been asked as to what is the causal or a critical parameter that induces the maintenance of pain. Chronic pain is not just nociception. Pain interacts with and reorganizes the brain. Injuries in humans may or may not lead to chronic pain, suggesting that something genetic in the brain needs to be considered in addition to the injury. Many participants indicated that much can be learned from genetic models in mice that might inform research on the human condition.

In summary, participants in this breakout session raised the following issues with regard to animal models of pain:

- Many existing animal models of pain might be more useful if researchers ask the right questions.
- Pain is more than just a sensation and appropriate measures are needed in existing animal models to address this complex issue.
- In particular, several participants were interested in identifying mechanisms for inciting pain versus maintenance of pain and

    understanding the mechanisms of chronic pain in humans.
- The usefulness of mouse genetic models and corresponding animal and clinical neuroimaging biomarkers was also discussed.

ANIMAL MODELS ADDRESSING NEURODEVELOPMENT

In the open discussion following the breakout group summaries, a participant raised another subarea of neuroscience research as an offshoot of the discussions of models for schizophrenia and addiction—animal models of what may essentially be developmental disorders.

Moore pointed out that although models of schizophrenia in adult animals are used for pharmacologic studies, knowledge of the epidemiology of schizophrenia has led to the development of models where the perturbation is made quite early in development, when risk factors for the disease come into play. Although the perturbations are made early in development, behavioral and neurological outcomes traditionally have not been studied until those animals were adults, presumably because that is when the disease emerges in humans. Only recently are researchers starting to think about looking at different stages in disease development and potential strategies for prevention.

Moore noted that people who are at high risk for the psychopathology associated with schizophrenia are not asymptomatic before they become psychotic. Rather, they have phenotypes that could be identified, characterized, and targeted for treatment (Kaur and Cadenhead, 2010). That treatment may delay or prevent psychosis might significantly impact functional outcome even though the person is still undergoing a psychotic episode. Researchers can start looking for signs earlier in people who have a first degree relative with schizophrenia and can look at prodromal patients who have been clearly identified as at risk using well-characterized and accepted scales, and ask what treatment is needed prior to the onset of symptoms.

Moore suggested that there should be less focus on predicting who may become psychotic and trying to prevent that and focusing more on treating the pathology affecting them at any particular time in their lives. There is a real need for a developmental perspective to schizophrenia, she said, and animal models can help elucidate this.

Many changes in the brain occur during adolescence, but that does not mean that adolescence is a pathology. Perhaps the parts of the brain

that are changing at the fastest rate during adolescence may be the most vulnerable. If those areas overlap with the circuits that are implicated in anxiety, drug abuse, or depression, for example, it may provide clues to the points of vulnerability in that circuit at that time. These are still basic research questions.

Markou noted that people who start tobacco smoking in adolescence have the hardest time quitting. Therefore, it would be important to extend animal models of addiction to this developmental stage as well.

4

Perspectives on Standardization

This session of the workshop was designed to explore the ways in which standardization might impact research that incorporates animal models. Perspectives were provided by three speakers who discussed the challenges of developing, implementing, and disseminating standards along with the potential benefits and risks. Andrew Holmes described some of the controversy surrounding standardization of behavioral models and shared examples of interlaboratory standardization studies that have led to differing results. Timothy Bussey described the development of automated testing methods that would reduce interference introduced by experimenters in both animal and human studies. Lennart Mucke presented several examples of successful translation of findings in animal models to human and offered reasons why translation sometimes fails. He also provided his perspective on how optimization of experimental procedure through best practices for preclinical research might be an alternative to standardization of models.

As background for the discussion, session moderator Walter Koroshetz, deputy director of the National Institute of Neurological Disorders and Stroke (NINDS), referred to several recent reports that raise concerns about the reproducibility of published scientific data. For example, researchers at the pharmaceutical company Bayer reported that of 67 projects the company acquired based on "exciting published data," two-thirds were abandoned in the target validation stage because Bayer scientists could not sufficiently replicate the published data (Prinz et al., 2011). Another report suggests that many of the published findings of positive effects in animal models of potential treatments for amyotrophic

lateral sclerosis are most likely "noise … as opposed to actual drug effect" (Scott et al., 2008).

Koroshetz offered his own perspective on some of the goals of standardization:

- Improve best laboratory practices to decrease the publication of spurious results.
- Facilitate the reproducibility of results.
- Facilitate the dissemination of valuable animal models into more laboratories.
- Improve comparability across studies using "identical" animal models (requires knowledge of laboratory-to-laboratory variability).
- Restore trust before disaster strikes.

He also suggested several potential risks to keep in mind:

- The increased burden posed by over-standardization could stifle innovation.
- Research might gravitate toward standardized models, thereby restricting development of better models or the testing of multiple models.
- There could be decreased generalizability due to convergence of studies on a limited number of standardized models.

Standardization is not an "all-or-none" question but rather a "when and how much" consideration, Koroshetz said, and he referred workshop participants to recent recommendations from NINDS for experimental design, minimizing bias, results reporting, and results interpretation.[1]

CHALLENGES TO STANDARDIZATION
OF BEHAVIORAL MODELS

Andrew Holmes, chief of the Laboratory of Behavioral and Genomic Neuroscience at the National Institute on Alcohol Abuse and Alcoholism, discussed standardization of behavioral models in the context of preclinical models and assays of anxiety. Holmes described several tests that have been the basis for much of the preclinical research in anxiety

---

[1]See www.ninds.nih.gov/funding/transparency_in_reporting_guidance.pdf.

over the past 60 years, including the open field test, elevated plus-maze, and light/dark box. He explained that these approach/avoidance tests are based on the simple premise that small prey animals such as rats and mice have an innate aversion to exploring open, brightly lit areas where the risk of predation is presumably high, yet at the same time they have a natural drive to explore novel, potentially fruitful environments where they might find food, mates, or new territory.

The conceptual framework of these tests is straightforward, but each laboratory conducting anxiety testing uses what they believe to be the best apparatus and testing approach. The question then, Holmes said, was whether this variability affects the ability to reproduce findings across laboratories and across studies. To illustrate the complexities of this issue Holmes highlighted three studies. As background, he noted that it has been known for many decades that genetically inbred, isogenic strains of mice differ in various phenotypes, including measures related to anxiety. Using these inbred strains restricts the amount of variability in the population and presumably increases the ability to detect influences due to an environmental or a procedural difference.

The first study Holmes described compared the results of standard tests and assays for anxiety across four different laboratories involved in a consortium project (Mandillo et al., 2008). It was acknowledged that differences in equipment and apparatus could be a possible confound in standardization. Each laboratory was allowed to use the apparatuses already in place, and there were no attempts to equate variables such as housing or the vendor from which the mice were purchased. In one test, for example, using the percentage of time spent in the center of the open field as a measure of anxiety-like behavior, there were marked differences between mouse strains within a laboratory. Yet, although the magnitude of the differences varied among laboratories, trends were preserved. The authors concluded that despite differences in equipment, vendors, and housing across laboratories, the results were reproducible and robust. They also suggested possible confounds that might limit tighter replication, including experimenter experience, animal husbandry, apparatus differences, and clarity of the standard operating procedure used.

In the second case highlighted by Holmes, the investigators went to "extraordinary lengths to equate the test apparatus, protocols, and all possible features of animal husbandry" that they could control (Crabbe et al., 1999, p. 1670). Across a battery of different tests, the sites sought to

ensure that testing was done in the same order, on animals of the same age, at the same time of day, etc.

In one test, mice were assessed for time spent in the open arms of the elevated plus-maze. The authors found that at one site, for example, BALB/c mice spent more time in the open arms than C57BL/6 mice, indicative of a lower level of anxiety-like behavior in the BALB/c mice. A different site found the exact opposite effect, with C57BL/6 mice showing lower levels of anxiety on the same test. In contrast, in another test measuring voluntary alcohol consumption, the results for both strains at these two sites were remarkably consistent.

Crabbe and colleagues concluded that despite their efforts to equate the testing environments across laboratories, there were still significant effects of site for nearly all variables. They also noted the challenges of behavioral research standardization because there are many differing opinions on the "best" way to assay behavior. As a result, they went on to say that "it is not clear whether standardization of behavioral assays would markedly improve future replication of results across laboratories" (Crabbe et al., 1999 p. 1672). This statement was quite provocative and distressed many researchers, Holmes noted.

A third study Holmes described asked if standardization is not beneficial, would systematic nonstandardization paradoxically improve reproducibility? To test this, Richter et al. (2010) attempted to mimic different laboratory environments in their own setting. They ran four different experiments, systematically varying two potentially influential factors in each. For example, they compared C57BL/6 and BALB/c mice in the open field test under standardized conditions and under similar conditions, but with two factors at play: the size of the housing cage and the illumination level during the test (small cage, high light; large cage, high light; small cage, low light; and large cage, low light).

They found that under the standardized conditions, the magnitude of the difference between strains was variable across the experiments, whereas in the experiments where select parameters were systematically varied, they observed remarkable consistency in the strain differences. This was also quite a provocative result and led to a lot of debate in the field, Holmes said. The authors reasoned that standardized experiments can generate spurious results because over-standardization can artificially inflate the sensitivity of the procedure to the point where researchers are more likely to find false-positive or false-negative results. So while the results may be clean, the generalizability of that result to other situations may be limited. The authors suggested that varying some conditions

(e.g., age of animals, housing conditions) may improve reliability and generalizability of results. They also noted that this may apply more broadly, beyond behavioral studies.

These three studies all have some merit and add appreciably to the debate about standardization, Holmes said. As noted by Crabbe et al. (1999), standardization will be difficult, because there are many ideas about what "best" practice is. Moving forward, Mandillo et al. (2008) suggested that taking the experimenter out of the experiment would reduce one source of variability, that is, automated equipment may help to reduce subjectivity in scoring. Even if standardization will improve the reproducibility of behavioral tests, Holmes concluded, we also need to develop novel endpoints that might be less liable to these issues.

DEVELOPING TRANSLATABLE COGNITIVE ASSAYS

Timothy Bussey, professor in the department of experimental psychology at the University of Cambridge, focused his comments on behavioral cognitive assessment in animal models.

Bussey offered his opinion on an ideal cognitive testing method:

- **Automated:** Advantages of automation include high throughput or the ability to test large cohorts of animals simultaneously across multiple behavioral measures; minimal experimenter contact with animals during testing; labor saving; consistency and accuracy of task parameters and measures; data saved automatically; standardization.
- **Non-aversive and low-stress:** Stress and/or aversive stimuli can affect behavioral testing. Minimizing both when they are not specifically part of the study is important.
- **Multidimensional:** Standardization of all tasks carried out in the same apparatus, using the same stimuli and rewards while requiring the same responses. In this way, an animal can be tested on a battery of cognitive tasks and establish a cognitive profile of that animal.
- **Translational:** Make tasks as similar as possible to those used to test human populations.

One approach to increasing translation of results from animal models to clinical trials is to start by looking at how humans are tested. Increasingly, automated tests are used for human cognitive testing; for example,

the Cambridge Neuropsychological Test Automated Battery (CANTAB), uses a touchscreen. This approach, Bussey noted, benefits from all of the advantages of automation described above. Touchscreens offer tight contiguity between stimuli and responses, increasing learning and minimizing confounds compared to approaches where a person must divide attention between the computer screen and keyboard. The CANTAB battery also uses nonverbal stimuli that could conceivably be presented to an animal in a similar testing situation.

In fact, researchers are using cognitive methods that present computographic stimuli to animals. Bussey described a study on Huntington's disease in which a mouse is presented with two pictures on a computer screen that can detect a touch by the animal's nose. The task is called visual discrimination learning, which is simply the discrimination between two stimuli, one novel and one learned, on the computer screen. A specialized apparatus incorporates a computer monitor, touchscreen and food magazine that dispenses a pellet when the animal responds to the challenge correctly.

In another example, Bussey described a test of spatial and non-spatial learning and memory, visual reversal learning, and attention using the triple transgenic Alzheimer's disease (3xTgAD) mouse model that showed attention impairment compared to wild type mice (Romberg et al., 2011).

Finally, Bussey shared data from paired associate learning tests used to distinguish among mouse strains with knockout mutations of scaffolding proteins associated with the N-methyl D-aspartate (NMDA) receptor (Nithianatharajah et al., 2013). In this test, the animal must learn that a particular shape belongs in a particular location. Although some knockout strains perform no differently from the wild type control animals, one strain (with a knockout mutation in postsynaptic density protein 93 or PSD-93) never achieved performance above chance levels.

The next step was to translate these mouse testing methods to humans. Most of the human subjects did eventually learn the task over the course of many trials, Bussey said. However, study participants who were known to have deletions of the disks large homolog 2 gene (DLG2) that codes for the PSD-93 scaffold protein, some of whom had schizophrenia, were generally unable to learn the task.

Although these are preliminary studies, they demonstrate the potential for relevant translation between the animal models and the human clinical studies by using an automated, standardized apparatus and methods.

STANDARDIZATION FROM THE PERSPECTIVE OF ALZHEIMER'S DISEASE MODELS

Experimental models are used to better understand nature, began Lennert Mucke, director of the Gladstone Institute of Neurological Disease at the University of California, San Francisco. Although mice are clearly not people, they face similar challenges (e.g., parenting, finding food, navigation). Experimental models need not simulate every aspect of a disease or disorder, but do need to have some critical features in common, he said.

Alzheimer's disease is a very complex condition that Mucke described as a multifactorial proteinopathy including, but not limited to, different assembly states of amyloid beta peptides, mislocalization of tau and alpha-synuclein, localization of apolipoprotein E (ApoE) both inside and outside of cells, inflammatory changes, and vascular changes. Models of Alzheimer's disease, including transgenic mouse models, have been very informative in dissecting this complexity, Mucke said. Much Alzheimer's research has focused on the structural alterations found in the human disease (e.g., amyloid plaque formation), as well as network disruption, synaptic deficits, and network failure.

Extrapolation from Animal Models to Humans

Mucke described several tests of learning and memory that can be used to evaluate therapeutic manipulations in animal models, including the Morris water maze for spatial learning and memory, novel object recognition and passive avoidance learning. As an example, a study by Cissé et al. (2011) compared control mice and human amyloid precursor protein (hAPP) transgenic mice and assessed the effect of hippocampal injection of a lentiviral vector overexpressing ephrin-B2 (EphB2), a tyrosine kinase that is depleted in hAPP mice and in humans with Alzheimer's disease. In all tests, untreated hAPP mice demonstrated learning and memory deficits, while hAPP mice treated with EphB2 performed similarly to controls. However, Mucke noted, these behavioral measures are sensitive to interference. For example, when the light/dark cycle in the animal housing facility fails and the lights stay on for extended periods of time, the animals become stressed and will not perform.

To study how navigational deficits relate to human dementia, Mucke and colleagues created a human maze in the hallways of their facility. Patients were tasked with route learning (forward and reverse), landmark

recognition, and photograph location and ordering. They found that approximately 70 percent of patients with early-stage Alzheimer's disease and 50 percent of patients with mild cognitive impairment got lost on reverse routing, a navigation deficit that could not be predicted from mini-mental state exam scores of these groups (deIpolyi et al., 2007). Mucke observed smaller right posterior hippocampal and parietal volumes in patients who got lost. Interestingly, this same right-posterior hippocampal region has been shown to be expanded in London cab drivers who have been on the job for a long time, together suggesting a strong association of this region with human navigation.

In comparing mice and humans, Mucke made several observations in the mice that had not been described previously in the human condition. For example, in the dentate gyrus of the hAPP mice, there was decreased calbindin and overexpression of collagen VI compared to controls. There was also activation of group IVA cystolic phosopholipase A2 (IVA cPLA2) and increased met-enkephalin in the hippocampus, and decreases in specific sodium channel subunits in the parietal cortex. Mucke subsequently looked for and found these same molecular abnormalities in humans with Alzheimer's disease, supporting the potential predictive value of these animal models (Verret et al., 2012).

One example of extrapolation of therapeutic findings from a mouse model to the human condition was the finding that immunization against amyloid beta clears amyloid plaques (Nicoll al., 2006; Schenk et al., 1999). As another example, Mucke described studies in mice that show how the antiepileptic drug, levetiracetam, when given chronically to hAPP mice, normalizes their long-term potentiation deficits in the hippocampus. There are also significant improvements, although not complete reversal, in performance in the Morris water maze. Studies in humans have shown that this drug also has beneficial effects in people with amnestic mild cognitive impairment.[2]

Increasing Success Through Best Practices

Having presented several examples of successful translation of findings in animal models to human Alzheimer's disease, Mucke offered a list of possible reasons why translation may fail:

---

[2]Discussed further by Gallagher in Chapter 5.

- species differences;
- aging issues (neurodegenerative diseases develop over decades in humans, much longer than the lifespan of a mouse);
- human disease is more complex than what is modeled in animals;
- general problems with human and animal studies; and/or
- faulty hypothesis underlying the model.

Mucke highlighted some particular problems with the translation between animal and human studies of Alzheimer's disease. First, the genetic heterogeneity of the patient population may obscure treatment effects. A better understanding of this heterogeneity might allow identification of subpopulations that respond to a particular treatment unlike the general population. Mucke added that animal models, in contrast to humans, typically have predictable genetic backgrounds and many other variables that can be controlled.

Another issue is that Alzheimer's disease is a multifactorial condition and the success of a cause-specific treatment depends on the relative impact of the cause. In other words, if a disease has multiple contributing factors with different weights of impact on overall pathogenesis, it may matter which factor is blocked and it may be necessary to block more than one to see a significant impact. If a transgenic mouse model simulates only one of the causes, a treatment effect may be observed, but in the larger context of the human disease, there may only be partial benefit.

Mucke referred workshop participants to a recent review on best practices for preclinical animal studies in Alzheimer's disease (Shineman et al., 2011). He offered his own list of what, in his experience, are important basic best practices for animal models:

- Blind-code all analyses.[3]
- Carefully match experimental and control groups (e.g., sex, age, other characteristics).
- Conduct rigorous statistical approaches.
- Reproduce experimental results in independent cohorts at different times.
- Use multiple outcome measures, including measures that are functionally relevant to humans.
- Regularly test animal models for quality control (e.g., genetic drift, loss of phenotype).

---

[3]Coding of data by someone other than the researchers so that analysis can be performed in an unbiased manner.

- Validate across models and in the human condition.
- To be useful, negative data require sensitive positive controls. False negatives are easy to obtain and they are as misleading as false positives.
- Do not ignore or suppress data that might contradict dogma.

With regard to standardization, Mucke stated that standardization makes sense for well-established basic principles. However, until truly perfect models and assays have been developed, optimization trumps standardization. Premature or overzealous standardization efforts can prevent progress, he concluded.

DISCUSSION

Following the presentations, participants expanded on the topics of reproducibility, statistics, and mouse strains as they relate to standardization and the use of animal models.

Reproducibility

In response to the Bayer paper described by Koroshetz, many participants, from both academic and industrial laboratories reported difficulty reproducing results. Different assays, models, and compounds all affect reproducibility. Mucke noted that complex sets of data are often hard to reproduce, and success may require sending researchers to the other laboratory to learn protocols. Certain aspects of these protocols may not be trivial, and failure to reproduce another laboratory's data may simply be that the methods are not adapted as carefully as they should be.

A participant cautioned against making generalizations about the inability to repeat work from academic laboratories. Mucke concurred and added that it would be helpful if there was a way to bring together the researchers who obtained discrepant findings and have them work through the discrepancies. Koroshetz noted that for spinal cord injury studies, when NINDS was not able to obtain the same results as the original investigators, they brought those investigators into the NINDS laboratories to help reproduce the results.

Mucke expressed concern that, while an academic researcher is usually committed to studying one area, such as behavior, for the length of

his or her career, industry researchers are often reassigned to new projects as company needs dictate. This turnover of scientists may impact the depth of expertise with animal models and/or methods used in a particular field of research. A participant with a pharmaceutical background countered that pharmaceutical researchers are equally dedicated scientists who work hard to maintain their expertise.

A participant pointed out that this workshop is focused on translation and exploring avenues to reduce clinical trial failures. Several participants continued to say that if it takes so much effort for two laboratories to reproduce the same finding, the likelihood of being able to then take that finding and reproduce it in a clinical trial would seem very low. While researchers may come together so that two laboratories using a model or test are doing everything the exact same way, with every variable controlled for, this might not be possible in the subsequent clinical trials. Perhaps the focus should be on profoundly reproducible findings and not those for which all experimental conditions need to be carefully nuanced.

Statistics

A concern was raised that most basic science graduate students and postdoctoral fellows do not learn statistics. In addition, journal editors face challenges finding statisticians who are available and able to review basic cell biology in manuscripts. Mucke emphasized the value of involving a biostatistician in both the planning of a study and in the data analysis.

Animal Models with Uniform Versus Outbred Backgrounds

Holmes noted that in the early days of transgenic animal models, there were concerns that genetically isogenic backgrounds would be artificially homogeneous and therefore not relevant. These concerns were countered by claims that it would be difficult to see positive results against a background of high noise.

Across all behavioral domains, researchers continue to use inbred isogenic strains, but, increasingly, are using only one or two specific background strains, with the vast majority of studies done using

C57BL/6 mouse strains. This does raise questions about generalizability beyond the model.

Complex genetic strains have been derived from crossing a dozen or so different inbred strains, producing a population that is genetically as complex as a human population, Holmes said. These strains have a place in research for answering specific questions, such as the genetic basis of a particular phenotype. The choice of strain depends on the question being asked. When studying the mechanism of a disease or the effect of a drug, there is value in controlling the genetic background of the animals. It was also noted that the same strain of animal (e.g., C57BL/6) from different breeders can have both genetic and behavioral differences.

5

Perspectives on Corresponding Animal and Clinical Endpoints

Sharon Rosenzweig-Lipson of IVS Pharma Consulting, and session chair, began by posing several questions for consideration: Will corresponding endpoints be useful for predicting clinical efficacy? What is an animal model meant to predict and what is the corresponding endpoint intended to predict? For example, decreases in amyloid beta can be measured both in animals and in patients with Alzheimer's disease. However, having this corresponding animal and clinical endpoint is not necessarily sufficient to make a prediction about the clinical efficacy of a potential therapeutic. Similarly, specific behavioral changes in animals may correspond to changes in humans, but these may, or may not, translate into a prediction of disease course. There is value to translation of clinical endpoints, she said, but it is important to understand what that value is.

ROLE OF MATCHING ENDPOINTS

Invited panelists used specific case examples to discuss the role of corresponding endpoints, and the impact of experimental parameters on corresponding endpoints and bidirectional translation. Neal Swerdlow described prepulse inhibition as an example of the ability to study the same endpoint in both an animal model and humans. Larry Steinman discussed experimental autoimmune encephalomyelitis (EAE) as an example of how differences in the way an animal model is tested can have profound differences on the findings. Michela Gallagher described how

45

neuroimaging tools have demonstrated that functional components of hippocampal circuits are very similar between animal models and humans.

PREPULSE INHIBITION

Neal Swerdlow, professor in the department of psychiatry at the University of California, San Diego, used the prepulse inhibition assay as a discussion case for the role of corresponding endpoints. Prepulse inhibition is defined as the automatic inhibition of the startle reflex, the contraction of the facial and skeletal musculature in response to an intense, abrupt stimulus, when the startling stimulus is preceded by a weak lead stimulus or prepulse. A primary measure of the startle reflex is movement of the orbicularis oculi muscle, or eye blink, often determined using surface electrodes attached to the muscles around the eye. In the laboratory, prepulse inhibition is an operational measure of sensorimotor inhibition, the inhibition of a motor response by a weak sensory event.

Prepulse inhibition is markedly diminished in a number of different neuropsychiatric disorders, including schizophrenia, Huntington's disease, Tourette's syndrome, Asperger's syndrome, fragile X syndrome, and obsessive-compulsive disorder (reviewed in Braff et al., 2001). In schizophrenia, for example, patients show deficits in prepulse inhibition regardless of whether the startling stimulus is a tactile (e.g., an air puff) or acoustic stimulus. Although this phenotype is not specific to schizophrenia, it is robust and replicable.

Commonalities and Differences

From an experimental perspective, the stimulus delivery and response acquisition hardware and software that are used for prepulse inhibition testing are very similar across species. The most obvious difference in testing is physical restraint; human subjects voluntarily sit in a chair during testing while mice are enclosed in a tube.

The response characteristics are strikingly similar across species, including sensitivity to stimulus parameters (e.g., prepulse, intensity, and interval), cross-modal inhibition, habituation, and latency facilitation. That is primarily because the startle reflex involves neural circuitry that is common across all mammalian species, Swerdlow explained (Swerdlow et al., 1999, 2008). There are some obvious, although rela-

tively subtle, differences in response characteristics; for example, axons are longer in humans and therefore reflex latencies tend to be longer. Waveform morphology can differ depending on whether the electrodes are collecting a whole-body response (as in the mouse model) versus a single muscle group (eye blink in humans).

There is also evidence that similar biological substrates are involved in regulating prepulse inhibition across species. There is interesting sexual dimorphism, he noted, with males tending to be more inhibited than females across species (mice, rats, and humans). Prepulse inhibition is also a highly heritable phenotype.

Predictive Models

Some of the practical uses of prepulse inhibition testing relate to its predictive validity of antipsychotic drug effects. Prepulse inhibition is disrupted by dopamine agonists such as apomorphine. Swerdlow described early work that he and others conducted which showed that administration of the typical antipsychotic, haloperidol, or the atypical antipsychotic, clozapine, produced dose-dependent normalization of prepulse inhibition in apomorphine-treated animals (Swerdlow and Geyer, 1993). This ability of a compound to reverse the disruptive effects of a dopamine agonist on prepulse inhibition in an animal has been used as a predictive model of antipsychotic efficacy in humans.

While establishment of these cross-species comparable endpoints has created robust systems for predicting clinical efficacy of antipsychotic therapies for schizophrenia, the larger picture is that the system facilitates the identification of "me-too" drugs, Swerdlow said. New compounds may have a different profile but drugs identified in using this system are all basically antipsychotics that affect positive symptoms.

Whether drug effects on prepulse inhibition have corresponding endpoints across species is dependent on many variables (e.g., species and strain, stimulus parameters, drug dose and route of administration, concomitant drug effects on startle magnitude, subpopulations). Most of these, Swerdlow explained, can be controlled or addressed post hoc to ensure matching endpoints.

For example, Long-Evans and Sprague-Dawley outbred rats have very different prepulse inhibition profiles. While Long-Evans rats show only slightly less prepulse inhibition than Sprague-Dawley rats, their response to a dopamine agonist is dramatically different. At shorter

prepulse intervals, Long-Evans animals show a robust potentiation of prepulse inhibition, while Sprague-Dawley animals show a profound disruption (Swerdlow et al., 2004). This is a highly reliable strain difference that was first viewed as noise, but is now known to be related to differences in biology. Importantly, studying the biology of this difference has led to much useful information that Swerdlow pointed out would not be known if only one standardized strain of animal was used.

Construct Models

There are corresponding endpoints across species in terms of the neural circuitry. The primary startle circuit is fairly constant, with more variability and interesting differences in the downstream circuitry, Swerdlow explained. Researchers have shown that activation of basal ganglia and cortical regions are relevant for regulating prepulse inhibition in rodents. Correspondingly, a number of human disorders that display deficits in prepulse inhibition have identifiable abnormalities within portions of the basal ganglia or limbic cortical circuitry. In other words, the human anatomy maps well to the rat anatomy of circuit regulation.

A number of developmental animal models will produce a deficit in prepulse inhibition, such as the neonatal ventral hippocampal lesion model. As one example of construct validity of these models, these deficits can be corrected in a dose-dependent manner with clozapine. This endpoint correlates well to the human condition, where control subjects display about half as much prepulse inhibition compared to patients with schizophrenia, a deficit that can be normalized substantially by clozapine (Kumari et al., 1999). Swerdlow also described corresponding endpoints relative to disease gene effects on prepulse inhibition across species. Patients with Huntington's disease, for example, show profound deficits in prepulse inhibition (Swerdlow et al., 1995), a phenomenon that has been reproduced in a transgenic mouse model of the disease (Carter et al., 1999).

The Role of Corresponding Endpoints

In summary, the conditions (e.g., eliciting stimuli, response acquisition) for studying startle and prepulse inhibition across species are nearly identical (with the obvious difference of physical restraint of animals);

response characteristics are comparable; and there is evidence for similar biological substrates across species. The prepulse inhibition assay has predictive validity in developing and testing the activity of antipsychotics and for developing typical and atypical antipsychotics.

That said, the anatomy, neural circuitry, and neural substrates of schizophrenia are very complex. While drugs can be developed to control some of the simpler symptoms of this disorder it is less clear that these predictive models will be helpful in developing interventions that offer long-term benefit in terms of function, interventions that may act through a completely different mechanism. The diffuse neuropathology in schizophrenia may reflect events very early in development, years or decades before patients seek medical intervention (Halliday, 2001). Swerdlow suggested that even a "perfectly corresponding" animal model cannot generate a therapeutic (drug, gene, protein, etc.) that will substantially restore healthy neural function to patients with schizophrenia, addressing the variable web of absent and misguided neural connections that have developed over the course of a lifetime. In developing therapeutics for schizophrenia, our "endpoint" should reflect these limits, he concluded.

EXPERIMENTAL AUTOIMMUNE ENCEPHALOMYELITIS

The EAE animal model was first described in 1933 by Rivers and colleagues while they were seeking to understand how some viral infections lead to neurologic reactions. Nearly 80 years later, EAE remains "one of the most enduring models of human disease," said Larry Steinman, professor in the departments of Neurological Sciences and Pediatrics at Stanford University (see Steinman, 2003).

EAE is not a model of multiple sclerosis, Steinman stressed, but there are similarities, and EAE is often used for the study of demyelination and examination of potential therapeutics for multiple sclerosis. Steinman shared two case examples of how experimental parameters can impact outcome; one is highlighted here.

Steinman shared the study of tumor necrosis factor (TNF) blockers for treatment of autoimmune diseases. One approach does not suit all autoimmiune diseases, he noted. It is known that blockade of TNF is effective in treating about 70 percent of patients with rheumatologic diseases (e.g., rheumatoid arthritis, psoriasis, inflammatory bowel disease), while it is not effective for treatment of multiple sclerosis and can even

exacerbate the disease in some patients (Steinman et al., 2012). Early research in this area, however, illustrates the impact of experimental design on outcomes.

Based on data showing an association between elevated levels of TNF in spinal fluid and disease progression in multiple sclerosis (e.g., Sharief and Hentges, 1991), it was thought that TNF blockade might have therapeutic value. In pursuit of this hypothesis, two groups published studies with very different conclusions. Feldmann and colleagues demonstrated control of established EAE in mice by inhibition of TNF using a monoclonal antibody (Baker et al., 1994). Another group, however, induced EAE in animals with a disruption in the TNF gene, suggesting that TNF was not essential for development of demyelinating lesions. They also found that TNF treatment reduced the severity of EAE in the animals (Liu et al., 1998).

Steinman explained that a potentially critical difference between the experiments was that the Feldman group used complete Freund's adjuvant in induction of EAE while the Bernard group used adoptive transfer. As it turns out, Freund's adjuvant induces production of TNF-alpha and causes leakage of the blood-brain barrier (Müssener et al., 1995; Rabchevsky et al., 1999). Steinman suggested that what Feldmann and colleagues may have seen was an amelioration of EAE due to the TNF-enhancing effect of the complete Freund's adjuvant. Both of these studies were "EAE experiments," Steinman pointed out; however, the results were completely opposite.

Around the same time, clinical studies of TNF blockade in humans with multiple sclerosis also showed exacerbation of disease instead of reduction of lesions (Lenercept, 1999; van Oosten et al., 1996). Steinman added that the TNF antagonist, Enbrel, now carries a label warning about the increased risk of demyelination.

Steinman concluded that an appropriate animal model should have a strong link to human disease. It is also important to avoid inclusion of any unnecessary steps or components, such as Freund's adjuvant. EAE can be a good model system. However, there are many different EAE models and Steinman noted that more than 10,000 publications on EAE are listed in PubMed.

Clinical Trials

The bigger issue, Steinman suggested, is not animal models, but how to facilitate faster, less expensive trials in humans. Animal models will

always be imperfect, Steinman opined. Even successful use of a model cannot predict all possible outcomes, such as the risk of PML with natalizumab treatment. Steinman went on to state that human clinical trials will ultimately guide therapy. He added that many approved drugs could be repurposed, but finding someone to support and conduct the trials can be challenging. Koroshetz and another participant noted that the National Institute of Neurological Disorders and Stroke has a Network of Excellence in Neuroscience Clinical Trials that was set up to facilitate testing of new therapies in patients with neurological disorders. Academic investigators, industry, advocacy groups, and others with a novel therapeutic can apply to conduct a study within the network.[1]

IMPROVING BIDIRECTIONAL TRANSLATION FOR NERVOUS SYSTEM DISORDERS

Expanding on the previous discussion by Swerdlow, Michela Gallagher, professor of psychology and neuroscience at Johns Hopkins University, agreed that prepulse inhibition is a good example of corresponding endpoints across species and can facilitate the study of neuronal circuitry. There are new in vivo imaging tools that provide information about the functional position of networks in the normal state and in disease states.

Gallagher described her work with an aged rat model of memory loss. Gallagher noted that although this is referred to as an animal model of aging, it is important to understand that it is not a surrogate of aging like some other models; it actually *is* aging. Studying the hippocampal circuitry using this model predicts circuit overactivity and its localization. Gallagher noted that in this model, there is no neuronal loss and the numbers of synapses are maintained in old animals with memory loss. The hippocampal circuit most affected in terms of integrity of synaptic connections is the entorhinal cortex layer 2, which sends input to the dentate gyrus and CA3. Electrophysiological recording experiments showed that CA3 pyramidal neurons have elevated firing rates in this animal model. In animals with memory loss, those neurons fail to encode new information (Wilson et al., 2003). In this condition, overactivity is a sign of dysfunction.

In the hippocampus of an aged brain, it is predicted that these neurons encode less distinctive representations when animals experience

---

[1]See http://www.ninds.nih.gov/research/clinical_research/NEXT-flyer.htm.

overlapping elements, a property referred to as pattern separation or the encoding of a new environment so it has little overlapping property. Because the majority of the synaptic inputs to CA3 neurons come from auto-associative networks,[2] this elevated activity drives a complementary process referred to as pattern completion. Cognitively, Gallagher explained, there is a shift from pattern separation to pattern completion, which does not encode something new, but rather retrieves something old.

This was observed empirically in aged animals with memory loss by recording an ensemble of neurons in the CA3 circuit. The recordings show spatially localized neuronal firing in a familiar environment. When young animals are moved to a new environment, they exhibit the phenomenon of pattern separation, either in terms of firing rates or neurons involved in that representation. This distinguishes one episode and environment distinctively from another. In aged rats with memory loss that are presented with a new environment, there is a failure of CA3 neurons to rapidly encode a new representation. Aged rats with normal spatial memory have CA3 neurons that encode new information comparable to young animals.

Gallagher explained that there are modalities for testing people that can also capture this kind of pattern-separation/pattern-completion process. One example is running a recognition task where patients identify a visual stimulus as old, new, or similar but not identical to something already seen, while undergoing functional MRI (Bakker et al., 2008). A correct response of "similar" reflects pattern-separation ability, while an incorrect response that the stimulus is "old" indicates pattern completion. Using this approach, Bakker et al. (2012) provide evidence for the role for the human dentate gyrus/CA3 region in pattern separation. High-resolution neuroimaging tools have shown that hippocampal overactivity in patients with mild cognitive impairment is isolated to the human dentate gyrus/CA3 region (Yassa et al., 2010).

---

[2]Autoassociative neural networks are feedforward nets trained to produce an approximation of the identity mapping between network inputs and outputs using backpropagation or similar learning procedures (Kramer, 1992).

Practical Application

In the aged animal model, the functional contribution of hippocampal overactivity is studied further by using a variety of treatments to try to lower activity (e.g., viral transfection of an inhibitory peptide, drugs). When overactivity in the CA3 region was reduced, the performance of the network improved, Gallagher explained. In the process, it was found that an atypical antiepileptic, levetiracetam, could restore behavioral performance and network function in aged animals with memory loss. Levetiracetam preferentially reduces the activity of neurons that are in burst-firing mode, and when old animals have increased firing rates, they generate more spikes per burst.

Gallagher and colleagues took this finding forward into a human study and found that a subclinical dose of levetiracetam reduced hippocampal overactivation and improved task-dependent memory performance in patients with mild cognitive impairment (Bakker et al., 2012). Therefore, it would appear that hippocampal overactivity is a condition of network dysfunction, not a compensatory beneficial recruitment of resources in the hippocampus.

Further experiments are needed, Gallagher said, and they will be complex. Analysis of data sets from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and other studies indicate that the degree of increased activation of the hippocampus predicts subsequent cognitive decline, and can predict conversion to Alzheimer's disease (Putcha et al., 2011). An unanswered question is whether there is any correlation or causal relationship between the loss of the layer 2 entorhinal cortex neurons that occurs in prodromal Alzheimer's disease and hippocampal overactivation, as these neurons form the input pathway to the dentate gyrus and the CA3 region. Analysis of ADNI data on cortical thickness of the entorhinal cortex shows that during mild cognitive impairment there is ongoing thinning of the entorhinal cortex that might represent the neurodegeneration that has been seen in autopsy.

In conclusion, Gallagher noted that this is just one example of how neuroimaging tools have allowed us to understand that the functional components of these circuits are very similar in their core functions across animal models and humans.

6

Addressing the Translational Disconnect

Despite the diversity of the animal models discussed thus far, session chair Mark Geyer, professor in the department of psychiatry at the University of California, San Diego, pointed out they all share some fundamental problems. In this session, panelists contemplated approaches that might help to bridge the gap between preclinical animal studies and human clinical trials. Geyer noted that new strategies are needed to better enable bidirectional translation of knowledge between preclinical and clinical researchers and development of common terminology and sharing of resources across disciplines and stakeholders.

Richard Ransohoff opened the session by discussing the experimental autoimmune encephalomyelitis (EAE) model a case example of some of the challenges of translating findings from bedside to bench and back again.

This was followed by presentations of current efforts to promote discussion about the translational disconnect.[1] Deanna Barch discussed the need to provide concrete opportunities for basic and clinical scientists to come together and learn from one another. Gerry Dawson described P1vital, a clinical research organization focused on standardizing, improving, validating, and sharing of human experimental medicine tools in the precompetitive space. Consortium members can then use these tools in their intellectual property–protected drug development programs. Hugo Geerts explained how quantitative systems pharmacology can be an effective translation tool, applying mathematical model-based decision support to drug development, and Thomas Steckler provided further

---

[1]This selection of presentations are not intended to be a complete listing of all programs addressing translational issues for nervous system disorders.

55

information on New Medications in Depression and Schizophrenia (NEWMEDS) as an example of a government-facilitated, precompetitive public–private partnership focused on developing new models and methods for drug development.

DEVELOPING BETTER ANIMAL MODELS OF ETIOLOGY AND PATHOPHYSIOLOGY

Richard Ransohoff, director of the Neuroinflammation Research Center at the Cleveland Clinic, discussed EAE and multiple sclerosis as a case example to illustrate some of the challenges in translating research from bedside to bench and back again.

Multiple sclerosis is a chronic, inflammatory, demyelinating disease that is specific to humans. There are no spontaneous conditions that occur in any other animal that recapitulate multiple sclerosis to the extent that it could serve as a model for study. As discussed by Steinman (Chapter 5), EAE is the most common animal model used for the study of multiple sclerosis. EAE as a model has become very highly refined, Ransohoff explained, evolving from inducing the disease within whole brain, to white matter, to myelin, to myelin proteins, and now, myelin peptides; and from studying monkeys, to rabbits and guinea pigs, to rats, to mice, and now most often, C57BL/6 mice. Studying EAE has helped to expand knowledge of autoimmunity and tolerance, immunity and inflammation, cytokines, chemokines, the blood-brain barrier, microglia in the brain, oligodendrocyte cell injury, and the capacity and mechanisms for remyelination. EAE is poorly predictive for therapeutics, however, as illustrated earlier by Steinman.

At the phenomenological level, mice with EAE lose weight and develop a range of neurobehavioral impairments. Importantly, EAE in C57BL/6 mice is a monophasic illness and therefore, the model lacks ability to study progressive or relapsing multiple sclerosis, making it a useful, but only partial, model. Against the spectrum of a lifetime with MS, the EAE model only captures a few weeks, recapitulating the initiation phase in which something breaks the tolerance and autoimmunity is expressed in the myelin, leading to a single demyelinating event.

Ransohoff mentioned other animal models of multiple sclerosis, each with varied strengths and limitations. There are viral models (e.g., Theiler's murine encephalomyelitis virus, mouse hepatitis virus, herpesvirus), but the focus of these studies is often on the virology, not

demyelination/remyelination mechanisms. Another model uses the mitochondrial toxin, cuprizone, which causes very predictable oligodendrocyte apoptosis followed by remyelination, and offers a good model to study demyelination/remyelination, he said. There are also emerging models with unknown potential, such as transgenic mice with inducible oligodendrocyte cell death. Finally, a new mouse model of spontaneously developing EAE does recapitulate many of features of relapsing-remitting multiple sclerosis (Berer et al., 2011). The model is difficult and slow, involving multiple transgenic mouse strains and manipulating the gut microbiome. Ransohoff suggested that while it may be the "EAE gold standard," it is unclear whether the platform will be useful for studying much other than the science of disease.

Ransohoff offered his thoughts on research practices found in some EAE studies in order to introduce concepts for discussion:

- If a mouse with EAE dies in the course of an experiment, some researchers include the dead mouse every day until the experiment is over. Ransohoff suggested it should be removed from the experiment. It is also incorrect to count a mouse that does not get sick every day of the experiment.
- Extensive reporting of downstream events in mice that do not show inflammation of the CNS is not informative. Animals that simply fail to become sick are not necessarily demonstrating neuroprotection.
- How limp the tail appears is not a useful parameter.
- Studies of prophylactic treatment are irrelevant for multiple sclerosis.
- There is often poor-quality tissue analysis of demyelination. For example, loss of Luxol Fast Blue staining for myelin does not necessarily indicate demyelination. Other factors can impact uptake of stains, including edema, axonal destruction, poor tissue preparation, infiltration of cells, etc.
- Lack of blinding during rating of disease severity and lack of correction for cage effects during interventions.
- Using phosphate buffered saline as the control for intraperitoneal injection of protein therapeutics instead of irrelevant antibodies or proteins.
- Comparing C57BL/6 mice from a commercial vendor with germline knockout mice of mixed background rather than littermate controls.

- Underuse of chemical demyelination (e.g., cuprizone) to study mechanisms of demyelination and remyelination.
- Failure to follow longitudinally the fate of authentically demyelinated axons.
- Underapplication of adoptive transfer.
- Minimal exploitation of spontaneous EAE models.

In closing, Ransohoff said that animal models have been helpful and have materially contributed to the science of disease. EAE is difficult to do well, and its application for direct drug development has been weak. The potential for remyelination or neuroprotective therapy is unknown. Although much has been learned about myelin production, maintenance, and repair, researchers do not necessarily know how to apply that knowledge.

To optimize the chances of translational success for this model, he proposed that there be a central or regional facility where EAE is done properly and in a standardized fashion, particularly for drug development. Discussion of such facilities should consider whether spontaneous EAE models would also be housed there and whether there should also be contract facilities for demyelination/remyelination models, where accurate quantitative tissue analysis is critical.

EFFORTS TO ADDRESS THE TRANSLATIONAL DISCONNECT

Opportunities to Bring Researchers Together for Consensus Building

Deanna Barch, professor of psychology, psychiatry, and radiology at Washington University, offered an illustration of what she called "the three-way problem" of linking basic and clinical science: The "basic animal researcher" studies spatial learning and memory using the Morris water maze. The "basic human researcher" may study the same mechanisms or constructs, but uses very different terminology and paradigms, and potentially different tools or methods. The clinical researcher may study thinking, memory, and functional abilities in patients with psychosis. This researcher is distinct from the other two and is not necessarily developing or using tools in humans that translate to the animal models, or vice versa.

Getting this trio of researchers to work together can be a challenge because there are many silos and very different languages. Barch offered several suggestions for enhancing interactions among basic and clinical researchers studying animals and human subjects:

- Provide concrete opportunities for basic and clinical scientists to exchange and develop ideas and to learn each others' "language." As examples, Barch mentioned the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) initiative,[2] the Cognitive Neuroscience Treatment to Improve Cognitionin Schizophrenia (CNTRICS) initiative,[3] and this Institute of Medicine workshop.
- Identify and clearly define the constructs of interest in terms that both the basic and clinical researchers can understand. Foster an understanding of how and why constructs have been operationalized differently across species.
- Provide concrete funding mechanisms to encourage basic and clinical research collaboration and funding support for tool development.
- Develop a consensus on what it means to measure a homologous process or construct and identify ways of improving that homology, either in existing paradigms or in new paradigms, with the goal of improving predictive utility.
- Raise awareness among the human researchers about what is needed to make their paradigms more amenable and attractive to animal researchers.
- Be willing to challenge existing ideas about animal models.

Bringing Validated CNS Experimental Medicine Methods from Academia to Industry

Seeing an opportunity to help bridge the gap between his clinical and preclinical colleagues, Gerry Dawson co-founded P1vital,[4] a clinical research organization focused on experimental medicine for central nervous system (CNS) disorders. Dawson is chief scientific officer of the

---

[2]See http://www.matrics.ucla.edu.
[3]See http://cntrics.ucdavis.edu.
[4]See http://www.p1vital.com.

company, which is headquartered in the department of psychiatry at the University of Oxford (in the United Kingdom [UK]). The specific idea, he explained, was to take the existing CNS experimental medicine methods for schizophrenia, depression, cognition, and anxiety that are used in academia and make them accessible for use by drug companies through management and coordination by P1vital, including the validation of models.

P1vital assembled an academic network of five UK academic institutions with expertise in CNS research (University of Oxford, Institute of Psychiatry London, University of Manchester, Cardiff University, and Imperial College London), and a precompetitive consortium of five pharmaceutical companies that provided financial and practical support for the validation studies (AstraZeneca, GSK, Lundbeck, Organon [a Merck subsidiary], and Pfizer Inc.).

Dawson described a consortium study on schizophrenia to illustrate the P1vital process. Schizotype is a psychological concept that describes a continuum of personality characteristics and experiences related to psychosis and schizophrenia. The basic hypothesis of the study is that high schizotypes can be identified from those who are low or mean scoring using a simple signal-detection test that exposes individuals to white noise and voices. High schizotype participants will have more false alarm responses (e.g., perceiving spoken human voices embedded in white noise when no voice is present) and faster-to-respond, or "jumping-to-conclusions" decision-making processes. Upon hearing voices, high schizotypes will display the same areas of activation on functional magnetic resonance imaging (fMRI) as patients having hallucinations.

College-age participants were given a nicotine patch as a putative cognition enhancer and a placebo capsule, or a placebo patch and either a placebo capsule, an amisulpride capsule, or a risperidone capsule. Combined data from three test sites showed very significant drug-by-group differential effects. In high schizotypes, amisulpride improved performance, and in low schizotypes it made it worse.

To try to back translate to animal models, Dawson and colleagues first brought forward an animal model system, the Morris water maze. Human participants were observed with fMRI while navigating a virtual reality water maze, or arena task, via 3-D goggles and a joystick. In a preliminary experiment, with only high and mean schiotypes, testing encoding and retrieval, hippocampal activation was observed in young males, but not in elderly males. In assessing schizotypes, it was found

that mean and high schizotypes performed the task comparably. High schizotypes showed decreased hippocampal activation, via fMRI, during encoding (i.e., fewer resources were recruited), and increased activation during retrieval compared with controls. This indicates that while there were no behavioral differences between the groups there were differences evident by fMRI. One participant noted that this would suggest a potential biomarker for schizophrenia that would be present prior to onset of symptoms.

Similarly, young people who have a family history of depression complete simple tasks such as N-back successfully, but fMRI shows they need to recruit more resources to do so. As these individuals get older, Dawson said, this will become more difficult and they will become more vulnerable. The question, then, is when to intervene with drugs and/or therapy.

Dawson noted that P1vital has a variety of methods validated thus far and the pharmaceutical consortium members are using them in their drug-discovery processes. He added that the pharmaceutical industry has many compounds with specific and well-characterized mechanisms of action that were assessed in primary efficacy clinical trials and subsequently shelved for lack of development resources. The next step for P1vital will be to select four to five unique compounds from these collections and assess them in a range of these validated experimental models for signals of efficacy. This effort will be sponsored by a European College of Neuropsychopharmacology (ECNP) led experimental medicine network of companies and academic groups. In this regard, Geyer added that AstraZeneca, which has decreased its focus on psychiatry research, has prepared a list of compounds that it will make available for research under an agreement with the company.

Quantitative Systems Pharmacology as a Translational Tool

Hugo Geerts, chief scientific officer for In Silico Biosciences, discussed quantitative systems pharmacology as a translational tool. To begin, he took stock of what the pharmaceutical industry could potentially learn from other successful industries, such as microelectronics and aeronautics. First, these other industries, Geerts explained, have formalized their "collective knowledge," which helps them to move from information to knowledge to building prototypes. These groups use an

advanced modeling and simulation approach to capture community-wide expertise and knowledge.

Next, these companies embrace complexity. The focus is on circuit analysis because the networks give rise to emergent properties that are not explained by single targets. In many cases, Geerts said, multitarget pharmacology is needed to really affect the complete outcome of an emergent property in the brain. However most pharmaceutical companies prefer to work on very selective molecules. Geerts added that nonlinear processes, including those going on in the brain at any given time, are very hard to capture without mathematical modeling.

In addition, these other industries focus more on failure analysis, learning from negative outcomes. In many cases with failed clinical trials, products are simply dropped and the results never published. But those failed trials offer a lot of information (e.g., Why did that trial fail? Was it the drug? The target? The patient population?).

How can the successes in other industries be applied to CNS research and drug development? As observed throughout the workshop, animal models are helpful in unraveling individual biological processes in great detail, Geerts said. They are less helpful for translating this knowledge into the clinic. There are a number of different reasons for this, he said. Human neural networks and clinical outcomes are much more complex than animal networks and experimental readouts. Drugs may behave differently in animal models versus humans, due to factors such as pharmacology, genotypes, comedication, or comorbidities. There is growing realization that complex CNS diseases need multitarget pharmaceutical approaches (Swinney and Anthony, 2011). Thus while the reductionist approach is good for understanding the biology of disease, it may not be the best approach for bringing drugs to the clinic.

A large amount of untapped clinical data exists for CNS drug development and a possible solution is an approach Geerts termed "computer-aided research and development." This approach combines the best of the animal world with the clinical world, in an animal-independent implementation, he explained.

One of the major issues in bridging the gap between biomarkers and clinical functioning is that in many cases, it is a very large jump from the individual biochemical change to the cognitive performance. In Alzheimer's disease, for example, the biochemical marker may be amyloid pathology, which affects the synaptic interactions. However, those synapses are part of pyramidal cells, and those pyramidal cells are part of

networks. Changes in emergent network properties are then assessed with clinical scales, such as cognitive performance.

Geerts listed several possible reasons why animal models might not always translate to the clinic:

- Differences in drug metabolism, both in clearance and absorption, and in the formation of metabolites.
- Different drug affinities for human versus animal targets, which Geerts noted is often underestimated.
- Absence of functional genotypes in animals, which in humans can affect pharmacology.
- Incomplete pathology in transgenic mice.
- Different neurotransmitter wiring. Human regions of the brain have many different receptor densities as compared to animals.
- Placebo responses in humans.

To address these issues, Geerts suggested a systems pharmacology approach. Systems pharmacology considers the interaction of a drug with its target not only at increasing structural levels (e.g., cells, organs, animals, humans, populations), but also at the cellular and multicellular networks levels, trying to simulate in a quantitative fashion the effect of an intervention on emergent properties, using mathematical models and simulations (see NIH, 2011, for further discussion).

To make this approach usable and actionable for pharmaceutical development, In Silico Biosciences leverages the preclinical data from academia over the past 60 years and integrates information on receptor physiology, CNS drug pharmacology, target exposure, human pathology, and imaging to create a computational neuropharmacology translational model bridging preclinical and clinical research. The goal is to have calibrated and validated platform applications to take compounds from pharmacology to clinical readout for different CNS diseases.

Precompetitive Public–Private Consortium for Methods Development and Validation

Thomas Steckler, head of Translational Research CNS at Johnson & Johnson, expanded on the discussion by describing NEWMEDS (Novel Methods Leading to New Medications in Depression and Schizophrenia), one project of the Innovative Medicines Initiative (IMI).[5] IMI is Eu-

---

[5] See http://www.imi.europa.eu.

rope's largest public–private initiative aiming to speed up the development of better and safer medicines for patients. Steckler noted that the lack of predictive validity remains the most frequent reason for attrition in drug development (Kola and Landis, 2004) and animal models are often blamed for failure of a compound to show efficacy in the clinic (Wang et al., 2011). In working to address this relative to the development of psychiatric drugs, the NEWMEDS consortium is focused on three key challenges:

1. How can we better understand the disorders (schizophrenia and depression) and make sense of the genetic and molecular findings, the pathways of the neuronal circuits, and symptom clusters?
2. Are there novel approaches that can improve success rates for new compounds taken into humans? Are there new animal models that would be more predictive of the efficacy of compounds in the clinic? Can imaging tools (e.g., PET, fMRI) and experimental medicine approaches be used to increase success?
3. Are there ways of doing clinical trials more efficiently (e.g., clinical databases, novel endpoints)?

Composed of work packages (WPs) that address issues across the drug development process, NEWMEDS creates a translational network. For example, WP07, "Identifying risk pathways via CNV (copy-number variation) genetics," is linked to WP04, "Cross-species and functional imaging models for drug discovery" to determine whether there are different activation patterns in patients depending on the CNVs they carry. Both are also linked to WP03, "Human cognitive testing," to the component of WP04 focused on rodent fMRI phenotyping, to WP01, "Linking animal and clinical models via electrophysiology," and to WP02, "Animal models of cognitive dysfunction that relate to clinical endpoints," as there are mouse models that carry these CNVs to be tested not only for behavior, but also electrophysiologically and via imaging. Proteomics and metabolomics (WP09) are used to help in characterizing the rodent model.

NEWMEDS activities also interact with other IMI projects, such as PharmaCog (focused on Alzheimer's disease) and the EU-AIMS (focused on autism). All three projects, for example, use touchscreen technology for development of model paradigms.

Activities within the consortium involve multimodal approaches; sharing of expertise, knowledge, methodology, and resources; and translation and back-translation. Across the consortiums there is overlapping molecular biology, neuronal circuits, symptom domains, and technologies.

Important aspects of NEWMEDS are replicability, reliability, and data sharing. Steckler noted that the consortium has lead to 15 published papers and 60 conference publications, as well as the development of Web-based tools, such as a clinical significance calculator for biomarkers in depression (Uher et al., 2012).

Steckler shared his personal assessment of the strengths, weaknesses, opportunities, and threats for the NEWMEDS initiative (see Figure 6-1). The strong scientific focus, multimodal approach, complementary capabilities, cross-species and back-translation activities, and high level of scientific and technical expertise are strengths of the NEWMEDS approach. The commitment to the initiative by the different partners and the openness to sharing data are the most valuable aspects of the consortium, he said. There are many opportunities to gain knowledge, communicate across partners, and share the workload.

A primary limitation, however, is that the initiative is limited to the European Union. Steckler also noted that there is room for improvement in the interactions between the different work packages. A participant added that although there is much U.S. interest in NEWMEDS, companies have raised concerns about the level of bureaucracy.

The most significant threat, according to Steckler, is the instability of the pharmaceutical industry, with three pharmaceutical partners already opting out of the consortium because they decided to drop some or all of their psychiatry research programs. They may still provide expert input, but they are no longer actively generating data.

Strengths

- **Scientific focus:** disease/neurobiology/symptom clusters
- **Multi-modal approach, complementary capabilities**
- Cross-species, translational and back-translational activities
- High level of scientific and technical **expertise** among partners
- Resource sharing
- Experienced collaborators, strong leadership
- High degree of **commitment, openness,** and collaborative attitude
- Right number and balance of partner organizations
- Appropriate budget estimates
- Duration of 5 years may enable progress

Weaknesses

- In-kind contributions **mostly limited to EU**
- EU bureaucracy
- Reporting requirements
- Scope for improvements in **interaction between WPs,** esp. preclinical–clinical
- Scope for improvements in cross-IMI project interactions
- Scope for improvements in data sharing
- Duration of 5 years may be too short to achieve goals given complexity of disorders

Opportunities

- **Knowledge gain:** difficult to achieve by individual organization
- **Communication, esp. among EFPIA partners**
- Confidence building in internal strategy and processes **(what works)**
- Streamlining of processes, avoidance of error repetition **(what doesn't work)**
- **Shared workload** of method development, **cross-site assay validation**
- Cost efficiency
- Potential for target identification/validation
- Highly motivating and stimulating

Threats

- Under-resourcing due to budget cuts in industry
- Project sometimes perceived as an additional burden if internal staff has to shoulder workload
- Risk to lose focus on internal projects
- Initial budget estimation relatively inflexible; limited scope for adjustments (not really an issue for NEWMEDS)
- Change of priorities of EFPIA partners (long duration)
- **Instability of pharmaceutical industry**
- With risk of increased burden for those remaining

FIGURE 6-1 Strengths, weaknesses, opportunities, and threats for the NEWMEDS initiative.
NOTE: EFPIA = Federation of European Pharmaceutical Industries and Associations.
SOURCE: Steckler presentation, March 29, 2012.

7

Summary of Workshop Topics

To bring the workshop to a close, each session chair provided a brief synopsis of his or her session, reflecting on recurring themes and thoughts for the future.[1]

CURRENT ANIMAL MODELS

There is much that can be done in animal models that cannot be done in nonanimal models and humans, and there are regulatory requirements for testing in animals before investigational compounds can be tested in humans. Studies in animals allow, for example, evaluation of drug targets, testing of pharmacokinetics, metabolism, distribution of investigational compounds, and prediction of the dose that will be maximally efficacious and yet still tolerable and safe. Animal studies expand the understanding of the nervous system diseases and disorders in a defined system and allow researchers to draw links to clinical pathways and formulate hypotheses for human testing,

In summary, said session chair Stevin Zorn, although we can learn much from animal models, it was emphasized in the presentations and discussion that what we learn depends heavily what questions are asked,

---

[1]The topics highlighted in this chapter are based on the summary remarks made by each session chair during the final workshop session and at the end of his or her session. Additional comments by participants related to the closing remarks are also included. As noted in Chapter 1, comments included here should not be construed as reflecting any group consensus or endorsement by the Institute of Medicine or the Forum on Neuroscience and Nervous System Disorders.

67

how those questions are asked, and how the results are interpreted. When failures in translation occur, some potential questions to ask are: was it the animal model itself, the analysis, the clinical trial, or another factor? In this regard, the importance of training in skillful study design and appropriate statistical analysis was discussed.

Humans are uniquely different from animals in many ways and there really is no such thing as an animal model of human disease (e.g., an "animal model of depression"). It is important to recognize that animal models are only models of what is being modeled. In other words, these animals are modeling specific perturbations in an effort to understand the biology and assess potential therapies.

Bringing a new drug to market has numerous challenges and significant costs and many potential therapies fail to meet expectations in Phase III clinical trials. Solving the problem of attrition of investigational new drugs will require a "renaissance," Zorn suggested. Preclinical and clinical scientists need to reunite and redefine what is needed to enhance translation from preclinical models to human clinical trials. Precompetitive alliances and cross-sector collaborations were mentioned as potential approaches. An issue that was raised by multiple participants was that effective, cross-discipline discussion will require a better understanding of each other's vocabulary.

Also discussed was the need for better definition of the behaviors and physiological measures examined in both the animal model and the human condition. Clearly defining what is measured is important in terms of gauging reproducibility and validity of model systems. There was much discussion about the importance of measuring "the right thing," but exactly what that is for a given model or disease was, and will continue to be, passionately debated.

Another factor impacting the use of animal models is publication bias, including the tendency in the literature to publish more positive than negative findings; the publication of poorly designed, executed, or analyzed studies that could contribute to uninformative conclusions; and cultural assumptions about what constitutes "good science."

To foster discussion it was suggested that perhaps it is time to design a whole new set of animal models based on emerging knowledge (e.g., from imaging and genetics), rather than trying to refine existing models. It was also suggested that much can now be learned from sophisticated, non-invasive studies in the living human brain that can inform development of animal models.

BOX 7-1
Potential Methods for Increasing Translation

- Development of hypothesis-driven experiments
- Sample size calculations
- Blind coding of analyses
- Randomization of treatment assignments
- Blinding of experimenter to treatment assignments
- Careful matching of control and experimental groups (e.g., sex, age, strain)
- Rigorous statistical approaches
- Reproduction in independent cohorts
- Multiple outcome measures
- Matching basic and clinical endpoints
- Regular testing for genetic drift or loss of phenotype
- Use of sensitive positive controls
- Control of testing parameters (e.g., time of testing, lighting conditions)
- Automated testing
- Low-stress, non-aversive testing environments

NOTE: This list was identified and summarized by the rapporteurs. This list is not comprehensive and should not be construed as reflecting any group consensus.

DISORDER-FOCUSED BREAKOUT DISCUSSIONS

To foster more in-depth analysis of the translational success of animal models, the second session of the workshop featured concurrent breakout discussions on six areas of neuroscience research: neurodegeneration, Alzheimer's disease, stroke, schizophrenia, addiction, and pain. Based on the breakout summaries provided by each group moderator when the full workshop reconvened (see Chapter 3), session co-chair Richard Hodes offered his perspective on the common themes and differences across the groups.

Variable State of Understanding of Pathophysiology

The understanding of underlying pathophysiologic processes is quite different for various nervous system disorders examined, Hodes said.

Our understanding of the genetic component of Alzheimer's disease, for example, has generated therapeutic targets that have been reproduced in animal models of particular disease features. This knowledge is both a strength and a weakness, Hodes cautioned. While there are potential targets to capitalize on, it is important to remember, as Zorn noted above, that the full disease has not been modeled, only the target of interest. This is in contrast to other conditions such as schizophrenia, where there was discussion about the usefulness of current models and defining exactly what should be modeled.

Therefore, the next steps for improving translation of animal models may differ depending on the current state of knowledge about the basic biology of diseases and disorders.

Adequacy of Models

Hodes recalled the summary from the stroke breakout group, where it was noted that adequate animal models of stroke exist, but translation of the science from animal model to humans has failed. It was suggested that this discordance between animal and human studies in stroke could be due to bias in both animal and clinical study design or to the failure of animal models to adequately mimic clinical disease. This was characteristic of much of the discussion in that there are a variety of possibilities to explain where the faults may be when models do not successfully translate. Hodes cautioned against letting the animal model become the "standard" and lamenting that the clinical state is not cooperating with the model.

Risk That Research Is Constrained by Models

As alluded to above, when models are considered adequate (e.g., arterial occlusion models of stroke) or compelling (e.g., transgenic mouse models of Alzheimer's disease pathology), there is a risk that research then becomes restricted to those areas or models. In the case of Alzheimer's research, for example, the amyloid hypothesis has dominated. Only very recently has there been more interest in tauopathy and in mouse models that recapitulate tau-dependent neurodegeneration. It is important to keep an open mind when working with established models, and not become locked into them in a counterproductive way.

Public–Private Cooperation

Hodes observed that the discussion of the strengths and weaknesses of collaborative efforts tended to group cooperation into two types: parallel partnerships and cooperative relationships that involve sequential roles. Some discussions, for example, focused on concerns that academic institutions and not-for-profit organizations carry out the basic research, and then hand it off to drug discovery and development, as if these are two discrete components in the process and not one continuous pathway.

STANDARDIZATION

Issues surrounding the standardization of animal models were discussed from the perspective of preclinical models of anxiety, cognitive assays, and Alzheimer's disease models. Session chair Walter Koroshetz pointed out an overarching concern for these and all models is that the process of tool development and standardization of models is not easily funded or staffed.

Koroshetz observed that inter-laboratory standardization may be more difficult for some tests than others. It is important to know what the different performance characteristics are for a particular test. For one of the studies, even with great efforts to standardize all parameters among test sites, there was inter-site variability in the results of one type of behavioral assessment (elevated plus maze), but consistency among test sites with another (voluntary alcohol consumption).

Another topic of discussion was concern about over-standardization artificially inflating significance and reducing generalizability. This was shown in a study in which heterogeneity was introduced systematically in the comparison of two strains of mice (by altering housing cage size, and lighting during testing). Under highly standardized conditions, the magnitude of the difference among strains was variable across the experiments. However, when select parameters were heterogeneous, there was remarkable consistency in the strain differences. There was also discussion about the use of primarily one or two particular inbred isogenic mouse strains for behavioral studies and whether that raises questions about generalizability beyond the model.

In discussing better ways in which to assess animal behavior, it was noted that there can be tremendous interference introduced by the experimenter. One approach is to "take the experimenter out of the experi-

ment," automating as much as possible to minimize experimenter contact with the animal and to reduce variability. The use of visual touchscreens for cognitive testing of both mice and humans was described as an example of standardization, and translation of testing methods.

The discussion also expanded on the point made in the prior two sessions that animal models do not simulate every aspect of a complex disease, but they are very useful for dissecting out particular pieces. This applies not only to molecular pathology, but also to underlying perturbations to networks and functional circuits.

Koroshetz reiterated a list of best practices for preclinical animal studies offered by Mucke, which included blind coding of all analyses and allocations, carefully matching experimental and control groups, rigorous statistical approaches, reproducing results in independent cohorts at different times and in different conditions, using multiple outcome measures, quality control of animal models, validating across models and in the human condition, and including sensitive positive controls to help eliminate false negatives. Several participants noted that a positive control need not necessarily be a compound, but rather, some demonstration that the assay is as sensitive as expected to pharmacological manipulations.

Overall, there was spirited discussion about the value of standardization, with concerns raised that premature standardization might not be helpful and can stifle innovation. It was suggested that improving experimental procedures may be the most helpful to the field going forward. Many participants stressed the importance of best practices, training scientists in well-established principles of experimental design and analysis (e.g., statistical power), and bringing researchers together to share and compare approaches.

CORRESPONDING ANIMAL AND CLINICAL ENDPOINTS

Session chair Sharon Rosenzweig-Lipson observed that even though there may be corresponding preclinical and clinical endpoints for an aspect of disease, scientists may not know whether that aspect predicts the whole disease or disease reversal. In other words, during discussions of corresponding endpoints, researchers take a step forward in translation, but not necessarily in prediction of therapeutic efficacy. It is important to establish what the corresponding endpoint is intended to predict.

In this session, panelists discussed the role of corresponding endpoints, the choice of endpoints, and bidirectional translation for the study

of nervous system disorders. Prepulse inhibition was offered as an example of the ability to study the same endpoint in an animal model and in human testing. Relatively similar manipulations alter the phenotype in a corresponding manner in both animals and humans. The prepulse inhibition assay has predictive validity for testing the activity of antipsychotics and for developing typical and atypical antipsychotics. Such a corresponding endpoint is valuable in the development of therapies for specific symptoms, in this case therapies acting at specific nodal points within the complex circuitry of schizophrenia.

The experimental autoimmune encephalomyelitis (EAE) model was described as an example of both success and failure within the same drug development program. The model predicted the efficacy of a humanized monoclonal antibody for the treatment of multiple sclerosis but failed to predict a serious adverse event. In a second example, differences in the way the EAE model was used by two different laboratories resulted in completely opposite results regarding the role of tumor necrosis factor (TNF) in multiple sclerosis. Clinical trials were conducted based on the animal study showing control of EAE by inhibition of TNF, but the trial results soon confirmed the animal study showing exacerbation of disease when TNF is blocked.

Another example described how neuroimaging tools have allowed us to understand and exploit the fact that the functional components of hippocampal circuits are very similar in their core functions across animal models and humans.

Every time we use an animal model to make a prediction, Rosenzweig-Lipson said, we need to know the level of understanding of the underlying pathophysiology on which the model is based, the validity of the model, and the level of risk in using the model to make a prediction. For some of the cognition models, for example, there is no "gold standard" model and the risk in making predictions using such a model may be very high. The risk may be lower for models based on a stronger understanding. The key questions are how big is the risk and how good is the prediction. Rosenzweig-Lipson suggested that there should to be honesty in dialogues about predictive value so that later, when there is a failure, it is understood that there was, for example, only a 20 percent certainty that the model was going to make a good prediction. In some cases it is not the models that need to be improved, she said, but the dialogue about the models. One participant noted that the use of multiple models and reduction of relying on a single model might reduce risk through a layered strategy.

THE BASIC AND CLINICAL SCIENCE GAP

Several examples of efforts to bridge the translation gap between preclinical models and clinical trials were discussed, including National Institute of Mental Health–funded consensus-building initiatives (MATRICS and CNTRICS); a for-profit consultancy stepping into the gap between academia and industry to bring validated models to drug developers (P1vital); a company using quantitative systems pharmacology as a translation tool, applying mathematical model-based decision support to drug development (In Silico Biosciences); and a European Union government–facilitated, precompetitive public–private partnership to address specific issues in drug development (NEWMEDS).

Session chair Mark Geyer highlighted several take-away messages from the session. He continued on the theme that discussions about animal models should focus on what is being predicted. This means not attempting to predict Phase III clinical trials outcomes, but more toward predicting Phase IIA results.

As has been done in CNTRICS, Geyer suggested it is important to take the following steps:

- Be clear about what is being measured.
- Determine how to measure it in a human, both in healthy and affected individuals.
- Design tasks that are simple, manageable clinical tools for experimental medicine.
- Design tasks that are also nonverbal and amenable to study in animals.

The structure of NEWMEDS and other IMI initiatives in the European Union are enabling researchers to interact both laterally (across institutions and companies) to share data and ideas and vertically (from preclinical through to clinical). Geyer suggested that the United States can learn from this and improve upon it.

CLOSING REMARKS

Following the session summaries, workshop co-chair Hodes called for final comments and suggestions from participants for going forward.

Participants continued to explore the impact of the terminology used to describe models, and offered suggestions for further discussion.

Model Nomenclature as a Confounder

Despite the broad recognition by many participants that animal models need to be recognized as models of only an aspect of a disease and do not represent an entire disease pathophysiology and phenotype, it was acknowledged that most models are generally referred to as an animal model of a particular disease (e.g., "an animal model of schizophrenia"). Expanding on the discussion in session IV (Chapter 5), it was suggested that equating an animal model of a particular phenotype to a human disease, and discussing the results of a study as a "treatment" or "cure" for the phenotype or disease, can be very misleading. The animal, in fact, never had the full human disease and was not cured. It was also suggested that researchers are misleading each other with these "therapeutic misconceptions." In this context, it was suggested that authors of manuscripts, reviewers, and journal editors should carefully examine the weight assigned to published results.

It was reiterated that clinical trials are conducted for a particular disease and geared toward a drug label indication for treating specific symptoms in the context of the disease. Some participants, however, suggested engaging regulators and others in discussions on this topic during early stages of development, qualification, and validation of biomarkers and endpoints.

One can back-translate from the disease to the animal model, but that does not mean it is an animal model of the disease. In depression, for example, one model used shows an effect on emotional processing, which is a prelude to mood change. This rat model is not about mood; it is about how the animal evaluates rewards and how researchers evaluate emotional stimuli.

SUMMARY

The use of animal models has led to a better understanding of nervous system disorders and diseases and the development of new therapeutics. However, given that there are still few treatment options for many diseases, this workshop sought to concentrate on several important questions: What leads to poor translation of animal models to clinical prac-

tice? Is it the models themselves, research expectations, how models are used to make predictions and decisions, or perhaps the level of knowledge about underlying pathophysiology for any given disease? At each step of the therapeutic development pathway, speakers and participants suggested specific areas for improvement, including the validation of targets, the design of experiments, how results are statistically analyzed, and the way in which positive and negative results are reported. Many participants supported increasing cross-sector collaboration, strengthening training programs and improving the reproducibility of research with a goal of improving translational efforts. Finally, several participants pointed to the need to merge new tools, technologies, and techniques with current research methods, including animal models, as a way to accelerate therapeutic development for nervous system diseases and disorders.

A

References

Arrowsmith, J. 2011a. Trial watch: Phase II failures: 2008-2010. *Nat Rev Drug Discov* 10(5):328-329.

Arrowsmith, J. 2011b. Trial watch: Phase III and submission failures: 2007-2010. *Nat Rev Drug Discov* 10(2):87.

Baker, D., D. Butler, B. J. Scallon, J. K. O'Neill, J. L. Turk, and M. Feldmann. 1994. Control of established experimental allergic encephalomyelitis by inhibition of tumor necrosis factor (TNF) activity within the central nervous system using monoclonal antibodies and TNF receptor-immunoglobulin fusion proteins. *Eur J Immunol* 24(9):2040-2048.

Bakker, A., C. B. Kirwan, M. Miller, and C. E. Stark. 2008. Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science* 319(5870):1640-1642.

Bakker, A., G. L. Krauss, M. S. Albert, C. L. Speck, L. R. Jones, C. E. Stark, M. A. Yassa, S. S. Bassett, A. L. Shelton, and M. Gallagher. 2012. Reduction of hippocampal hyperactivity improves cognition in amnestic mild cognitive impairment. *Neuron* 74(3):467-474.

Berer, K., M. Mues, M. Koutrolos, Z. A. Rasbi, M. Boziki, C. Johner, H. Wekerle, and G. Krishnamoorthy. 2011. Commensal microbiota and myelin autoantigen cooperate to trigger autoimmune demyelination. *Nature* 479(7374):538-541.

Braff, D. L, M. A. Geyer, and N. R. Swerdlow. 2001. Human studies of prepulse inhibition of startle: normal subjects, patient groups, and pharmacological studies. *Psychopharmacology (Berl)* 156(2-3):234-258.

Carter, R. J., L. A. Lione, T. Humby, L. Mangiarini, A. Mahal, G. P. Bates, S. B. Dunnett, and A. J. Morton. 1999. Characterization of progressive motor deficits in mice transgenic for the human Huntington's disease mutation. *J Neurosci* 19(8):3248-3257.

Cissé, M., B. Halabisky, J. Harris, N. Devidze, D. B. Dubal, B. Sun, A. Orr, G. Lotz, D. H. Kim, P. Hamto, K. Ho, G. Q. Yu, and L. Mucke. 2011. Reversing EphB2 depletion rescues cognitive functions in Alzheimer model. *Nature* 469(7328):47-52.

Clozel, M. 2011. Between confidentiality and scientific exchange: the place of publication in drug discovery and pharmaceutical research. *Sci Transl Med* 3(67):67cm2.

Conn, P. J., and B. L. Roth. 2008. Opportunities and challenges of psychiatric drug discovery: Roles for scientists in academic, industry, and government settings. *Neuropsychopharm* 33(9):2048-2060.

Crabbe, J. C., D. Wahlsten, and B. C. Dudek. 1999. Genetics of mouse behavior: Interactions with laboratory environment. *Science* 284(5420):1670-1672.

deIpolyi, A. R., K. P. Rankin, L. Mucke, B. L. Miller, and M. L. Gorno-Tempini. 2007. Spatial cognition and the human navigation network in AD and MCI. *Neurology* 69(10):986-997.

Easterbook, P. J., J. A. Berlin, R. Gopalan, and D. R. Matthews. 1991. Publication bias in clinical research. *Lancet* 337(8746):867-872.

Halliday, G.M. 2001. A review of the neuropathology of schizophrenia. *Clin Exp Pharmacol* 28(1-2):64-65.

IOM (Institute of Medicine). 2010. *Extending the spectrum of precompetitive collaboration in oncology research: Workshop summary*. Washington, DC: The National Academies Press.

Kaur, T., and K. S. Cadenhead. 2010. Treatment implications of the schizophrenia prodrome. *Curr Top Behav Neurosci* 4:97-121.

Kilkenny, C., W. J. Browne, I. C. Cuthill, M. Emerson, and D. G. Altman. 2010. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol* 8(6):e1000412.

Kola, I., and J. Landis. 2004. Can the pharmaceutical industry reduce attrition rates? *Nat Rev Drug Discov* 3(8):711-715.

Kramer, M. A. 1992. Autoassociative neural networks. *Comp Chem Eng* 16(4):313-328.

Kumari, V., W. Soni, and T. Sharma. 1999. Normalization of information processing deficits in schizophrenia with clozapine. *Am J Psychiatry* 156(7):1046-1051.

Landis, S. C., S. G. Amara, K. Asadullah, C. P. Austin, R. Blumenstein, E. W. Bradley, R. G. Crystal, R. B. Darnell, R. J. Ferrante, H. Fillit, R. Finkelstein, M. Fisher, H. E. Gendelman, R. M. Golub, J. L. Goudreau, R. A. Gross, A. K. Gubitz, S. E. Hesterlee, D. W. Howells, J. Huguenard, K. Kelner, W. Koroshetz, D. Krainc, S. E. Lazic, M. S. Levine, M. R. Macleod, J. M. McCall, R. T. Moxley 3rd, K. Narasimhan, L. J. Noble, S. Perrin, J. D. Porter, O. Steward, E. Unger, U. Utz, and S. D. Silberberg. 2012. A call for transparent reporting to optimize the predictive value of preclinical research. *Nature* 490(7419):187-191.

Lenercept (The Lenercept Multiple Sclerosis Study Group and The University of British Columbia MS/MRI Analysis Group). 1999. TNF neutralization in MS: Results of a randomized, placebo-controlled multicenter study. *Neurology* 53(3):457-565.

Liu, J., M. W. Marino, G. Wong, D. Grail, A. Dunn, J. Bettadapura, A. J. Slavin, L. Old, and C. C. Bernard. 1998. TNF is a potent anti-inflammatory cytokine in autoimmune-mediated demyelination. *Nat Med* 4(1):78-83.

Mandillo, S., V. Tucci, S. M. Hölter, H. Meziane, M. A. Banchaabouchi, M. Kallnik, H. V. Lad, P. M. Nolan, A. M. Ouagazzal, E. L. Coghill, K. Gale, E. Golini, S. Jacquot, W. Krezel, A. Parker, F. Riet, I. Schneider, D. Marazziti, J. Auwerx, S. D. Brown, P. Chambon, N. Rosenthal, G. Tocchini-Valentini, and W. Wurst. 2008. Reliability, robustness, and reproducibility in mouse behavioral phenotyping: A cross-laboratory study. *Physiol Genomics* 34(3):243-255.

Markou, A., C. Chiamulera, M. A. Geyer, M. Tricklebank, and T. Steckler. 2009. Removing obstacles in neuroscience drug discovery: The future path for animal models. *Neuropsychopharmacology* 34(1):74-89.

Morgan, P., P. H. Van Der Graaf, J. Arrowsmith, D. E. Feltner, K. S. Drummond, C. D. Wegner, and S. D. Street. 2012. Can the flow of medicines be improved? Fundamental pharmacokinetic and pharmacological principles toward improving Phase II survival. *Drug Discov Today* 17(9-10):419-424.

Müssener, A., L. Klareskog, J. C. Lorentzen, and S. Kleinau. 1995. TNF-alpha dominates cytokine mRNA expression in lymphoid tissues of rats developing collagen- and oil-induced arthritis. *Scand J Immunol* 42(1):128-134.

Nicoll, J. A., E. Barton, D. Boche, J. W. Neal, I. Ferrer, P. Thompson, C. Vlachouli, D. Wilkinson, A. Bayer, D. Games, P. Seubert, D. Schenk, and C. Holmes. 2006. Abeta species removal after abeta42 immunization. *J Neuropathol Exp Neurol* 65(11):1040-1048.

Nieuwenhuis, S., B. U. Forstmann, and E. J. Wagenmakers. 2011. Erroneous analyses of interactions in neuroscience: a problem of significance. *Nat Neurosci* 14(9):1105-1107.

NIH (National Institutes of Health). 2011. *Quantitative and systems pharmacology in the post-genomic era: New approaches to discovering drugs and understanding therapeutic mechanisms*. http://www.nigms.nih.gov/News/Reports/201110-syspharma.htm (accessed July 3, 2011).

Nithianatharajah, J., N. H. Komiyanna, A. McKechanie, M. Johnstone, D. H. Blackwood, D. St. Clair, R. D. Emes, L. N. van de Lagemaat, L. M. Saksida, T. J. Bussey, and S. G. N. Grant. 2013. Synaptic scaffold evolution generated components of vertebrate cognitive complexity. *Nature Neurosci* 16(1):16-24.

Paul, S. M., D. S. Mytelka, C. T. Dunwiddie, C. C. Persinger, B. H. Munos, S. R. Lindborg, and A. L. Schacht. 2010. How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov* 9(3):203-214.

Prinz, F., T. Schlange, and K. Asadullah. 2011. Believe it or not: How much can we rely on published data on potential drug targets? *Nat Rev Drug Discov* 10(9):712.

Putcha, D., M. Brickhouse, K. O'Keefe, C. Sullivan, D. Rentz, G. Marshall, B. Dickerson, and R. Sperling. 2011. Hippocampal hyperactivation associated with cortical thinning in Alzheimer's disease signature regions in non-demented elderly adults. *J Neurosci* 31(48):17680-17688.

Rabchevsky, A. G., J. D. Degos, and P. A. Dreyfus. 1999. Peripheral injections of Freund's adjuvant in mice provoke leakage of serum proteins through the blood-brain barrier without inducing reactive gliosis. *Brain Res* 832(1-2):84-96.

Richter, S. H., J. P. Garner, C. Auer, J. Kunert, and H. Würbel. 2010. Systematic variation improves reproducibility of animal experiments. *Nat Methods* 7(3):167-168.

Rivers, T. M., D. H. Sprunt, and G. P. Berry. 1933. Observations on attempts to produce acute disseminated encephalomyelitis in monkeys. *J Exp Med* 58(1):39-53.

Rogawski, M. A., and H. J. Federoff. 2011. Disclosure of clinical trial results when product development is abandoned. *Sci Transl Med* 3(102):102cm29.

Romberg, C., M. P. Mattson, M. R. Mughal, T. J. Bussey, and L. M. Saksida. 2011. Impaired attention in the 3xTgAD mouse model of Alzheimer's disease: Rescue by donepezil (Aricept). *J Neurosci* 31(9):3500-3507.

Sarter, M., and M. Tricklebank. 2012. Revitalizing psychiatric drug discovery. *Nat Reviews* 11:423-424.

Schenk, D., R. Barbour, W. Dunn, G. Gordon, H. Grajeda, T. Guido, K. Hu, J. Huang, K. Johnson-Wood, K. Khan, D. Kholodenko, M. Lee, Z. Liao, I. Lieberburg, R. Motter, L. Mutter, F. Soriano, G. Shopp, N. Vasquez, C. Vandevert, S. Walker, M. Wogulis, T. Yednock, D. Games, and P. Seubert. 1999. Immunization with amyloid-beta attenuates Alzheimer-disease-like pathology in the PDAPP mouse. *Nature* 400(6740):173-177.

Schulz, K. F., D. G. Altman, and D. Moher, for the CONSORT Group. 2010. CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomized trials. *BMJ* (340):697-702.

Scott, S., J. E. Kranz, J. Cole, J. M. Lincecum, K. Thompson, N. Kelly, A. Bostrom, J. Theodoss, B. M. Al-Nakhala, F. G. Vieira, J. Ramasubbu, and J. A. Heywood. 2008. Design, power, and interpretation of studies in the standard murine model of ALS. *Amyotroph Lateral Scler* 9(1):4-15.

Sena, E. S., H. B. van der Worp, P. M. W. Bath, D. W. Howells, and M. R. Macleod. 2010. Publication bias in reports of animal stroke studies lead to major overstatement of efficacy. *PLoS Biol* 8(3):e1000344.

Sharief, M. K., and R. Hentges. 1991. Association between tumor necrosis factor-alpha and disease progression in patients with multiple sclerosis. *N Engl J Med* 325(7):467-472.

Shineman, D. W., G. S. Basi, J. L. Bizon, C. A. Colton, B. D. Greenberg, B. A. Hollister, J. Lincecum, G. G. Leblanc, L. B. Lee, F. Luo, D. Morgan, I. Morse, L. M. Refolo, D. R. Riddell, K. Scearce-Levie, P. Sweeney, J. Yrjänheikki, and H. M. Fillit. 2011. Accelerating drug discovery for Alzheimer's disease: Best practices for preclinical animal studies. *Alzheimers Res Ther* 3(5):28.

Steinman, L. 2003. Optic neuritis, a new variant of experimental encephalomyelitis, a durable model for all seasons, now in its seventieth year. *J Exp Med* 197(9):1065-1071.

Steinman, L., J. T. Merrill, I. B. McInnes, and M. Peakman. 2012. Optimization of current and future therapy for autoimmune diseases. *Nat Med* 18(1):59-65.

Swerdlow, N. R., and M. A. Geyer. 1993. Clozapine and haloperidol in an animal model of sensorimotor gating deficits in schizophrenia. *Pharmacology Biochemistry and Behavior* 44(3):741-744.

Swerdlow, N. R., J. Paulsen, D. L. Braff, N. Butters, M. A. Geyer, and M. R. Swenson. 1995. Impaired prepulse inhibition of acoustic and tactile startle response in patients with Huntington's disease. *J Neurol Neurosurg Psychiatry* 58(2):192-200.

Swerdlow, N. R., D. L. Braff, and M. A. Geyer. 1999. Cross-species studies of sensorimotor gating of the startle reflex. *Ann NY Acad Sci* 87:202-216.

Swerdlow, N. R., J. M. Shoemaker, S. Crain, J. Goins, K. Onozuka, and P. P. Auerbach. 2004. Sensitivity to drug effects on prepulse inhibition in inbred and outbred rat strains. *Pharmacol Biochem Behav* 77(2):291-302.

Swerdlow, N. R., M. Weber, Y. Qu, G. A. Light, and D. L. Braff. 2008. Realistic expectations of prepulse inhibition in translational models for schizophrenia research. *Psychopharm* 199(3):331-388.

Swinney, D. C., and J. Anthony. 2011. How were new medicines discovered? *Nat Rev Drug Discov* 10(7):507-519.

ter Riet, G., D. A. Korevaar, M. Leenaars, P. J. Stark, G. J. F. Van Noorden, L. M. Bouter, R. Lutter, R. P. Oude Elferink, and L. Hooft. 2012. Publication bias in laboratory animal research: A survey on magnitude, drivers, consequence and potential solutions. *PLoS ONE* 7(9):e43404.

Uher, R., K. E. Tansey, K. Malki, and R. H. Perlis. 2012. Biomarkers predicting treatment outcome in depression: What is clinically significant? *Pharmacogenomics* 13(2):233-240.

van der Worp, H. B., D. W. Howells, E. S. Sena, M. J. Porritt, S. Rewell, V. O'Collins, and M. R. Macleod. 2010. Can animal models of disease reliably inform human studies? *PLoS Med* 7(3):e1000245.

van Oosten, B. W., F. Barkhof, L. Truyen, J. B. Boringa, F. W. Bertelsmann, B. M. von Blomberg, J. N. Woody, H. P. Hartung, and C. H. Polman. 1996. Increased MRI activity and immune activation in two multiple sclerosis patients treated with the monoclonal anti-tumor necrosis factor antibody cA2. *Neurology* 47(6):1531-1534.

Verret, L., E. O. Mann, G. B. Hang, A. M. Barth, I. Cobos, K. Ho, N. Devidze, E. Masliah, A. C. Kreitzer, I. Mody, L. Mucke, and J. J. Palop. 2012. Inhibitory interneuron deficit links altered network activity and cognitive dysfunction in Alzheimer model. *Cell* 149(3):708-721.

Wang, S. J., H. M. Hung, and R. O'Neill. 2011. Adaptive design clinical trials and trial logistics models in CNS drug development. *Eur Neuropsychopharmacol* 21(2):159-166.

Wilson, I., S. Ikonen, R. McMahan, M. Gallagher, H. Eichenbaum, and H. Tamila. 2003. Place cell rigidity is associated with poor spatial learning in aged rats. *Neurobio Aging* 24:297-305.

Yassa, M. A., S. M. Stark, A. Bakker, M. S. Albert, M. Gallagher, and C. E. Stark. 2010. High-resolution structural and functional MRI of hippocampal CA3 and dentate gyrus in patients with amnestic Mild Cognitive Impairment. *Neuroimage* 51(3):1242-1252.

B

Workshop Agenda

**Improving Translation of Animal Models for Nervous
System Disorders:
A Workshop**

March 28-29, 2012

**Institute of Medicine
Keck Building, Room 100
Washington, DC**

Background: Nervous system disorders and diseases are highly prevalent and substantially contribute to the national disease burden. Animal models have significantly increased our understanding of nervous system disorders. Yet, in spite of these advances, there still remains a large gap in treatment options that are high in efficacy but low in side effects for many diseases. And for some diseases there are no treatment options. More than 80 percent of research projects fail to reach clinical trials. Of those nervous system drugs that do make it to clinical trials, only 8 percent end up being approved. These statistics translate to drug approval rates that are 50 percent lower than drugs for other therapeutic areas. Given the tremendous disease burden associated with nervous system diseases and disorders, the goal of this workshop is to bring together key stakeholders to discuss potential opportunities for maximizing the translation of effective therapies from animal models to clinical practice.

83

Meeting Objectives:

- Discuss key issues that contribute to poor translation of animal models in nervous system disorders.
    - o Examine case studies that highlight successes and failures in the development and application of animal models.
- Consider strategies to increase the scientific rigor of preclinical efficacy testing.
    - o Explore the benefits and challenges to developing standardized animal and behavioral models.
    - o Identify methods to facilitate development of corresponding animal and clinical endpoints.
- Identify methods that would maximize bidirectional translation between basic and clinical research.
- Determine the next steps that will be critical for improvement of the development and testing of animal models of disorders of the nervous system.

DAY ONE

1:30 p.m.   Opening Remarks

            RICHARD HODES, *Co-Chair*
            STEVEN PAUL, *Co-Chair*

SESSION I: EVALUATION OF CURRENT ANIMAL MODELS

Session Objective: Identify critical limitations impacting translation of therapies from animal models to clinical practice. Explore current expectations of animal models to predict therapeutic efficacy. Determine the impact of generalization of animal model capabilities. Examine the role of animal model–derived data in making decisions about moving therapeutics into clinical trials.

1:40 p.m.   Overview and Session Objectives

            STEVIN ZORN, *Session Chair*

1:45 p.m.	Examination of Current Expectations for Animal Models

> STEVEN PAUL
> Director
> Helen and Robert Appel Alzheimer's Disease
>   Research Institute
> Weill Cornell Medical College

2:00 p.m.	Choice and Validation of Animal Models for CNS Drug Discovery

> MARK TRICKLEBANK
> Director and Senior Research Fellow
> Eli Lilly and Co.

2:15 p.m.	Impact of Publication Bias

> KATRINA KELNER
> Editor
> *Science Translational Medicine*

2:30 p.m.	Q&A with Speakers

SESSION II: CASE STUDIES

Breakout Objective: Conduct in-depth analysis of six case studies in which animal models have ranged in translational success. Specifically, breakout groups will focus on three key questions: (1) Would this research area benefit from a new or improved standardized animal or behavioral models (e.g., testing conditions, reference standards)? (2) Do animal and human endpoints match for this case study? (3) What is needed to bridge the gap between animal models and clinical science?

3:00 p.m.	Overview and Session Objectives

> RICHARD HODES and STEVEN PAUL
> *Session Chairs*

3:10 p.m.	BREAK OUT INTO GROUPS

3:25 p.m.   Breakout 1: Animal Models for Neurodegeneration

> ROBERT FERRANTE, *Moderator*
> Professor
> Departments of Neurological Surgery, Neurology,
>   and Neurobiology
> University of Pittsburgh
>
> TIM COETZEE, *Discussant*
> Chief Research Officer
> National Multiple Sclerosis Society

Breakout 2: Animal Models for Alzheimer's Disease

> BRADLEY HYMAN, *Moderator*
> John B. Penny Jr. Professor of Neurology
> Harvard Medical School
>
> RICHARD HODES, *Discussant*
> Director
> National Institute on Aging
>
> SHARON ROSENZWEIG-LIPSON, *Discussant*
> IVS Pharma Consulting

Breakout 3: Animal Models for Stroke

> CONSTANTINO IADECOLA, *Moderator*
> George C. Cotzias Distinguished Professor of
>   Neurology and Neuroscience
> Weill Cornell Medical College
>
> WALTER KOROSHETZ, *Discussant*
> Deputy Director
> National Institute for Neurological Disorders and
>   Stroke

STEVEN PAUL, *Discussant*
Director
Helen and Robert Appel Alzheimer's Disease
    Research Institute
Weill Cornell Medical College

Breakout 4: Animal Models for Schizophrenia:
                Cognition Enhancement and Antipsychotic
                Efficacy

HOLLY MOORE, *Moderator*
Associate Professor
Clinical Neurobiology in Psychiatry
Columbia University

MARK GEYER, *Discussant*
Professor
Department of Psychiatry
University of California, San Diego

STEVIN ZORN, *Discussant*
Executive Vice President
Neuroscience Research
Lundbeck USA

Breakout 5: Animal Models for Addiction

ATHINA MARKOU, *Moderator*
Professor
Department of Psychiatry
University of California, San Diego

ALAN LESHNER, *Discussant*
Chief Executive Officer
American Association for the Advancement of
    Science

GERARD MAREK, *Discussant*
Project Director
Neuroscience Development
Abbott Laboratories

Breakout 6: Animal Models for Pain

> A. VANIA APKARIAN, *Moderator*
> Professor
> Neuroscience Institute
> Northwestern University

> DAVID SHURTLEFF, *Discussant*
> Acting Deputy Director
> National Institute on Drug Abuse

4:15 p.m.   Breakout Groups Report Out and Panel Discussion with Participants (15 minutes per group)

- What common themes were identified in:
  - standardization needs
  - endpoints
  - basic science/clinical research gap

6:00 p.m.   ADJOURN

DAY TWO

Note: Continental breakfast will be available at 8:00 a.m.

SESSION III: THE VALUE OF STANDARDIZATION

Session Objective: Explore key components of animal model science that would benefit from standardization, such as behavioral paradigms. Examine the benefits and challenges to developing standardizations for animal and behavioral models. Discuss potential methods for dissemination of standards.

8:30 a.m.   Overview and Session Objectives

> WALTER KOROSHETZ, *Session Chair*

8:40 a.m.   Standardization in Preclinical Models of Anxiety:
Necessary But Not Sufficient

> ANDREW HOMES
> Chief
> Laboratory of Behavioral and Genomic
> Neuroscience
> National Institute on Alcohol Abuse and Alcoholism

8:55 a.m.   Developing New Methods for Cognitive Translation
from Rodent to Human

> TIM BUSSEY
> Professor
> Department of Experimental Psychology
> University of Cambridge

9:10 a.m.   AD Models and the Risk/Benefit Ratio of
Standardization

> LENNART MUCKE
> Director, Gladstone Institute of Neurological
>   Disease
> Professor, Department of Neurology
> University of California, San Francisco

9:25 a.m.   Discussion with Speakers and Participants

10:00 a.m.   BREAK

SESSION IV: CORRESPONDING ANIMAL AND CLINICAL ENDPOINTS

Session Objective: Discuss methods to facilitate development of equivalent or surrogate animal research and clinical trial endpoints. Explore the value of surrogate endpoints for nervous system disorders. Identify components that require recapitulation in both animal studies and clinical trials.

10:15 a.m.   Overview and Session Objectives

               SHARON ROSENZWEIG-LIPSON, *Session Chair*

10:30 a.m.   Prepulse Inhibition: Corresponding Endpoints, But to
               What End?

                    NEAL SWERDLOW
                    Professor
                    Department of Psychiatry
                    University of California, San Diego

10:45 a.m.   Choice of Endpoints—EAE to Approved Drug—One
               Huge Success; One Massive Failure

                    LARRY STEINMAN
                    Professor
                    Department of Neurology and Neurological Sciences
                    Stanford University

11:00 a.m.   Improving Bidirectional Translation for Nervous System
               Disorders

                    MICHELA GALLAGHER
                    Krieger-Eisenhower Professor of Psychology and
                         Neuroscience
                    Department of Psychological and Brain Sciences
                    Johns Hopkins University

11:15 a.m.   Discussion with Speakers and Participants

11:45 a.m.   LUNCH (*will be provided for all participants*)

SESSION V: THE BASIC AND CLINICAL SCIENCE GAP

Session Objective: Explore methods for increasing bidirectional
application of research findings between basic and clinical researchers.
Examine regulatory requirements that may either facilitate or impede
translation of animal models. Identify methods to increase confidence in

the movement from animal models to clinical trials, including replication of studies.

1:00 p.m.	Overview and Session Objectives

> MARK GEYER, *Session Chair*

1:10 p.m.	Developing Better Animal Models of Etiology and Pathophysiology

> RICHARD RANSOHOFF
> Director
> Neuroinflammation Research Center
> Cleveland Clinic

1:30 p.m.	Panel Discussion on the Session Topic (15 minutes/speaker)

> DEANNA BARCH
> Professor of Psychology, Psychiatry and Radiology
> Washington University
>
> GERRY DAWSON
> Chief Science Officer
> P1vital
>
> HUGO GEERTS
> Scientific Liaison Officer
> In Silico Biosciences
>
> THOMAS STECKLER
> Senior Scientific Director
> Neuroscience Drug Discovery
> Johnson & Johnson

2:30 p.m.	Discussion with Speakers and Participants
3:00 p.m.	BREAK

SESSION VI: FUTURE DIRECTIONS AND NEXT STEPS

<u>Session Objective:</u> Define important, yet practical, expectations for animal models in nervous system disorders. Identify opportunities and key stakeholders necessary for the success of improving translation of animal models. Identify key components of the infrastructure support that will be required for implementation.

3:15 p.m.   Overview and Session Objectives

        RICHARD HODES and STEVEN PAUL
        Session Chairs

3:25 p.m.   Session Synopsis and Next Steps

        STEVIN ZORN, *Session I Chair*

        RICHARD HODES, *Session II Co-Chair*
        STEVEN PAUL, *Session II Co-Chair*

        WALTER KOROSHETZ, *Session III Chair*

        SHARON ROSENZWEIG-LIPSON, *Session IV Chair*

        MARK GEYER, *Session V Chair*

4:25 p.m.   Discussion with Speakers and Participants

5:00 p.m.   ADJOURN

C

Registered Attendees

Jane Acri
National Institute on Drug
  Abuse

Neeraj Agarwal
National Eye Institute

Alex Bailey
Food and Drug Administration

Ana Basso
Abbott

Sue Bogner
Institute for the Study of Human
  Error, LLC

Nathalie Breysse
Lundbeck Research USA

Lauren Briese
Physicians Committee for
  Responsible Medicine

Richard Brown
Dalhousie University

Neil Buckholtz
National Institute on Aging

Daniel Burch
CeNeRx BioPharma

Jiu-Chiuan Chen
University of Southern
  California

Theresa Chen
Food and Drug Administration

Wen Chen
National Institute on Aging

Timothy Coetzee
National MS Society

Changhai Cui
National Institutes of Health

Hirsch Davis
National Institute on Drug
  Abuse

93

Michael Didriksen
H. Lundbeck

Jamie Driscoll
National Institute of Mental
Health

Subash Duggirala
Centers for Medicare &
Medicaid Services

Emmeline Edwards
National Center for
Complementary and
Alternative Medicine

Samer Eid
Merck Research Laboratories

Danielle Evers
Office of Science and
Technology Policy

Matthew Fell
Merck & Co.

Stephanie Fertig
National Institute of
Neurological Disorders and
Stroke

John Glowa
National Center for
Complementary and
Alternative Medicine

Jim Gnadt
National Institute of
Neurological Disorders and
Stroke

Ivana Grakalic
National Institute on Alcohol
Abuse and Alcoholism

Lindsey Grandison
National Institute on Alcohol
Abuse and Alcoholism

Amelie Gubitz
National Institute of
Neurological Disorders and
Stroke

Swapnil Gupta
Downstate Medical Center

Magali Haas
One Mind for Research

Michael Haas
BioCentury Publications Inc.

Michael Henry
St. Elizabeth's Medical Center

David Hill
Merck & Co.

David Howland
CHDI Foundation

Carol Hubner
National Institute on Drug
Abuse

William Hurst
Sanofi

Lynn Hyde
Merck Research Laboratories

Steve Hyman
The Broad Institute

Thomas Insel
National Institute of Mental
  Health

John Kehne
Translational
  Neuropharmacology
  Consulting

Elena Koustova
National Institutes of Health

Andrea Kudwa
Psychogenics Inc.

Audrey Kusiak
Department of Veterans Affairs

Story Landis
National Institutes of Health

Emer Leahy
Psychogenics Inc.

Wei Liang
Food and Drug Administration

Yu Lin
National Institutes of Health

Roger Little
National Institute of Mental
  Health

Qi-Ying Liu
National Institute on Alcohol
  Abuse and Alcoholism

Minda Lynch
National Institute on Drug
  Abuse

David McCann
National Institute on Drug
  Abuse

Ryan Merkley
Physicians Committee for
  Responsible Medicine

Eric Mohler
Abbott

Johnathan Moreno
University of Pennsylvania

Jill Morris
National Institute of
  Neurological Disorders and
  Stroke

Richard Morris
National Institutes of Health

Mahadev Murthy
National Institute on Aging

Richard Nakamura
Center for Scientific Review,
  National Institutes of Health

Mark Nedelman
Ekam Imaging

Antonia Noronha
National Institute on Alcohol
  Abuse and Alcoholism

Samia Noursi
National Institute on Drug
  Abuse

Kathie Olsen
ScienceWorks

Deborah Olster
National Institutes of Health

Alexander Ommaya
Department of Veterans Affairs

Javier Perez
Virginia Institute for Psychiatric
  and Behavioral Genetics

Suzana Petanceska
National Institute on Aging

Jonathan Pollack
National Institute on Drug
  Abuse

John Porter
National Institute of
  Neurological Disorders and
  Stroke

Sylvie Ramboz
Psychogenics Inc.

Lorenzo Refolo
National Institute on Aging

Soundar Regunathan
National Institute on Alcohol
  Abuse and Alcoholism

Robert Riddle
National Institute of
  Neurological Disorders and
  Stroke

Becky Robinson
Food and Drug Administration

Joni Rutter
National Institute on Drug
  Abuse

Paul Schnur
National Institute on Drug
  Abuse

Guy Seabrook
Janssen Pharmaceuticals

Mercedes Seraban
Food and Drug Administration

Ming Shah
National Institute on Drug
  Abuse

Mariela Shirley
National Institute on Alcohol
  Abuse and Alcoholism

David Shurtleff
National Institute on Drug
  Abuse

Beth-Anne Siebe
National Institute of
  Neurological Disorders and
  Stroke

Phil Skolnick
National Institute on Drug
  Abuse

D. Stephen Snyder
National Institute on Aging

Erica Spotts
National Institute on Aging

Michael Steinmetz
National Eye Institute

Catherine Stoney
National Institutes of Health

Amber Story
National Science Foundation

Margaret Sutherland
National Institute of
  Neurological Disorders and
  Stroke

Anna Taylor
National Institute of
  Neurological Disorders and
  Stroke

David Thomas
National Institute on Drug
  Abuse

Jason Uslaner
Merck & Co.

Susan Volman
National Institute on Drug
  Abuse

Kenneth Warren
National Institute on Alcohol
  Abuse and Alcoholism

David White
National Institutes of Health

Vicky Whittemore
National Institute of
  Neurological Disorders and
  Stroke

Lois Winsky
National Institute of Mental
  Health

Ellen Witt
National Institute on Alcohol
  Abuse and Alcoholism

Da-Yu Wu
National Institutes of Health

Troy Zarcone
National Institute on Alcohol
  Abuse and Alcoholism

Danna Zimmer
Center for Biomolecular
  Therapeutics

Philippe Zitoun
NOTOCORD