Computational Psychiatry and the Challenge of Schizophrenia

John H. Krystal[*,1–4], **John D. Murray**[1], **Adam M. Chekroud**[5,6], **Philip R. Corlett**[1,5], **Genevieve Yang**[2],
Xiao-Jing Wang[7], and **Alan Anticevic**[1]

[1]Department of Psychiatry, Yale University School of Medicine, New Haven, CT; [2]Department of Neuroscience, Yale University School of Medicine, New Haven, CT; [3]Department of Neuroscience, Yale-New Hospital, New Haven, CT; [4]VA National Center for PTSD, VA Connecticut Healthcare System, West Haven, CT; [5]Department of Psychology, Yale University, New Haven, CT; [6]Department of Psychology, Spring Health, New York, NY; [7]Center for Neural Science, NYU, New York, NY

*To whom correspondence should be addressed; Department of Psychiatry, Yale University School of Medicine, 300 George St. Suite 901, New Haven, CT 06520-8055, US; tel: 203-785-6396, fax: 203-785-6196, e-mail: john.krystal@yale.edu

Schizophrenia research is plagued by enormous challenges in integrating and analyzing large datasets and difficulties developing formal theories related to the etiology, pathophysiology, and treatment of this disorder. Computational psychiatry provides a path to enhance analyses of these large and complex datasets and to promote the development and refinement of formal models for features of this disorder. This presentation introduces the reader to the notion of computational psychiatry and describes discovery-oriented and theory-driven applications to schizophrenia involving machine learning, reinforcement learning theory, and biophysically-informed neural circuit models.

Key words: computational psychiatry/computational neuroscience/schizophrenia/delusions/medication selection/working memory/machine learning

Introduction

Computational approaches may assist the field of schizophrenia research in dealing with its own disordered thoughts. Schizophrenia researchers engaged in discovery-oriented research are overwhelmed by the complexity of large datasets and the challenge of "convergence science,"[1] ie, the effort to integrate many types of data in the pursuit of deep insights into etiology, pathophysiology, and treatment. The field also has struggled to develop formal theory that might guide research conducted within a hypothesis-testing framework. The application of recently developed analytic and mathematical modeling approaches to psychiatry, "computational psychiatry,"[2–4] is facilitating both discovery-oriented and hypothesis-based research. This commentary is intended to introduce readers to computational psychiatry and its applications to schizophrenia and thereby raise

awareness and broaden the adoption of these computational approaches in schizophrenia research.

Computational Psychiatry

Discovery-oriented computational psychiatry approaches have their roots in statistics and computer science.[5] This research is exemplified by the application of machine learning techniques, such as random forests, support vector machines, linear discriminant analysis, and k-means clustering. These approaches are most useful when analyzing highly complex datasets, as they enable one to uncover relationships that are not evident from summary or simple statistics and then to make accurate predictions based on these statistics. Examples of this approach are the effort to predict the response to treatments[6] or to cluster patients on the basis of biological or behavioral traits to supplement or supplant current diagnostic groupings.[4]

Another important approach in computational psychiatry is the application of a Bayesian statistical framework to reinforcement learning (RL) theory.[2,4,5] In Bayesian statistics, predictions (ie, statistical models) are updated with each new observation. Each salient new piece of information teaches you something new because it is, to some degree, different from one's prior expectations, in other words there is a mismatch between the new information and the prior model. This difference, known as a "prediction error," causes one to update one's prior model. In a simplified way,[2] this process may be represented using the following equation, $\delta_t = r_t - V_t$, where δ_t is the prediction error, r_t is the salient new information, and V_t is your prior, what you expected to see at time t. The rate of new learning may be represented as $\alpha\delta_t$, where α may be mediated via neuromodulators like dopamine or acetylcholine. These models provide a powerful framework

for studying disturbances in reward learning and concept formation. However, as will be discussed later, "model-based" forms of RL are complemented by "model-free" forms that establish relatively rigidly expressed forms of thought or behavior, such as habits or fixed beliefs.[7]

A second important domain of theory-driven computational psychiatry research emerged as a clinical translation of computational neuroscience. Computational neuroscience emerged with advances in neurophysiology, computing capacity, informatics, and computational modeling.[8] Increasingly, neural simulations emerged that were informed by the specific biophysical properties of actual neural circuits and that represented signaling alterations that were implicated in psychiatric disorders. One early example of this was the modeling of deficits in the maintenance of persisting prefrontal neural activity, thought to underlie working memory deficits in schizophrenia.[2] The emergence of advances in mechanistic psychiatric research employing psychopharmacology, genomics, and molecular brain imaging, and functional neuroimaging created new opportunities to apply neural simulations derived from preclinical models to data from individuals diagnosed with psychiatric disorders. In this way, computational psychiatry could facilitate, for the first time, the development of formal hypotheses regarding the nature of microcircuit dysfunction in psychiatric disorders.[2-4] In so doing, these approaches set the stage for deep understanding of brain-behavior relationships and the development of novel therapeutics.

Three Brief Exemplars of Applications of Computational Psychiatry to Schizophrenia

Increasingly, computational psychiatry approaches[9,10] are applied in both discovery-oriented and theory-oriented studies of schizophrenia. For example, a recent discovery-oriented study applied machine learning to data from a large randomized study, the European First Episode of Schizophrenia Trial (EUFEST).[11] It identified a set of variables that predicted multiple clinical outcomes with over 70% accuracy at 4 weeks and 1 year of treatment.

Computational psychiatry approaches building on temporal difference RL approaches have contributed advances in theory, including a cognitive theory of delusions.[7] In this model, simplified here, aberrant prediction errors in the right lateral prefrontal cortex are generated by a number of factors associated with psychosis, such as heightened salience of irrelevant stimuli as a consequence of increased striatal dopamine release, distortion of information in memory, and other factors.[7] Distortions in the resulting beliefs produce "top-down," ie, cortically driven distortions of perceptual processes that reinforce aberrant expectancies. The cycle of altered perceptual experiences and false inference formation recurrently reinforce each other as a result of the need to attempt to reconcile sensory experiences, memories, and expectancies.

This repetitive process appears to shift the learning of expectancies, ie, the formation of beliefs, from a model-based to model-free form of learning. In other words, the beliefs lose the properties of "regular" beliefs, which are recalled voluntarily and updated flexibly, and acquire the properties of habits, ie, they are fixed and not falsifiable. This aberrant process may be maintained by high levels of dopamine D_2 receptor signaling, as D_2 receptor antagonists can alleviate delusions.

Lastly, a biophysically-informed computational model of cortical microcircuits sheds light on disturbances in microcircuits that might arise in schizophrenia as a consequences of deficits in glutamate signaling, particular signaling mediated by NMDA glutamate receptors (NMDA-R). In these models, NMDA-Rs contribute to recurrent excitation in layer 3 of the lateral prefrontal cortex that underlies the maintenance of information in working memory.[2] Reductions in NMDA-R signaling undermine sustained activity, compromising working memory-related PFC activity and working memory performance. These findings are consistent with subsequent studies of NMDA-R antagonist effects on working memory-related PFC activity in humans and reduced working memory-related PFC activity associated with schizophrenia. However, these models also suggest that reduced excitatory drive of interneurons would impair working memory in schizophrenia by undermining the suppression of "noisy" neural activity at rest and during working memory. Loss of spatial tuning also reduces the precision of memory.[12] Again, these disturbances in cortical activity could be demonstrated in healthy subjects administered ketamine and in schizophrenia. Further, this model predicted that in patients who have reduced working memory precision, a metabotropic-2 glutamate receptor agonist (mGluR2) would improve memory function by restoring the balance of excitation and inhibition in disinhibited networks. This hypothesis has yet to be tested. However, mGluR2 agonists do attenuate ketamine-related impairment in working memory in humans[13] and may have therapeutic effects in subgroups of schizophrenia patients.[14]

Toward Computational Schizophrenia Research

Computational psychiatry is not a panacea, but it constitutes both a conceptual and a practical advance. Advances in this area highlight the limitations of prior analytic approaches and missed opportunities to develop formal theories related to schizophrenia etiology, pathophysiology, and treatment. In order to achieve the opportunities created by computational psychiatry, clinical investigators may need to establish new collaborations with computational neuroscientists, statisticians, computer scientists, as well as scientists and engineers in related fields. This breaking down of the intellectual silos that separate fields should benefit everyone. Further, there is

an opportunity to train a new generation of translational and clinical investigators in the emerging computational methods. For them, the field of computational psychiatry will be a foregone conclusion.

References

1. Sharp PA, Langer R. Research agenda. Promoting convergence in biomedical science. *Science*. 2011;333:527.
2. Wang XJ, Krystal JH. Computational psychiatry. *Neuron*. 2014;84:638–654.
3. Montague PR, Dolan RJ, Friston KJ, Dayan P. Computational psychiatry. *Trends Cogn Sci*. 2012;16:72–80.
4. Redish AD, Gordon JA. *Computational Psychiatry: New Perspectives on Mental Illness*. Cambridge, MA: The MIT Press; 2016.
5. Huys QJ, Maia TV, Frank MJ. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat Neurosci*. 2016;19:404–413.
6. Chekroud AM, Zotti RJ, Shehzad Z, et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. *Lancet Psychiatry*. 2016;3:243–250.
7. Corlett PR, Taylor JR, Wang XJ, Fletcher PC, Krystal JH. Toward a neurobiology of delusions. *Prog Neurobiol*. 2010;92:345–369.
8. Sejnowski TJ, Koch C, Churchland PS. Computational neuroscience. *Science*. 1988;241:1299–1306.
9. Deserno L, Heinz A, Schlagenhauf F. Computational approaches to schizophrenia: a perspective on negative symptoms [published online ahead of print December 13, 2016]. *Schizophr Res*.
10. Dowd EC, Frank MJ, Collins A, Gold JM, Barch DM. Probabilistic reinforcement learning in patients with schizophrenia: relationships to anhedonia and avolition. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2016;1:460–473.
11. Koutsouleris N, Kahn RS, Chekroud AM, et al. Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. *Lancet Psychiatry*. 2016;3:935–946.
12. Murray JD, Anticevic A, Gancsos M, et al. Linking microcircuit dysfunction to cognitive impairment: effects of disinhibition associated with schizophrenia in a cortical working memory model. *Cereb Cortex*. 2014;24:859–872.
13. Krystal JH, Abi-Saab W, Perry E, et al. Preliminary evidence of attenuation of the disruptive effects of the NMDA glutamate receptor antagonist, ketamine, on working memory by pretreatment with the group II metabotropic glutamate receptor agonist, LY354740, in healthy human subjects. *Psychopharmacology (Berl)*. 2005;179:303–309.
14. Kinon BJ, Millen BA, Zhang L, McKinzie DL. Exploratory analysis for a targeted patient population responsive to the metabotropic glutamate 2/3 receptor agonist pomaglumetad methionil in schizophrenia [published online ahead of print March 19, 2015]. *Biol Psychiatry*.