Science Bulletin 68 (2023) 1606–1610

Contents lists available at ScienceDirect

Science Bulletin

journal homepage: www.elsevier.com/locate/scib

Perspective

Bridging the gap between artificial intelligence and mental health

Tangsheng Lu [a,1], Xiaoxing Liu [b,1], Jie Sun [c], Yanping Bao [a], Björn W. Schuller [d], Ying Han [a,*], Lin Lu [a,b,e,*]

[a] National Institute on Drug Dependence and Beijing Key Laboratory of Drug Dependence Research, Peking University, Beijing 100191, China
[b] Peking University Sixth Hospital, Peking University Institute of Mental Health, NHC Key Laboratory of Mental Health (Peking University), National Clinical Research Center for Mental Disorders (Peking University Sixth Hospital), Chinese Academy of Medical Sciences Research Unit, Peking University, Beijing 100191, China
[c] Pain Medicine Center, Peking University Third Hospital, Beijing 100191, China
[d] GLAM-Group on Language, Audio & Music, Imperial College London, London SW7 2AZ, UK
[e] Peking-Tsinghua Center for Life Sciences and PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing 100871, China

Mental disorder is one of the greatest health challenges of our time, attracting increasingly more public awareness and causing high disability and huge socioeconomic burdens. According to the 2019 Global Burden of Disease led by the Institute for Health Metrics and Evaluation, 12.5% of the global population has been suffering from a mental disorder. Globally, mental illness contributes to 32.4% of years lived with disability and 13.0% of disability-adjusted life-years [1]. However, most individuals suffering from severe mental health conditions actually do not receive mental health services [2]. This can be attributed to various challenges, including the stigma of mental disorders, a scarcity of mental health resources, and the subjective nature of diagnosing mental disorders. In recent decades, with the advancement of machine learning (ML), natural language processing (NLP), and computer vision, researchers have gained the opportunity to develop artificial intelligence (AI) tools for clinical fields to alleviate the shortage of medical resources. In addressing these challenges, AI offers distinct advantages. The development of AI technologies enables the implementation of digital healthcare, which can help alleviate the strain on healthcare resources. Moreover, AI has the potential to accurately identify mental disorders, overcoming the issue of social stigma associated with mental health questionnaires. Additionally, data-driven AI can provide more objective evidence for the diagnosis and treatment of mental disorders. However, effective AI tools in the field of mental disorders are still relatively scarce. This phenomenon partially results from the practical challenges in translating AI into clinical practice (also known as the AI chasm), which is caused by the limited transparency, suitability, and adaptability of AI [3]. Another contributing factor to the AI chasm is the ambiguous definition of mental disorders per se (Fig. 1).

Currently, the diagnosis of mental disorders is based on symptom-based classification, requiring mental health professionals to face-to-face observe the states of patients in clinical practice. Skilled clinicians do not primarily rely on abstract psychological concepts for basic diagnosis, but must also understand the complex psychological characteristics of their patients [4]. Thus, mental clinicians depend more on their tacit knowledge which has been accumulated through long-term clinical experience. However, this kind of knowledge exhibits subjectivity and informality, which exists in specific individuals and environments. Specifically, the difficulty and inherent arbitrariness of establishing a diagnostic threshold for mental disorders are widely acknowledged. The complex etiology of mental disorders remains unclear, and the diagnosis and assessment of these disorders are influenced by cultural, social, and individual differences. However, AI-powered tools have the potential to overcome this issue. The utilization of AI and ML algorithms to analyze electrophysiological data and functional magnetic resonance imaging data, is thereby offering a definitive quantitative "gold standard" diagnostic test. This presents a notable challenge for AI tools, as it requires the development of robust algorithms that can accurately interpret complex and heterogeneous data. Besides, large language models (LLMs) have demonstrated impressive performance in NLP tasks. A review indicates a growing emphasis on NLP research for the detection of mental disorders [5]. Moreover, the application of advanced LLMs like GPT-4 and Bard could greatly enhance the processing and interpretation of electronic health records (EHRs). Furthermore, employing general-purpose chatbots for sentiment analysis enables the detection of user's emotions during conversations. Therefore, crossing the AI chasm could bring novel insight for the treatment of mental disorders. AI-based tools, if properly developed and implemented, could provide clinicians with valuable support for diagnosing and treating mental disorders, potentially leading to better patient outcomes.

Developing early detection, diagnosis, treatment, and redefining or subtyping of mental disorders using AI technology based on brain imaging data, EHRs, and novel monitoring systems is a challenging endeavor. The unclear pathogenesis of mental disorders has posed significant challenges for clinical diagnosis, while AI has already surpassed human performance in some domains such as image recognition tasks, indicating its immense potential to analyze brain imaging data through computer vision and ML [6]. Sufficient evidence suggests that using brain age estimation methods to extract features from brain imaging data allows for the accurate diagnosis of the brain's health status [7]. Automated

* Corresponding authors.
E-mail addresses: yinghan@bjmu.edu.cn (Y. Han), linlu@bjmu.edu.cn (L. Lu).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.scib.2023.07.015

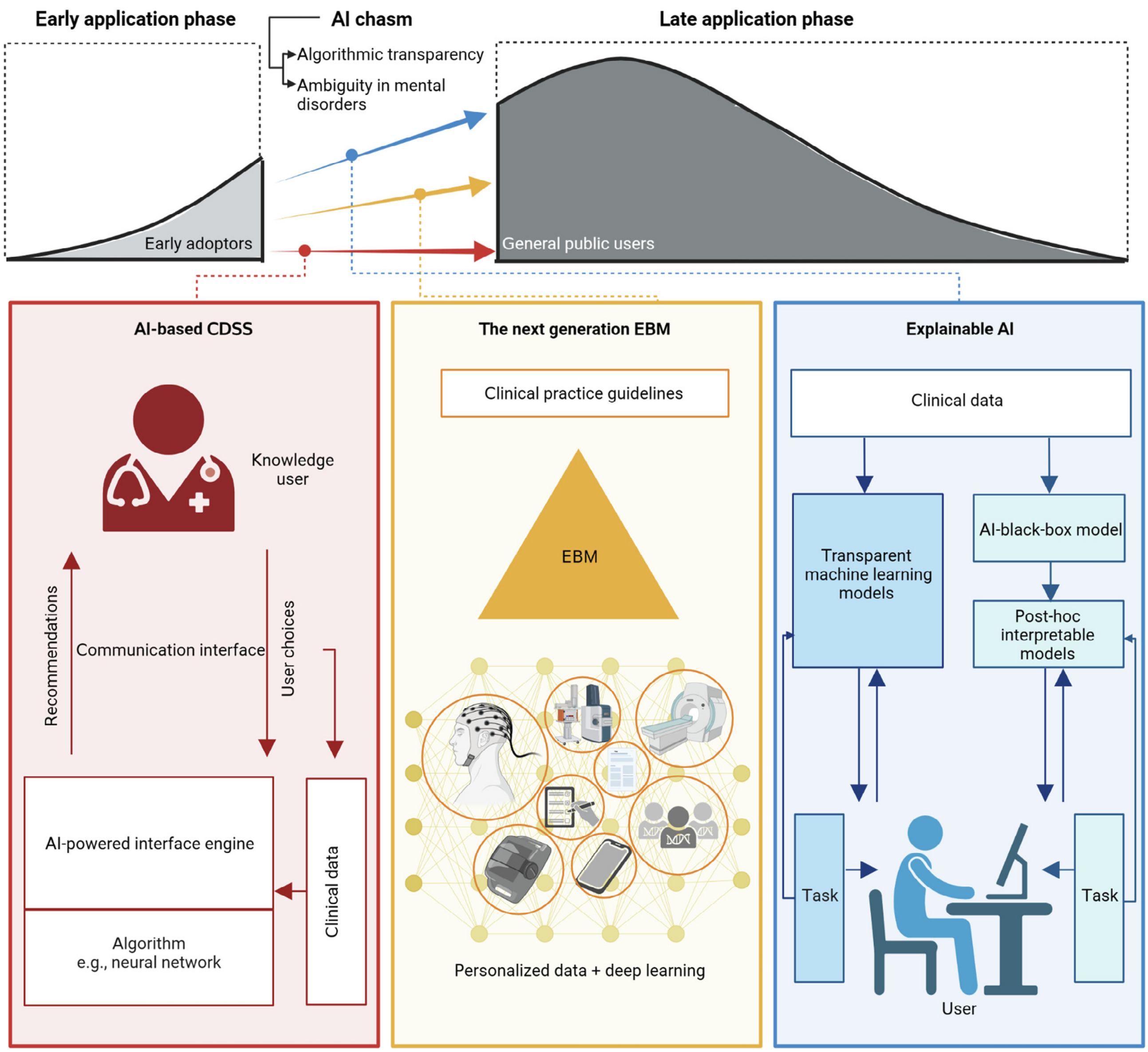


Fig. 1. Overview of three strategies for crossing the AI chasm in mental health. The AI chasm refers to the obstacles encountered during the development of artificial intelligence from early clinical research to its widespread clinical application. In mental health, the obstacles of the AI chasm not only involve the characteristics of AI technology itself but also the ambiguity in mental disorders. The first strategy involves the use of artificial intelligence-based clinical decision support system (AI-CDSS) to assist clinicians in decision-making, improving accuracy and efficiency. The second strategy is the integration and merging of all available data to achieve the next generation of evidence-based medicine (EBM). The third strategy focuses on the development of explainable AI, with two main directions of the construction of transparent machine learning models and interpretable models. (Created with BioRender.com).

computer-aided diagnosis, which includes brain imaging data pre-processing, feature extraction, and classification, provides a straightforward and concrete method for identifying brain disorders [8]. Therefore, the application of AI technology has revolutionized the processing and analysis of a large amount of brain signal data, making it possible to diagnose mental disorders early. Clinical trials have demonstrated that the US Food and Drug Administration (FDA)-approved AI agent, Wysa, can effectively relieve chronic pain and associated depression and anxiety [9]. This indicates the potential of AI in mental treatment. ML algorithms have also shown better results in disease classification. An 18-month longitudinal study used data-driven approaches and high-dimensional clinical data to expand the classification of mental disorders beyond symptom-based categories to include stratification based on functional outcomes, genetic markers, and trajectory [10]. In summary, AI has tremendous potential in the diagnosis, treatment, and classification of mental disorders, and can overcome bottlenecks in traditional approaches. AI may bring a qualitative leap forward to the field of mental disorders.

While AI has great potential in the clinical application of mental disorders, it is clear that the current strategies are far from optimal, as the limited clinical usage suggests. As of October 2022, the US FDA has approved 521 AI/ML-enabled medical devices, with the highest application of such devices in the field of radiology. However, only 14 AI/ML-enabled medical devices are related to neurology, including two for the early identification of autism spectrum disorder in children, three for assessing mild traumatic brain injuries, and two for sleep scoring and analysis. From searching on ClinicalTrials.gov in April 2023, a total of 1296 AI-related clinical research has been found, and only 102 studies were identified using "mental" and "AI" as keywords. However, few studies focus on mental disorders, such as only 25 research on anxiety and/or depression, 19 on dementia or Alzheimer's disease, 12 on sleep disorders, and 6 on opioid use disorders. Of these, a similar pattern is evident in the 60 projects funded by Chinese universities, hospitals, or research groups, given that only 2 of them applied for mental health as a subject, aiming at insomnia or major depressive disorder. This situation suggests that the use of AI in clinical practice is still restricted, and that it has yet to bring about the digital health revolution that many experts have predicted.

Clinical practice guidelines (CPGs) are based on evidence-based medicine (EBM), which aimed to integrate the expertise of clinicians, patient preference, and the best available scientific research to guide clinical management decisions. The concept of EBM has had a profound impact on various fields of healthcare over the past few decades, including medical education, clinical decision-making, and healthcare policy. The gold standard of EBM is to control for confounding factors in order to eliminate bias and infer causal relationships. However, AI technology, which converts complex data such as EHRs, patient-reported symptoms and outcomes, examinations, and treatment, into algorithms, challenges the traditional epidemiologic approach of EBM. The two paradigms of EBM and data-driven AI approaches have completely different disciplinary logics regarding "evidence" and "results". This indicates incommensurability between the two, and it is impossible to achieve a synthesis between two incommensurable paradigms on an epistemological level [11]. Based on the extent of AI's use in the field of mental illness, we propose three strategies for the current development of AI applications in mental disorders, with clinical medicine as the main axis and AI as an auxiliary tool, to address the incommensurability between EBM and AI (Fig. 1). These strategies aim to bridge the AI chasm and enable the seamless integration of AI technology into clinical practice.

AI-based clinical decision support system. AI-based clinical decision support systems (AI-CDSS) utilize biomedical data to analyze the probability of medical outcomes, aiming to enhance diagnosis, treatment, and prognosis. Understanding the determinants of AI-CDSS is crucial for guiding their development and implementation, as the factors influencing their utilization vary across different innovations. The clinical decision-making is the main content of clinical activities, heavily relying on CPGs. However, the most commonly used guidelines for the diagnosis of mental disorders are based on patients' self-reported symptoms, clinicians' observed behavioral patterns, and the process and related features of symptoms, leading to the subjectivity of mental disorders' diagnosis. To address this limitation, AI-based analysis of multimodal medical imaging, laboratory values, clinical notes, and omics data can increase the accuracy of diagnosis, treatment, and prognosis [12]. Consequently, an AI-CDSS can effectively detect and diagnose various mental disorders, unlike traditional knowledge-based CDSS. AI-CDSS provides real-time information and suggestions to clinicians based on established clinical guidelines and a large dataset. However, for AI-CDSS to gain trust and confidence of clinicians, its development and implementation must be carried out with care. While some AI-CDSS have shown exceptional performance in tests, it is crucial to consider potential risks associated with the model and algorithm when deploying them in practical and commercial applications. A literature review suggests that AI-CDSS tools based on AI or ML can be used in mental health systems, but their development is still in its early stages [13]. Successful implementation of AI-CDSS in mental healthcare requires identifying problems, selecting appropriate ML approaches, determining data and format requirements, gathering feedback, and validating the tool. Altogether, in the future, mental health professionals can use AI-CDSS and its associated evaluation tools to improve clinical decision-making and diagnostic accuracy, but they must remain mindful of potential risks associated with the technology.

The next generation of EBM. The next generation of EBM is an approach that surpasses the traditional hierarchy of evidence and leverages AI to play a role in data collection, analysis, and comprehensive synthesis. The principle of EBM is to reduce the reliance on unstructured clinical experience and to give greater weight to clinical research evidence. However, there are various challenges in enrolling and retaining study participants, such as a restricted number of clinical centers, inadequate patient recruitment strategies, limited patient access, and suboptimal trial design. These issues, coupled with the high costs of clinical trials, highlight the inefficiencies of the trial process and result in a crisis in clinical research. To address these challenges, the pharmaceutical industry is exploring AI-based solutions to optimize clinical trial processes and reduce expenses. For instance, home wearable devices and other digital technologies can be used to revolutionize trials by replacing structured, time-point, and in-clinic-specific assessments with continuous, objective measures. These devices can measure patient outcomes such as physical activity or quality of sleep conveniently. In addition, remote patient monitoring and management technologies enable more consistent engagement with trial participants and facilitate the sharing of health status information, allowing for more accurate collection of outcome data in clinical trials. Therefore, the next generation of EBM underpinned by AI will require a deep integration and consolidation of all available data, including audio and image data, EHRs, the medical internet of things, genomics, omics, published clinical research, and more [14]. These personalized data will be integrated into the design, execution, and data processing of clinical trials and may also lead to the revision of CPGs.

Explainable AI. Explainable AI (XAI) refers to a collection of tools and frameworks designed to assist humans in comprehending and interpreting the predictions generated by machine learning models. Specifically, in the context of AI-based medical decision-making, it is crucial to provide reasonable explanations that clinicians can understand. The issue of the "black box" of AI cannot be avoided, regardless of the role that AI plays in the healthcare system. The healthcare system is a high-risk field, and unexplainable models with risk characteristics are inevitable. This has led to healthcare professionals' cautious attitude and distrust towards AI technology products, increasing the demand for XAI. The appeal from practitioners in the field of mental illness is even stronger, mainly because data describing syndromes, outcomes, disorders, symptoms, and tentative etiologies, as well as the multifactorial social and psychological determinants of disorders, possess probabilistic relationships with each other [15]. Therefore, it is necessary to build XAI that can be understood and appropriately trusted by healthcare professionals. In the future, to enhance the transparency and explainability of AI models in mental health system, two strategies can be pursued. Firstly, the dependence of deep learning models on hyperparameters should be explored, and the underlying workings of deep learning should be understood to develop transparent ML models. Secondly, post-hoc explainability can be achieved through the development of model-agnostic techniques that can extract information from the prediction procedure of any model, thus increasing the transparency and explainability of AI systems. Therefore, the development of interpretable models that can be widely applied in medicine is important. Moreover, the usability of such explanations will need to be guaranteed and personalization of explanations seems highly needed. In fact, such explanations should go beyond mere visualization including also verbal or sonified and multimodal explanations.

The development of AI is highly dependent on big data, which is critical for training and improving decision-making processes. However, the quality of data is of utmost importance to ensure the performance and reliability of AI. Low-quality data can lead to incorrect outcomes and unreliable decisions. In the field of mental health, it is crucial to ensure standardized collection and supervision of various types of clinical data, such as standardized writing of EHRs, as well as the standardized collection and analysis of brain imaging data. Furthermore, it is essential to guarantee the integrity and accuracy of the data. By ensuring high-quality and standardized data, AI algorithms can be trained and validated to provide more reliable and accurate predictions.

The use of AI in clinical practice poses significant moral challenges that require careful consideration. One immediate challenge is obtaining informed consent, especially with the use of black-box algorithms that clinicians may struggle to comprehend. Clinicians need to educate patients about the complexities of AI, including the type of ML used by the system, the data inputs used, and the possibility of biases or other data-related issues. Another concern is the right to explanation, particularly when AI is used in diagnosis and treatment recommendations. Additionally, patient privacy must be balanced with the safety and effectiveness of AI. User agreements should include information on the future use of AI health Apps, which may be conditional on accepting changes to the terms of use, and should resemble informed consent documents. Thus, as the use of AI in mental health grows, it is necessary to establish legal and regulatory frameworks that safeguard privacy rights. In addition to existing ethical policies in the clinical domain, China, the European Union countries, and the US are actively engaged in discussions and initiatives regarding AI ethics, fairness, and privacy.

Although some progress has been made in the development of AI for mental illness, there is still a long way to go before AI can meet the high expectations of clinical practice. In order to achieve this goal, there are three strategies for the current development of AI implement in mental disorders. The first strategy is AI-CDSS, which mainly utilizes advanced statistics and computing to analyze clinical data, identify patterns, and make predictions to identify or suggest gaps, mistakes, safety issues, or care pathway improvements to the user. It may promisingly and potentially assess and diagnose mental disorders in the future. The second strategy is to promote the widespread application of AI in clinical practice, which mainly involves the full participation of AI in the process of clinical guidelines to form the next generation of EBM. Similarly, given the unique characteristics of mental disorders, their treatments may not fully conform to the assumptions of EBM. However, the next generation of EBM is also striving to develop evidence for evidence-based psychiatry. The third strategy aims to address the trust issues of the "black box" problem of AI and solve the problem of algorithmic transparency, providing a smooth path for the clinical application of AI. XAI is crucial in psychiatry due to the presence of probabilistic relationships among the data describing syndromes, outcomes, and disorders. In general, AI has been revolutionizing the field of mental healthcare. AI could serve as a new tool for planning mental health services and identifying and monitoring individual and population mental health issues. AI-driven tools can utilize digitized healthcare data, including EHRs, medical images, and biomedical data, and deepen the understanding of complex disease etiologies. Additionally, the trustworthiness, fairness, efficiency, and dependability data of AI, as well as data safety, laws, and regulations, need to be carefully considered. Furthermore, these three strategies can be synergized to advance the clinical application of AI.

Conflict of interest

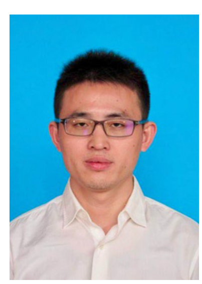
The authors declare that they have no conflict of interest.

Acknowledgments

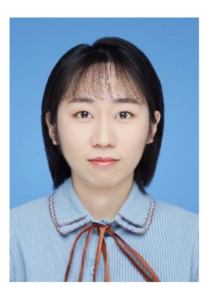
This work was supported by the STI2030-Major Projects (2021ZD0200800).

References

[1] Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. Lancet Psychiatry 2016;3:171–8.
[2] Wang PS, Aguilar-Gaxiola S, Alonso J, et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. Lancet 2007;370:841–50.
[3] Aristidou A, Jena R, Topol EJ. Bridging the chasm between AI and clinical implementation. Lancet 2022;399:620.
[4] Clark LA, Cuthbert B, Lewis-Fernandez R, et al. Three approaches to understanding and classifying mental disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC). Psychol Sci Public Interest 2017;18:72–145.
[5] Le Glaz A, Haralambous Y, Kim-Dufor DH, et al. Machine learning and natural language processing in mental health: systematic review. J Med Internet Res 2021;23:e15708.
[6] Monsour R, Dutta M, Mohamed AZ, et al. Neuroimaging in the era of artificial intelligence: current applications. Fed Pract 2022;39:S14–20.
[7] Mishra S, Beheshti I, Khanna P. A review of neuroimaging-driven brain age estimation for identification of brain disorders and health conditions. IEEE Rev Biomed Eng 2023;16:371–85.
[8] Lima AA, Mridha MF, Das SC, et al. A comprehensive survey on the detection, classification, and challenges of neurological disorders. Biology (Basel) 2022;11:469.
[9] Leo AJ, Schuelke MJ, Hunt DM, et al. Digital mental health intervention plus usual care compared with usual care only and usual care plus in-person psychological counseling for orthopedic patients with symptoms of depression or anxiety: cohort study. JMIR Form Res 2022;6:e36203.
[10] Dwyer DB, Kalman JL, Budde M, et al. An investigation of psychosis subgroups with prognostic validation and exploration of genetic underpinnings: the PsyCourse study. JAMA Psychiatry 2020;77:523–33.
[11] Chen X. Thomas Kuhn's latest notion of incommensurability. J Gen Phil Sci 1997;28:257–73.
[12] Kline A, Wang H, Li Y, et al. Multimodal machine learning in precision health: a scoping review. NPJ Digit Med 2022;5:171.
[13] Golden G, Popescu C, Israel S, et al. Applying artificial intelligence to clinical decision support in mental health: what have we learned? arXiv: 230303511, 2023.
[14] Subbiah V. The next generation of evidence-based medicine. Nat Med 2023;29:49–58.
[15] Joyce DW, Kormilitzin A, Smith KA, et al. Explainable artificial intelligence for mental health through transparency and interpretability for understandability. NPJ Digit Med 2023;6:6.



Tangsheng Lu is a Ph.D. candidate at National Institute on Drug Dependence and Beijing Key Laboratory of Drug Dependence Research, Peking University. He received his master's degree from ShanghaiTech University. His research interest mainly lies in pathological emotional memory.



Xiaoxing Liu now works at the Peking University Sixth Hospital. She received her Ph.D. degree in integrated life science from Peking-Tsinghua Center for Life Sciences and PKU-IDG/McGovern Institute for Brain Research, Peking University. Her research interest mainly lies in the neural circuit mechanism that underlies psychiatric disorders such as substance use disorders.

Ying Han is an assistant professor and principal investigator at National Institute on Drug Dependence and Beijing Key Laboratory on Drug Dependence Research, Peking University. Her research interest mainly focuses on investigating the neural mechanisms that underlie stress-induced mental disorders with the goal of discovering new targets for rapid antidepressant action and novel therapeutic interventions.

Lin Lu is a professor and the director of Peking University Sixth Hospital/Institute of Mental Health. He is also member of the International Narcotics Control Board, the director of the National Psychiatric Medical Center, National Clinical Research Center for Mental Disorders, Mental Health Center of Chinese Center for Disease Control and Prevention, and Peking University Clinical Psychology Center. He has been engaged in research on public health policies, pathogenesis, epidemiological characteristics, and intervention strategies for mental and psychological diseases and sleep disorders.