as measured by electrochemiluminescence assay, and blood cell phenotypic data obtained using the cyTOF platform. Finally, we address current shortcomings in methodology by applying a novel method to estimate specific binding to TSPO.

**Results:** Lower levels of TSPO was demonstrated in antipsychotic-naïve first episode psychosis patients. This unexpected result was confirmed in the individual-participant data meta-analysis. In follow-up analyses in a subgroup of our cohort, the results indicated a normalization of TSPO. Preliminary analyses show positive correlations between chemokines in CSF and central TSPO binding, whereas in controls these variables were negatively correlated. In the cyTOF analysis, where up to 50 different phenotypic markers can be obtained from a single blood cell sample, we observed significant aberrations that were located to the monocyte population. Finally, preliminary results applying the Simultaneous Method, which utilizes time activity-curves from multiple brain regions, show that specific binding estimates were in correspondence to in vivo blocking data, and a test-retest analysis showed good reliability and precision.

**Conclusions:** The TSPO PET data together with immune marker analysis support aberrations in both central and peripheral immune cell function in schizophrenia. Methods for accurately assessing specific binding to TSPO can be of help in increasing sensitivity in clinical studies.

### 14. IS BIGGER BETTER? PROMISES AND PITFALLS OF BIG DATA IN NEUROIMAGING OF PSYCHOSIS

**Ofer Pasternak**  
Harvard Medical School

Big neuroimaging datasets comprising hundreds or even thousands of subjects are becoming widely available, thanks to major collaborative efforts across multiple imaging centers and groups. Mining and analyzing Big-Data is also becoming feasible, owing to increased computational power and new implementations of machine learning algorithms, which can learn from data and generate predictions. Big-Data studies bear exceptional promise in disentangling complex psychiatric illness, including psychosis, where imaging correlates are often subtle and difficult to reproduce. Large datasets, combined with novel machine learning algorithms have opened avenues for delineating subtypes, as well as for predicting biological and clinical outcomes, including psychosis conversion and drug response. However, the use of Big-Data generates challenges in terms of design, construction and statistical analyses. One such problem presented by current studies relates to harmonizing neuroimaging signals across multiple centers or sites, as no current gold standard exists. Moreover, harmonizing disparate populations could introduce unwanted covariables, and potentially attenuate psychosis-related effects. Additional challenge relates to incomplete or incompatible clinical information between different study sites, which diminishes usable clinical measures and in turn, the clinical scope of Big-data studies.

This symposium is designed to address the practical and theoretical aspects of acquiring and analyzing large neuroimaging studies in psychosis and schizophrenia-spectrum disorders. We will describe new and exciting opportunities that come with Big-Data studies, as well as the pitfalls limiting current methods. Throughout, we will provide practical recommendations to maximize the potential of big-data. In addition, we will explore the current tradeoff between Big-data studies that bolster sensitivity and smaller studies, which enable nuanced investigation of homogenous populations and specific imaging markers to elucidate underlying pathologies in psychosis. We argue that both Big and small data play important roles in our effort to understand psychosis and its treatment.

This symposium will bring together five leading neuroimagers, who will present different aspects of Big versus small data studies, while providing critical cautionary remarks regarding the shortcomings of these methods:

1) Prof. Neda Jahanshad, PhD, of the Keck School of Medicine, University of Southern California, is a key player in the ENIGMA network, and its Big-Data that was constructed through meta analyses. She will present key findings from the ENIGMA studies, as well as important technical considerations that arise in the meta-analysis process, including combining clinical information across multiple study sites.

2) Prof. Bo (Cloud) Cao, PhD, of the Department of Psychiatry, University of Alberta, Canada, is an expert in the modification and application of machine learning approaches for imaging studies. He will present state-of-the-art prediction algorithms and describe the advantages of combining these algorithms with Big-Data in psychosis studies.

3) Prof. Jennifer Coughlin, MD, of the Department of Psychiatry and Behavioral Sciences, Johns Hopkins University applies novel PET radioligands that are developed to specifically target molecules relevant to biological pathophysiology. She will present her work in psychosis, highlighting the important role of smaller, yet more specific studies.

4) Prof. Ofer Pasternak, PhD, of the Departments of Psychiatry and Radiology, Harvard Medical School, is a developer of more-specific MRI measures, and has been designing acquisition protocols for large multi-site studies. He will present recent advances in the harmonization of diffusion MRI data, and discuss the trade-off between small imaging studies and large harmonized imaging studies.

The discussant will be Prof. Carrie Bearden, PhD, Department of Psychology, UCLA. Dr. Bearden had leading roles in a number of large imaging studies (e.g., NAPLS, ENIGMA, ABCD) as well as smaller studies. She will complement the panel by bringing in a more clinical point of view, informed of the practical needs of neuroscientists who are considering entering large multi-site studies.

### 14.1 LARGE-SCALE PSYCHIATRIC IMAGING STUDIES IN ENIGMA: WHAT IS GAINED AND WHAT ARE FUTURE CHALLENGES

**Neda Jahanshad*1**  
*Keck School of Medicine, University of Southern California*

**Background:** Over the last several years, researchers in the brain imaging field - and much of the biological sciences in general - began to realize that they were facing a reproducibility crisis. Null findings were rarely reported, and findings from small cohorts with nominally significant p-values were dominating publications with few or no attempts at replication. In the search for consistent patterns in the underlying neurobiology of disease, often studies of different cohorts did not agree on which brain regions were significantly different between diagnostic groups such as patients with schizophrenia and healthy controls.

Many studies had examined imaging data from less than 50 individuals, limited by the cost of data collection and the time needed to recruit and scan a large cohort. Underpowered studies can lead to both false positives and false negatives, which can damage efforts to advance science and misdirect future scientific efforts. Inflated effect sizes from small studies also pose a problem, as new studies planned using these effect sizes as benchmarks would vastly underestimate the sample sizes needed to detect true differences. The realization that findings were failing to replicate has prompted some funding agencies mandated that researchers plan to assess the rigor and reproducibility of findings. Gathering hundreds or thousands of patient imaging data samples on a single scanner, is often not feasible due to cost and recruitment constraints. In particular, for studies of rare disorders, there may only be a relatively few number of patients in a given geographic region, limiting the achievable sample size.

**Methods:** By pooling data across centers that are or have been recruiting targeted populations, consortia such has ENIGMA have been able to achieve larger sample sizes, and greater statistical power to detect reliable effects. Data and analytical harmonization methods are now allowing...