THE HARTWELL FOUNDATION

2019 Individual Biomedical Research Award

Ingo Helbig, MD

Assistant Professor
Department of Neurology and Pediatrics

University of Pennsylvania

Delineating Subgroups and Outcomes in Neurodevelopmental Disorders through Computational Phenotypes

Approximately 15% of children in the United States ages 3 to 17 yrs of age are affected by neurodevelopmental disorders (NDD), which variously include developmental delay, attention-deficit/hyperactivity disorder, autism spectrum disorder, dyslexia, intellectual disability, conduct disorders, seizures, as well as certain impairments in vision and hearing. NDD represent a significant burden to children and their families, as symptoms can last well into adulthood. The disease course of NDD is often very complex and though many have a strong genetic contribution, the cause often remains unknown. Some gene mutations have been described that overlap distinct forms of NDD, suggesting a shared genetic etiology (e.g., autism, intellectual disability and schizophrenia). In addition, thousands of common low-risk genetic variants are known that can collectively contribute to NDD susceptibility and comorbidities. Nonetheless, little is known about how genotypes correlate with shared attributes and/or clinical outcomes (phenotypes) and therefore, special needs care. The reason for this unmet need is not the lack of information or available technology but the format in which available data exists. Today, the ability to derive phenotypic data is largely limited to manual review of sparse and unstructured text in medical records, a process bottleneck complicated by the lack of an established computational framework (assumptions, principles and rules). To address this unmet need, Ingo proposes to focus on the analysis of available data rather than the process of new data generation; a systematic examination of large-scale longitudinal data derived from existing electronic medical records. Based on this conceptual framework, he will examine more than 13,000 NDD patient-parent trios for shared clinical and genetic features using the standardized vocabulary of phenotypic abnormalities (Human Phenotypic Ontology, NIH), which will enable the correlation of genotype and etiology for NDD, including defining clinically homogeneous subgroups that should represent prime targets for personalized precision medicine. Subsequently, from the electronic medical records of 5,000 children diagnosed with NDD, Ingo will derive longitudinal, multidimensional phenotypes based upon changes in specific clinical and age-dependent similarities over time. Effectively, the analysis will create a "timescape" for any gene mutation(s) that should indicate when clinical features (e.g., seizures, autism, etc.) for specific forms of NDD converge and are prominent and thus, enable a prognosis. If successful, combining genomic analysis of NDD with data analysis of large-scale electronic medical records will introduce a paradigm shift in biomedical genetics. Integrating clinical data in a computable format to discover relationships beyond the capacity of manual analysis will improve early diagnosis of NDD and predict outcome measures that should facilitate closing the gaps in special needs care and improve the lives of affected children and their families.