

Yale SCHOOL OF PUBLIC HEALTH

# DEPARTMENT OF BIOSTATISTICS

Presents the annual **Colin White Memorial Lecture**

“Scalable Analysis of Large Biobanks & Whole Genome Sequencing Studies”

**Dr. Xihong Lin**

Departments of Biostatistics and Department of Statistics  
Harvard University and Broad Institute



**Abstract.** Big data from genome, exposome, and phenome are becoming available at a rapidly increasing rate. Examples include Whole Genome Sequencing data, smartphone data, wearable devices, and Electronic Health Records (EHRs). A rapidly increasing number of large scale national and institutional biobanks have emerged worldwide. Biobanks integrate genotype, electronic health records, and lifestyle data, and is the trend of health science research. In this talk, I will discuss several analytic issues in analysis of large scale biobanks and population-based Whole Genome Sequencing (WGS) studies of common and rare genetic variants and EHRs. I will discuss scalable mixed model analysis using sparse genetic related matrix to account for relatedness and population structure; estimation of the number of ancestry principal components using Bulk Eigenvalue Matching Analysis (BEMA); and

geometric differences in PCA of multiple phenotypes and PCA of multiple genotypes. The discussions are illustrated using ongoing large scale whole genome sequencing studies of the Genome Sequencing Program of the National Human Genome Research Institute and the Trans-Omics Precision Medicine Program from the National Heart, Lung and Blood Institute, and the UK Biobank and FinnGen.

**Xihong Lin** is Professor and former Chair of the Department of Biostatistics, Coordinating Director of the Program in Quantitative Genomics at the Harvard T. H. Chan School of Public Health, and Professor of the Department of Statistics at the Faculty of Arts and Sciences of Harvard University, and Associate Member of the Broad Institute of Harvard and MIT.

**Monday, December 6, 2021**

**2:30 – 4:00 P.M. Eastern**

Please [register here](#) to receive the Zoom link to this **free virtual event**.

Dr. Colin White was a prominent biostatistician and former chair of the YSPH Department of Epidemiology & Public Health. The Colin White Memorial Lecture was established in 2013 to honor the work and life of Dr. Colin White.



Yi Dai

We will recognize the **2021 Colin White Memorial Scholarship** Recipients, **Yi Dai** and **Chichun (TJ) Tan**.

The Colin White Memorial Scholarship was created with a generous gift from Alan White in honor of his father, Dr. Colin White.



Chichun (TJ) Tan