University of California San Francisco

P HI L E

email: phillevn at gmail.com
personal website
LinkedIn and github
Google Scholar website

PROJECTS

Biostatistics/Bioinformatics

- Spatial scRNAseq with cell annotation cluster, Cell annotation and batch correction integration, Cell trajectory, velocity and development fate (notes how to, work in progress)
- Single Cell RNA sequence analysis (scRNAseq) and Colorectal cancer: We are working on a project to find gene set pathways, cell clusters and annotation, trajectory + velocity, spatial scRNA-seq, etc. that are associated with Colorectal cancer (CRC)
- Differential Gene expression, Differential Transcript Expression, and Differential Transcript Usage: I give some details and pipeline how to do differential expression at the gene level and transcript level and also for Differential Transcript Usage. I also provided a different way to do DTE analysis by the DESeq method instead of Swish. I also study the gene expression, gene set enrichment, gene set pathway from scRNAseq
- Time Series Analysis: I used time series to make predictions for new cases and new deaths of COVID-19 from Mar-Aug, 2020. My prediction was very close to actual cases within 20 days period
- Linear Mixed Effect model: I wrote a little program to find the estimates for linear mixed effect models by using the EM algorithm
- Did Increasing Continuity of Care Protects Patients with Chronic Disease from Emergency and Hospitalization Readmission? A Cohort Spatial-Temporal Study in Mississippi: We used spatial-temporal models to study the effects of locations on the risk of having Emergency visits and hospitalizations. This model is different from classical linear regressions in the sense that we account for the correlation effects of the neighborhood on locations due to Social Determinant Of Health factors
- Effect of disparities on continuity and healthcare utilization among patients with obesity-associated chronic conditions (OCC) and the subgroup with diabetes (OCC+T2D): In this project, we applied various types of generalized linear regressions together meta-analysis method to study the disparities of healthcare utilization for patients living in TN, MS, LA
- SNPs and BMI changing associations from longitudinal functional GWAS study: We apply a functional longitudinal GWAS method to find the associations between SNPs and BMI changing over time. Our method is different from longitudinal methods because we consider coefficients for genotypes as a function instead of a number in the mixed effect models. I also write a pipeline to perform the analysis for a supercomputer in the Linux system
- Associations of location and continuity of care index to 30-day unplanned hospital readmissions of patients with obesity-associated chronic conditions in Central Mississippi: This is joint work with physicians and Geographic information system (GIS) experts. We are building a model to understand factors that affect the risk of 30-day unplanned hospitalizations by using different health-related variables from GIS within a distance from patients' locations and patients' health records. Our model is special since we consider the social determinant of health factors at some fixed distance from patients' locations, and we also account for the effects of correlations of physical neighborhoods AI in Health Science
- Can we know whether a cell is cancer cell or normal cell from its single-cell RNA sequence? In this project, our goal is to classify each individual cell from a tissues sample of a cancer patient. There are always a mix of healthy and cancer cells in any sample, we would like to distinguish them individually instead of by using clustering which assign cell types if a cluster has some special markers.
- Classifying T cell receptors (TCRs) by using protein structures. Normally, researchers use the Amino Acid sequence of TCR to cluster them, however, those methods mostly focus on the order of letters of Amino Acid sequence and the distance of difference of letters at places. They ignored the protein structures of TCRs which is the backbone of the functions of TCRs - protein bindings. This project, we apply some new techniques in protein folding predictions and also natural language processing models.

AI in Health Science

- Using graph neural network to identify cancer cells from biopsy tumor samples by scRNA-seq of cells
- T cell receptor protein structures and their abundant association with cancer treatments
- Alzheimer severe prediction from MRI brain images: I did classify the severity of Alzheimer's Disease (AD) by using brain MRI images. I used a Bayesian CNN method to boost the accuracy rate, moreover, I also got the level of uncertainty of the predictions. The Bayesian method is important in the case that prediction input is not in the training population
- Pneumonia vs. normal from chest X-ray: I used Bayesian deep learning to classify pneumonia vs. normal from chest x-ray images. My project got a little better accuracy on a test set but just for a few epoch training and provide uncertainty level which is very important in medical prediction
- Blood cell detection and classification: In this project, I created a supervised deep learning model to detect and classify blood cell types in blood sample images
- Will I have Alzheimer's disease (AD) when I get old? In this project, our team is working on using genes, medical health records, and brain MRI images in the past to forecast whether someone will have AD in the next two or more decades. We are developing a deep learning statistical model to predict MRI images of the brain in a few decades by using biomarkers, health records of patients, and the first few brain MRI images scanned a few years apart. Our goal is to create a prediction model if someone will have AD in the future with high accuracy and provide some level of uncertainty. Notice that, someone is having AD now means he/she had it decades ago with no detected symptoms.

PROFESSIONAL SERVICES

- Journal Reviewer: Reviewer for Mathscinet Journal since 2016
- Founder/Organizer of Student Seminar at University of Mississippi Medical Center since 2021

EMPLOYMENT

San Francisco, CA	**University of California San Francisco**	**Sep 1, 2022 - current**
• Postdoctoral Fellow		
Jackson, MS	**University of Mississippi Medical Center**	**Fall 2018 – Aug 2022**
• Statistician Researcher		
Da Nang, Vietnam	**Duy Tan University**	**Fall 2018 – Feb 2022**
• Remote Part-time research collaborator		
Syracuse, NY	**Syracuse University**	**Fall 2016 – Spring 2018**
• Postdoctoral fellow in Mathematics		

EDUCATION

Jackson, MS	**University of Mississippi Medical Center**	**Fall 2018 – May 2020**
• Master in Biostatistics and Data Science at Department of Biostatistics and Data Science		
Columbia, MO	**University of Missouri Columbia**	**Fall 2011 – 2016**
• Ph.D in Mathematics		
• Advisor: Steve Hofmann		
Salt Lake, Utah	**University of Utah**	**Fall 2008 – 2011**
• MS in Mathematics		
Ho Chi Minh, Vietnam	**Vietnam National University**	**Fall 2001 – 2005**

- BS in Mathematics and Computer Science

TEACHING EXPERIENCE

Jackson, MS	**University of Mississippi Medical Center**	**Fall 2018 - Fall 2022**

- Workshop: Using Linux and R
- Workshop: Introduction to single cell RNA analysis

Syracuse, NY	**Syracuse University**	**Fall 2016-Spring 2018**

- Calculus I, II, III
- Introduction to Partial Differential Equations
- Putnam coach for Mathematics team (my student got ranked 59th for all students across US and Canada)

Columbia, MS	**University of Missouri**	**Fall 2011 – Summer 2016**

- Courses: Calculus I, II, III

AWARDS

Jackson, MS	**University of Mississippi Medical Center**	**2018-2022**

- ITCR 2022 workshop financial support award
- SISG2022: The 27th Summer Institute in Statistical Genetics Scholarship awards for courses: SM2214: Advanced Quantitative Genetics, SM2217: Computational Pipeline for WGS Data, SM2216: Pathway and Network Analysis for Omics Data
- Regions Research Award - School of Population Health, 2021
- Associated Student Body award in Education category, 2019

Columbia, MO	**University of Missouri Columbia**	**Fall 2011-Summer 2016**

- Invited to Hausdorff Research Institute for Mathematics, Bonn, Germany, for one month with full financial support – May-June 2014
- Highest scores for Ph.D Analysis Qualifying Exams in Mathematics – Spring 2011

PROFESSIONAL SERVICE

Jackson, MS	**University of Mississippi Medical Center**	**2018-2022**

- President of the UMMC chaper of American Statistics Association, 2019-2020
- Chair of Data Science Officer of School of Population Heath in the UMMC Associated Student Body, 2019-2020

SCIENTIFIC TALKS/PRESENTATIONS

- Geospatial analysis of colorectal cancer outcomes, behavioral risk factors, and preventive resource distribution in Mississippi: emphasis on disparities between rural and urban populations
- American College of Preventive Medicine's(ACPM) Preventive Medicine 2022 Conference, Denver Colorado June 15-18, 2022: "Effect of geographic and racial disparities on continuity of care and healthcare utilization among patients with obesity-associated chronic conditions"
- Symposium on Data Science & Statistics (SDSS), Virtual June 2-4, 2021 (Lighting talk in Data-Driven Healthcare): Did Increasing Continuity of Care Protect Patients with Chronic Disease from Emergency and Hospitalization Readmission? A Cohort Spatial-Temporal Study in Mississippi
- SGIM Southern Meeting, Feb 2021 (Poster): Does continuity of care protect patients with Obesity-Associate Chronic Conditions from Hospitalizations and Emergency Department visits?
- Northeast Analysis Network, Syracuse University, Sep 22, 2017: L^p risk wavelet density estimates for bias data, .
- Analysis seminar, Syracuse University, Oct 7, 2016: Recent progress for degenerate elliptic equations,

- Fall Eastern Sectional Meeting 2016, Bowdoin College, Brunswick, MESep 24-25, 2016: BMO solvability and absolute continuity of harmonic measure,
- Analysis Seminar, University of Missouri-Columbia, March 8, 2016: BMO solvability and absolute continuity of harmonic measure,
- SEAM 2016, University of South Florida, March 12-14, 2016: BMO solvability and absolute continuity of harmonic measure,
- Prairie Analysis Seminar 2015, Kansas State University, Sep. 25-26, 2015: *Quasi-linear PDEs and uniform rectifiability*,
- AMS Central Sectional Meeting, Michigan State University, Mar. 13-15, 2015: *L^p bounds for Riesz transforms, square root functions associated to degenerate elliptic operators*,
- SEAM 2015 - The 31st Southeastern Analysis Meeting, University of Georgia - Athens, Mar. 8-10, 2015: *Carleson measure estimates and the Dirichlet problem for degenerate elliptic equations*
- The Fifth Ohio River Analysis Meeting, Cincinnati, Ohio, Feb. 28 - Mar. 1, 2015: *Carleson measure estimates and the Dirichlet problem for degenerate elliptic equations*,
- Workshop on Harmonic Analysis, Partial Differential Equations and Geometric Measure Theory, ICMAT, Campus de Cantoblanco, Madrid January 12 - 16, 2015: *Some recent development for degenerate elliptic operators*,

PUBLICATIONS

Mathematics, Biostatistics and Health Science

- Continuity of care indices and their associations with unplanned hospital readmission within 30 days (first author, in preparation)
- Effect of geographic and racial disparities on continuity of care and healthcare utilization among patients with obesity-associated chronic conditions (first author, ready to submit)
- Effect of continuity of care on emergency department and hospital visits for obesity-associated chronic conditions: A federated cohort meta-analysis, Satya Surbhi, Ming Chen, Sohul A.Shuvo, Eboni Price-Haywood, Lizheng Shi, Joshua Mann, Yilu Lin, Phi L.Led, Jeffrey H.Burton, James E.Bailey, Journal of the National Medical Association, https://doi.org/10.1016/j.jnma.2022.07.001
- Finsler Trudinger-Moser inequalities on $\mathbf{R}^2$, with Duy, N.T., Sci. China Math. (2021). https://doi.org/10.1007/s11425-020-1820-5
- Cylindrical Hardy type inequalities with Bessel pairs, with Duy Nguyen, Operators and Matrices Volume 15, Number 2 (2021), 485–495 doi:10.7153/oam-2021-15-34
- Continuity of Care for Patients with Obesity-Associated Chronic Conditions: Protocol for a Multisite Retrospective Cohort Study, with others, JMIR Res Protoc 2020;9(9):e20788, PMID: 32902394 PMCID: PMC7511855 DOI: 10.2196/20788
- Hardy and Caffarelli-Kohn-Nirenberg inequalities with nonradial weights, with Nguyen Tuan Duy, Nguyen Thanh Son, Electronic Journal of Differential Equations, Vol. 2020 (2020), No. 33, pp. 1–10. ISSN: 1072-6691. URL: http://ejde.math.txstate.edu or http://ejde.math.unt.edu
- A note on the second order geometric Rellich inequality on half-space, with Duy, N.T., Lam, N., Monatsh Math 195, 233–248 (2021). https://doi.org/10.1007/s00605-020-01490-9
- Sharp affine Trudinger-Moser inequalities: A new argument, with N Duy and N Lam, Canadian Mathematical Bulletin, 1-14, 2020, DOI: https://doi.org/10.4153/S0008439520000806
- Quantum divergences with p-power means, with N Lam, Linear Algebra and its Applications 609, 289-307, 2020, https://doi.org/10.1016/j.laa.2020.09.009
- Hardy Inequalities and Caffarelli-Kohn-Nirenberg inequalities with radial derivative, with Nguyen Tuan Duy, Weijia Yin , 2020 Journal of Mathematical Inequalities, Volume 14, Number 2 (2020), 501–523 , dx.doi.org/10.7153/jmi-2020-14-32
- Sharp Trudinger-Moser inequalities with homogeneous weights, with Duy, Nguyen Tuan; Nghia, Le Trung; Electron J. Differential Equations 2019, N. 205, https://ejde.math.txstate.edu/Volumes/2019/105/duy.pdf

- Carleson measure estimates and the Dirichlet problem for degenerate elliptic equations, with Steve Hofmann and Andrew Morris, ANALYSIS & PDE, Volume 12, No. 8, 2019, DOI: 10.2140/apde.2019.12.2095
- BMO solvability and absolute continuity of harmonic measure, with Steve Hofmann, The Journal of Geometric Analysis, Volume 28, Issue 4, pp 3278–3299, https://doi.org/10.1007/s12220-017-9959-0
- The weak-A_∞ property of harmonic and p-harmonic measures implies uniform rectifiability, with Steve Hofmann, Kaj Nystrom, Jose Maria Martell, Analysis and PDEs. Vol 10. No. 3 2017, DOI: 10.2140/apde.2017.10.513
- Nonlinear versions of Stampacchia and Lax-Milgram theorems and applications, with Duong M. Duc and Nguyen H. Loc, Nonlinear Anal. 68 (2008), no. 4, 925931, https://doi.org/10.1016/j.na.2006.11.048

FUNDING

- Co-Investigator: CORNET project ($75,000), 2018 - 2020

COMPUTER LANGUAGES AND SKILL

- Google
- Thinking about "strange" ideas that no one did during my free time
- Building machine learning, deep learning, computer vision models for health science researches
- Computer language skill: Tensorflow, Tensorflow Probability, Keras, Scikit-learn, pytorch, SQL; SAS; R; Python; STATA; Julia (Beginning); Linux;Windows; MacOS; High Performance Computing; Google map APIs
- Genomics statistical analysis: GWAS, Functional GWAS (fGWAS) methods
- Bio-informatics: Differential Gene Expression (DESeq2 method), limma, EdgeR, Differential Transcript Expression (Swish, DESeq2 methods, gage, pathview, ...), Differential Transcript Usage (DRIMSeq + StageR method), Single-cell RNA seq (Seurat, scanpy, SingleR, machine learning, cell annotation, ...)
- Build Bayesian models by Stan
- Network Analysis and Graph Neural Networks deep learning