# Thien (Theo) Pham

400 N Negley Ave Apt 111, Pittsburgh, PA 15206 | +1.919.448.5730 | quythien14@gmail.com | linkedin.com/in/quythien


## EDUCATION

**University of Pittsburgh** August 2021 – May 2026
Ph.D. Candidate in Biostatistics (GPA: 3.85)

**University of North Carolina at Chapel Hill** September 2017 – May 2021
Bachelor of Science in Statistics, Bachelor of Science in Mathematics (Highest Distinction, GPA: 3.85)
Honors: Phi Beta Kappa, Dean's List

## SKILLS

**Programming Languages**: R, R Shiny, Python, STATA, SQL, and Bash (proficient); Java, C++, and SAS (familiar).
**Data Science:** Proficient in Statistical inference, Optimization and Stochastic models, Deep learning (PyTorch), and Machine learning (scikit-learn). Experienced in Data wrangling and Visualization, Bayesian statistics, Clinical trial, and Causal inference.
**Research Skills**: Developing methods for bio-markers detection in high-dimensional omics data, conducting causal inference and longitudinal analysis for both trial and correlated data, and constructing generative models for the integration of multimodal data.

## PROFESSIONAL EXPERIENCE

**University of Pittsburgh** August 2021 – Present
*UPMC - Graduate Research Assistant (Multi-omics Bioinformatics and Bayesian Research Assistant)*

- Utilized Hierarchical Bayesian models and Machine learning models (XGBoost, Random Forest, unsupervised clustering) to analyze high-dimensional omics data, facilitating the development of novel concordance metrics for cross-species analysis.
- Evaluated circadian biomarkers from multi-omics data in early-phase oncology trials using high-throughput computing to assess sex- and condition-specific effects across brain regions and cancer tissues in humans and mice.
- Streamlined multi-omics pipelines for preprocessing (STAR, Hisat2), enabling high-fidelity downstream analysis of diverse data types (short-read RNA-seq, single-cell RNA-seq, proteomics, etc.).
- Integrated multi-modal single-cell data (Scanpy, Seurat), using deep learning (Pytorch, VAE) and probabilistic models (Gaussian Mixture Model) to optimize cellular phenotyping with domain adaptation techniques.
- Experienced with parallel and cloud computing environments, including AWS and Google Colab, to facilitate scalable and efficient analyses of large-scale omics datasets.

*Department of Infectious Disease - Graduate Research Assistant* August 2023 – Present

- Applied mixed-effect models and Generalized Estimating Equations (GEE) to analyze clinical trial data, deriving insights to improve HIV treatment adherence and quality of life, particularly among women workers in industrial zones.
- Conducted a mediation analysis for longitudinal data to assess the causal short-term effects of dry needling (DN), using repeatedly measured pain intensity as mediators to demonstrate that reductions in pain intensity mediated DN's impact on disability.

**Carolina Center for Neurostimulation, Chapel Hill, NC** August 2019 – December 2019
*Undergraduate Research Assistant*

- Investigated cortical synchrony between the thalamus and cortex using Python and MATLAB to enhance models visualizing brain interactions, advancing brain stimulation research, and understanding of cortical dynamics.

**Hill-Rom, Cary, NC** June 2020 – August 2020
*Research Assistant Intern*

- Conducted comprehensive exploratory data analysis and literature review to provide the team with key insights for development of Machine learning models for predicting pediatric sepsis in ICU settings.

## Other Projects

*Deep learning course projects*

- Implemented deep learning models, including CNNs and RNNs, integrating VAEs and efficient sampling techniques to improve phoneme recognition and alignment with spoken inputs.
- Implemented diffusion models on ImageNet-100 with DDIM, VAE, and EMA, achieving better performance on reduced-class subsets while addressing high-class dataset challenges.

**Publications**

[1] Murillo, Carlos, et al. "Unraveling the Mechanisms Behind the Short-Term Effects of Dry Needling: New Insights from a Mediation Analysis with Repeatedly Measured Mediators and Outcomes." *Archives of Physical Medicine and Rehabilitation* (2024).

[2] [Submitted] Ha, Toan, and Pham, Thien, et al. "Exploring Relationships between Perceived Health Problems, Sexual and Reproductive Health, and Psychological Well-Being among Female Migrant Workers in Vietnam."

[3] [Submitted] Ha, Toan, and Pham, Thien, et al. "Reducing reasons for drinking among men living with HIV: A secondary analysis of a randomized controlled trial."

**AWARDS**

- Phi Beta Kappa Induction, **top 1%**, University of North Carolina at Chapel Hill 2021
- International Championship of Collegiate A Cappella – Semi-finalist, **rank 2nd** 2020
- International Championship of Collegiate A Cappella – Semi-finalist, **rank 2nd** 2019
- Outstanding student Scholarship, **top 1%**, Pham Ngoc Thach University of Medicine 2016
- Outstanding student Scholarship, **top 1%**, Pham Ngoc Thach University of Medicine 2015
- Excellent student at City level in Biology, **rank 2nd** 2014
- Excellent student at City level in Biology, **rank 2nd** 2013
- Excellent student at City level in Biology, **rank 1st** 2011

**TEACHING EXPERIENCE**

**University of Pittsburgh** **Years Taught**

*Department of Biostatistics, Graduate Teaching Assistant*

- BIOST 2086: Mixed Models Spring 2024
- BIOST 2050: Longitudinal and Clustered Data Analysis Fall 2023
- BIOST 2041: Introduction to Statistical Methods Spring 2022
- BIOST 2038: Foundation of Statistical Theory Fall 2021, Fall 2022

*Institute for Clinical Research Education, Graduate Teaching Assistant*

- CLRES 2020: Biostatistics Summer 2022, Summer 2024
- CLRES 2005: Computational Methods for Clinical Research Summer 2022, Summer 2024

**University of North Carolina – Chapel Hill**

*Department of Mathematics, Undergraduate Teaching Assistant*

- Math 547: Linear Algebra Fall 2019

**MEMBERSHIP & LEADERSHIP** **Years**

- American Statistical Association (Pittsburgh Chapter), *Member* 2021-Present
- Pittsburgh Men's Glee Club, *Member* 2023-Present
- Tar Heel Voices – UNC's co-ed a cappella group, *Fundraising Committee Member* 2019-2021
- Pham Ngoc Thach University of Medicine, School of Medicine, *Class President* 2014-2016