Xutao Wang

Department of Biostatistics
Boston University School of Public Health
801 Mass. Ave. 3rd Floor, Boston, MA, 02118

Email: xutaow@bu.edu
Website: xutao-wang.github.io

EDUCATION

Boston University 09/2019 - 06/2024
Ph.D. candidate in *Biostatistics* Boston, MA
Advisors: Dr. W. Evan Johnson and Dr. Prasad Patil

Harvard University 09/2017 - 05/2019
M.S. in *Computational Biology* Boston, MA
Advisor: Dr. Giovanni Parmigiani

University of Toronto 09/2012 - 06/2016
Honors B.S. with High Distinction in *Statistics* and *Molecular Biology* Toronto, ON

RESEARCH INTERESTS

My research focuses on developing and applying computational methodologies in transcriptomic data, especially in studying genomic biomarkers for tuberculosis diagnostics and combining models trained in heterogeneous data. Major research topics include:

- Infectious disease diagnostics
- Biomarker analysis
- Domain generalization
- Breast cancer research in African American women

I broadly collaborate with researchers from *Uganda*, *India*, and *Brazil* to study the host interface of tuberculosis (TB), with a focus on reproducibility and replicability of TB genomic biomarkers.

WORKING EXPERIENCE

GlaxoSmithKline (GSK) 05/2022 - 08/2022
Supervisor: Dr. Shashank Jariwala Cambridge, MA

- Executed an in-depth benchmark analysis on gene set scoring methods using sing-cell RNA-seq datasets with more than 10K samples.
- Conducted statistical analyses for selected scoring methods and prototyped a streamlined implementation framework within the R.
- Delivered a data-driven recommendation, aligning scoring methods with specific conditions.

University Health Network 11/2016 - 06/2017
Supervisor: Dr. Housheng Hansen He Toronto, ON

- Characterized genes differentially targeted by short hairpin RNA under CRISPR/Cas9 mediated screening.
- Performed drop out and clustering analyses for colon cancer cell-line screenings using MAGeCK/R.
- Designed single-guide RNA for CRISPR-Cas9 genome-scale knockout in human prostate cancer cells.

PUBLICATIONS

author* equal contribution

Under review

1. **Xutao Wang**, Arthur VanValkenberg, Aubrey R Odom-Mabey, Jerrold J Ellner, Natasha S Hochberg, Padmini Salgame, Prasad Patil, and W Evan Johnson. Comparison of gene set scoring methods for reproducible evaluation of multiple tuberculosis gene signatures. *bioRxiv*, pages 2023–01, 2023

2. **Xutao Wang**, Katie Harper, Pranay Sinha, William Evan Johnson, and Prasad Patil. Analysis of the cross-study replicability of tuberculosis gene signatures using 49 curated transcriptomic datasets. *bioRxiv*, pages 2023–12, 2023

Peer reviewed

3. Mollie E Barnard*, **Xutao Wang***, Jessica L Petrick, Gary R Zirpoli, Dennis Jones, W Evan Johnson, and Julie R Palmer. Psychosocial stressors and breast cancer gene expression in the black women's health study. *Breast Cancer Research and Treatment*, pages 1–14, 2023

4. Vaishnavi Kaipilyawar*, Yue Zhao*, **Xutao Wang***, Noyal M Joseph, Selby Knudsen, Senbagavalli Prakash Babu, Muthuraj Muthaiah, Natasha S Hochberg, Sonali Sarkar, Charles R Horsburgh Jr, et al. Development and validation of a parsimonious tuberculosis gene signature using the digital nanostring ncounter platform. *Clinical Infectious Diseases*, 75(6):1022–1030, 2022

5. Dylan Sheerin, Nashied Peton, William Vo, Cody Charles Allison, **Xutao Wang**, W Evan Johnson, Anna Kathleen Coussens, et al. Immunopathogenic overlap between covid-19 and tuberculosis identified from transcriptomic meta-analysis and human macrophage infection. *Iscience*, 25(6), 2022

TEACHING

Teaching Assistant

- CI 670: *Biostatistics with Computing*, Dept. of Medicine, Boston University — Fall 2023
- CI 670: *Biostatistics with Computing*, Dept. of Medicine, Boston University — Fall 2021

Workshop

- Host-pathogen interaction in TB and TB-HIV (presented remotely)
  Makerere University, Kampala, Uganda — June 2023

PRESENTATIONS

Conference talks

1. 54th Union World Conference on Lung Health, Paris, France — 11/2023
   Analysis of the cross-study replicability of tuberculosis gene signatures using 49 curated transcriptomic datasets

2. 36th New England Statistics Symposium, Boston, USA — 06/2023
   Multi-study learning for blood-based transcriptomic biomarkers in tuberculosis

3. RePORT India 10th Annual Meeting, Convened Virtually 02/2021
   Comparison of tuberculosis gene signatures using original models and gene set scoring methods

Seminar talks

1. Division of Infectious Disease, Rutgers New Jersey Medical School 05/2023
   Analysis of the cross-study replicability of tuberculosis gene signatures using 49 curated transcriptomic datasets
2. Department of Biostatistics, Boston University School of Public Health 04/2023
   Methods for reproducible evaluation of transcriptomic biomarkers in tuberculosis
3. Division of Computational Biomedicine, Boston University School of Medicine 01/2023
   Multi-study learning for blood-based transcriptomic biomarkers in tuberculosis
4. Department of Biostatistics, Boston University School of Public Health 04/2022
   Improving the predictive ability of existing TB gene signatures via ensemble learning
5. Division of Computational Biomedicine, Boston University School of Medicine 03/2022
   Analysis of the cross-study replicability of tuberculosis gene signatures using 49 curated transcriptomic datasets
6. Division of Computational Biomedicine, Boston University School of Medicine 05/2021
   Comparison of gene set scoring methods and original models on TB biomarkers
7. Tuberculosis Interdisciplinary Group, Boston Medical Center 09/2020
   curatedTBData: Clinically annotated data for tuberculosis transcriptomics

MENTORING

Please Lukau, Undergraduate student at Southern Utah University 01/2024-05/2024
Methods and web-based tools for multi-study learning and molecular biomarker exploration

Alex Sierra, Undergraduate student at Southern Utah University 01/2024-05/2024
Methods and web-based tools for multi-study learning and molecular biomarker exploration

Andrew Smith, Undergraduate student at Southern Utah University 01/2024-05/2024
Methods and web-based tools for multi-study learning and molecular biomarker exploration

Elie Wamana, Undergraduate student at Southern Utah University 01/2024-05/2024
Methods and web-based tools for multi-study learning and molecular biomarker exploration

Samantha MacDonald, Undergraduate student at Brigham Young University 05/2023-02/2024
Data curation and democratization for TB gene signatures

PROFESSIONAL SERVICE

Reviewer for Scientific Journals (# papers in parentheses):
Journal of the Royal Society Interface (1), BMC Bioinformatics (1)

SOFTWARE PACKAGES

1. **curatedTBData**: An R package focusing on curating and democratizing of tuberculosis transcriptomic studies

2. **TBSignatureProfiler**: An R package focusing on profiling RNA-Seq data using tuberculosis pathway signatures

SKILLS

Programming Languages	R, Python, MySQL, Java, Bash
Tools	Git/GitHub, LaTeX, High-Performance Computing Cluster

AWARDS

- The Alumni of Victoria College Award 2015
- The Isabel Bader In-Course Scholarship 2014
- Summer Research Award 2014
- Dean's List Scholar 2014 - 2016

REFERENCES

W. Evan Johnson, Ph.D.
Professor of Medicine, Division of Infectious Diseases
Director, Center for Data Science
Rutgers University - New Jersey Medical School

Prasad Patil, Ph.D.
Assistant Professor
Department of Biostatistics
Boston University School of Public Health

Julie Palmer, ScD
Karin Grunebaum Cancer Research Professor
Director, Slone Epidemiology Center
Boston University Chobanian & Avedisian School of Medicine