

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. in <i>Biostatistics</i>	Boston, MA
Advisors: Dr. W. Evan Johnson and Dr. Prasad Patil	
Harvard University	09/2017 - 05/2019
M.S. in <i>Computational Biology</i>	Boston, MA
Advisor: Dr. Giovanni Parmigiani	
University of Toronto	09/2012 - 06/2016
B.S. in <i>Statistics and Molecular Biology</i>	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
• Conducted a comprehensive benchmark analysis of six gene set scoring methods using simulated and publicly available single-cell RNA-seq datasets comprising over 10K samples. • Performed advanced statistical analyses and developed a streamlined implementation framework in R, enhancing the evaluation of multiple single-cell scoring methods. • Delivered a data-driven recommendation, aligning scoring methods with specific conditions.	

PUBLICATIONS

author* equal contribution

Under review

1. **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

2. **Xutao Wang**, Boyu Ren, Prasad Patil, and W Evan Johnson. Multi-study adaptive learning with applications to transcriptomic biomarkers in tuberculosis. *In preparation*, 2024
3. Andrew Smith, Alexander Sierra, Please Lukao, W Evan Johnson, and **Xutao Wang**. curat-edtbexplorer: facilitating novel multi-study tuberculosis biomarker development. *In preparation*, 2024

Peer reviewed

3. **Xutao Wang**, Arthur VanValkenberg, Aubrey R Odom, Jerrold J Ellner, Natasha S Hochberg, Padmini Salgame, Prasad Patil, and W Evan Johnson. Comparison of gene set scoring methods for reproducible evaluation of tuberculosis gene signatures. *BMC Infectious Diseases*, 24(1):610, 2024
4. 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
5. 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
6. 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

GRANT EXPERIENCE

I actively contributed to the proposal writing of Dr. W. Evan Johnson. My contribution included proposing specific aims and preparing preliminary results & main figures.

U01RFA-AI-23-038 (Johnson, Wei) pending
NIH/NIAID: *Development of Interactive Web server and Data Science tools for gene expression data and pathway profiling from Multiple Infectious Diseases*

R01AI175315-01 (Johnson, Salgame, Song) 02/01/2023 – 01/31/2028
NIH/NIAID: *Biomarker Signatures of TB Infection in Young Children with and Without HIV*

R21AI154387-01 (Johnson, Patil)

07/10/2020 – 06/30/2022

NIH/NIAID: *Signature of Profiling and Staging the Progression of TB from Infection to Disease***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-present
Methods and web-based tools for multi-study learning and molecular biomarker exploration
- Alex Sierra, Undergraduate student at Southern Utah University 01/2024-present
Methods and web-based tools for multi-study learning and molecular biomarker exploration
- Andrew Smith, Undergraduate student at Southern Utah University 01/2024-present
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, L ^A T _E X, High-Performance Computing Cluster, Shiny app design, PyTorch, Scikit-learn

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 |