Zhiguang Huo

Contact Information	4600 Bayard ST, APT 308 Pittsburgh, PA 15213	412-979-0592 zhh18@pitt.edu

Research Interests

Omics data integration, Machine learning and statistical learning, Genomics and epigenetics algorithms and applications, Bioinformatics, Optimization, Graphical model

Education

University of Pittsburgh, Pittsburgh, PA, US

Ph.D., Biostatistics, *Expected:* Summer 2017

- Thesis Topic: *Statistical integrative omics methods for disease subtype discovery and single cell methylation methods*
- GPA: *3.93/4.00*
- Advisors: George C. Tseng, ScD and Yong Seok Park, PhD

M.S., Physics, Apr 2012

- *GPA: 3.86/4.00*

Harbin Institute of Technology, Harbin, Heilongjiang, China

B.S., Physics, June 2011

- GPA: *90.43/100*

Research Experience

Research Assistant Dec 2011 to present
Department of Biostatistics,
University of Pittsburgh
Supervisor: George C. Tseng, ScD

Research Assistant Aug 2014 to present
Department of Biostatistics,
University of Pittsburgh
Supervisor: Yong Seok Park, PhD

Collaboration Feb 2016 to present
Department of Psychiatry,
University of Pittsburgh
Superviser: George C. Tseng, ScD
Collaborator: Colleen A. McClung, Ph.D, Marianne Seney, PhD, Ryan Logan, PhD

Collaboration Oct 2015 to present
Department of Pediatrics,
University of Pittsburgh
Supervisor: George C. Tseng, ScD
Collaborator: Nader Shaikh , MD, MPH

Collaboration Jul 2014 to Feb 2016
Department of Psychiatry,
University of Pittsburgh
Supervisor: George C. Tseng, ScD
Collaborator: David A. Lewis, MD, John F. Enwright, Ph.D., Dominique Arion, Ph.D.

Collaboration Jul 2014 to Sep 2015
Department of Pharmacology and Chemical Biology,
Magee-Womens Research Institute
Supervisor: George C. Tseng, ScD Collaborator: Steffi Oesterreich, PhD

Collaboration Jan 2013 to Sep 2014

Department of Pathology,
University of Pittsburgh
Supervisor: George C. Tseng, ScD Collaborator: Jianhua Luo, MD, PhD, Yan Ping Yu, MD, PhD

Collaboration Dec 2012 to Nov 2013
Department of Anesthesiology and Neurobiology,
University of Pittsburgh
Supervisor: George C. Tseng, ScD Collaborator: William R. Lariviere, PhD

Collaboration Mar 2012 to Aug 2012
Department of Environmental and Occupational Health,
University of Pittsburgh
Supervisor: George C. Tseng, ScD Collaborator: George D Leikauf, PhD

Refereed Journal Publications (Statistical)

1. **Zhiguang Huo**, Ying Ding, Silvia Liu, Steffi Oesterreich, and George Tseng. Meta-Analytic Framework for Sparse K-Means to Identify Disease Subtypes in Multiple Transcriptomic Studies. *Journal of the American Statistical Association*, 111, no. 513 (2016): 27-42.

2. **Zhiguang Huo**, Chi Song, George C. Tseng. (2016) Bayesian latent hierarchical model for transcriptomic meta-analysis to detect biomarkers with clustered meta-patterns of differential expression signals. Submitted to *Annals of Applied Statistics* (under second round of review).

3. **Zhiguang Huo**, George C. Tseng. (2016) Integrative Sparse K-means for disease subtype discovery using multi-level omics data. Submitted to *Annals of Applied Statistics* (under second round of review).

4. Li Zhu, Ying Ding, Cho-Yi Chen, Lin Wang, **Zhiguang Huo**, SungHwan Kim, Christos Sotiriou, Steffi Oesterreich and George C. Tseng. (2016) MetaDCN: meta-analysis framework for differential coexpression network detection with an application in breast cancer. *Bioinformatics* (accepted).

5. SungHwan Kim, Dongwan Kang, **Zhiguang Huo**, Yongseok Park, George C. Tseng. (2016) Meta-analytic principal component analysis. Submitted to *Annals of Applied Statistics* (under revision).

6. Silvia Liu, Wei-Hsiang Tsai, Ying Ding, Rui Chen, Zhou Fang, **Zhiguang Huo**, SungHwan Kim, Tianzhou Ma, Ting-Yu Chang, Nolan Michael Priedigkeit, Adrian V. Lee, Jianhua Luo, Hsei-Wei Wang, I-Fang Chung, George C. Tseng. (2015). Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data. *Nucleic Acids Research*, 10.1093/nar/gkv1234.

7. Xingbin Wang, Dongwan Kang, Kui Shen, Chi Song, Shuya Lu, Lunching Chang, Serena G. Liao, **Zhiguang Huo**, Naftali Kaminski, Etienne Sibille, Yan Lin, Jia Li and George C. Tseng. (2012) A Suite of R Packages for Quality Control, Differentially Expressed Gene and Enriched Pathway Detection in Microarray Meta-analysis. *Bioinformatics*, 28:2534-2536.

Manuscript in preparation (Statistical)

1. **Zhiguang Huo**, Shaowu Tang, YongSeok Park and George Tseng. Biomarker categorization and fast computing of adaptively weighted Fisher's method for meta-analysis in omics applications.

2. AW theory

3. Li Zhu, **Zhiguang Huo**, Tianzhou Ma, George Tseng. Bayesian indicator variable selection model with multi-layer overlapping groups.

4. Tianzhou Ma, **Zhiguang Huo**, . . ., George Tseng. MetaOmics - a Comprehensive Software Suite with Interactive Visualization for Transcriptomic Meta-Analysis

Refereed Journal Publications (Application)

1. Tiffany A. Katz, Serena G. Liao, Vincent J. Palmieri, Robert K. Dearth, Thushangi Pathiraja, **Zhiguang Huo**, Patricia Shaw, Sarah Small, Nancy E. Davidson, David G. Peters, George C. Tseng, Steffi Oesterreich, Adrian V. Lee. (2015) Targeted DNA methylation screen in the mouse mammary genome reveals a parity-induced hypermethylation of igf1r which persists long after parturition. *Cancer Prevention Research*, pages canprevres-0178.

2. Yan P. Yu, Silvia Liu, **Zhiguang Huo**, Amantha Martin, Joel B. Nelson, George C. Tseng and Jian-Hua Luo. (2015) Genomic copy number variations in the genomes of leukocytes predict prostate cancer clinical outcomes. *PloS one*, 10(8):e0135982.

3. Dominique Arion, **Zhiguang Huo**, John F. Enwright, John P. Corradi, George Tseng and David A. Lewis. Transcriptome alterations in prefrontal pyramidal neurons distinguish schizophrenia from bipolar and major depressive disorders. Submitted to *Biological Psychiatry*, (under second round of review).

Published Abstract

1. Oesterreich, S., Katz, T.A., Logan, G., Levine, K., Nagle, A., **Huo, Z.**, Tseng, G.C., Rui, H., Lee, A.V. and Butler, L.M., 2016. Abstract PD2-08: Potential role of prolactin signaling in development and growth of the lobular subtype of breast cancer. *Cancer Research*, 76(4 Supplement), pp.PD2-08.

2. Enwright, John, Dominique Arion, John Corradi, Aiqing He, **Zhiguang Huo**, George Tseng, and David Lewis. (2015) Transcriptome Profiling of Layer 3 Parvalbumin Neurons from the Dorsolateral Prefrontal Cortex of Schizophrenia Subjects. *NEUROPSYCHOPHARMACOLOGY*, vol. 40, pp. S400-S401.

Book and Book Chapter

1. George C. Tseng, **Zhiguang Huo** and Tianzhou Ma. Foundations for High-Throughput Omics Data Analysis: Methods, Theories and Applications. Chapman & Hall/CRC.

2. **Zhiguang Huo**, Shaowu Tang, YongSeok Park and George Tseng. (2015) MetaOmics: transcriptomic meta-analysis methods for biomarker detection, pathway analysis and other exploratory purposes. Book chapter in Integrating omics data: statistical and computational methods.

Awards

Student Awards

- American Statistics Association (ASA) Pittsburgh chapter — March 2016
  - Student of the year
- Department of Physics, Harbin Institute of Technology — May 2009
  - National Scholarship of P.R. China.

  (Awarded to the top 2 students in my Bachelors degree.)

Travel Awards

- SAMSI Research Triangle Park, NC.
  - Optimization Summer School — Aug 2016
  - Epigenetics Workshop — Mar 2015
  - Beyond Bioinformatics Workshop — June 2014

Teaching Experience (University of Pittsburgh)

Lecture

- BIOST2094 - Advanced R Computing — Jan 2017
  - 6 lectures on Advanced R Computation
- BIOST2025 - Special Studies in Bayesian Data Analysis — Oct 2016
  - 4 lectures on Advanced Bayesian Computation

Guest lecture

- BIOST2055 - Introductory high-throughput genomic data analysis I: data mining and applications — Mar 2016
  - Differential and isoform analysis of RNA-seq data
- BIOST2078 - Introductory high-throughput genomic data analysis II: theories and algorithms — Dec 2015
  - Reproducible research and parallel computing in R
- BIOST2078 - Introductory high-throughput genomic data analysis II: theories and algorithms — Dec 2014
  - Reproducible research

Teaching assistant

- BIOST 2078 - Introductory high-throughput genomic data analysis II: theories and algorithms — Sep 2014 - Dec 2014
- PHYS 0212 - Introduction to Laboratory Physics — Jan 2012 - Apr 2012
- PHYS 0212 - Introduction to Laboratory Physics — Aug 2011 - Dec 2011

Presentations

Statistical Meetings

- Poster, SAMSI optimization summer school, Research Triangle Park, NC — Aug 2016
  - Integrative Sparse K-means for disease subtype discovery using multi-level omics data.
- Poster, Pittsburgh ASA banquet, Pittsburgh, PA — Mar 2016
  - Integrative Sparse K-means for disease subtype discovery using multi-level omics data.
- Oral Presentation, JSM, Seattle, WA — Aug 2015
  - Meta-analytic framework for sparse K-means to identify disease subtypes in multiple transcriptomic studies.
- Poster, Pittsburgh ASA banquet, Pittsburgh, PA — Apr 2015
  - Meta-analytic framework for sparse K-means to identify disease subtypes in multiple transcriptomic studies.
- Oral Presentation, ENAR Conference, Miami, FL — Mar 2015
  - Meta-analytic framework for sparse K-means to identify disease subtypes in multiple transcriptomic studies.
- Poster, Dean's Day's competition, GSPH, University of Pittsburgh — Mar 2015
  - Discover and Characterize Invasive Lobular Breast Carcinoma Subtypes.
- Oral Presentation, ENAR Conference, Baltimore, MA — Mar 2014
  - Meta-analytic framework for sparse K-means to identify disease subtypes in multiple transcriptomic studies.
- Poster, Dean's Day's competition, GSPH, University of Pittsburgh — Mar 2014
  - Meta-analytic framework for sparse K-means to identify disease subtypes in multiple transcriptomic studies.

Department of Biostatistics, University of Pittsburgh

- Seminar talk — Nov 2015
  - How to use Latex to make slides

References

George C. Tseng

Professor
Department of Biostatistics (primary appointment)
Department of Human Genetics
Department of Computational & Systems Biology
University of Pittsburgh
Phone: 412-624-5318
E-mail: ctseng@pitt.edu

Yong Seok Park
Assistant Professor
Department of Biostatistics
University of Pittsburgh
Phone: 412-624-3028
E-mail: yongpark@pitt.edu

David A. Lewis, MD
Distinguished Professor of Psychiatry and Neuroscience,
Thomas Detre Professor of Academic Psychiatry, Chair
Department of Psychiatry
University of Pittsburgh
Phone: 412-246-6010
E-mail: lewisda@upmc.edu

Steffi
Colleen
Luo

Skills

Computer Programming:

- gitbub (1 years),
- R (5 years),
- Linux (5 years),
- Matlab (2 year),
- Python (1 year),
- Java, C++, C (1 year).

Running

- Bank of America Chicago Marathon, Chicago, IL 10/09/2016
  TIME 04:08:27
- First National Bank Pittsburgh Triathlon (Sprint), Pittsburgh, PA 08/14/2016
  TIME 1:10:30 (only biking + running)
- Dick's Sporting Goods Pittsburgh Marathon, Pittsburgh, PA 05/01/2016
  TIME 3:58:46
- First National Bank Pittsburgh Triathlon (Olympic standard), Pittsburgh, PA 08/09/2015
  Time 3:29:24
- Dick's Sporting Goods Pittsburgh Marathon, Pittsburgh, PA 05/03/2015
  TIME 3:49:38
- Dick's Sporting Goods Pittsburgh Marathon, Pittsburgh, PA 05/04/2014
  TIME 4:07:44
- Dick's Sporting Goods Pittsburgh Marathon, Pittsburgh, PA 05/05/2013
  TIME 4:48:49