

HAOYU ZHANG

677 Huntington Ave, Boston, MA 02115, USA

✉ haoyuzhang@hsph.harvard.edu

EDUCATION

Johns Hopkins University, Department of Biostatistics, **Baltimore, USA**
Ph.D. in Biostatistics, Advisor: Nilanjan Chatterjee and Ni Zhao *Sep 2014 - May 2019*

Zhejiang University, Department of Mathematics **Hangzhou, China**
Bachelor of Science in Statistics *Sep 2010 - Jun 2014*

PROFESSIONAL EXPERIENCE

Harvard University, Department of Biostatistics **Boston, USA**
Postdoctoral fellow, Advisor: Xihong Lin *Aug 2019 - Present*

National Cancer Institute **Rockville, USA**
Special volunteer, Advisor: Montserrat García-Closas *Aug 2019 - Present*

National Cancer Institute **Rockville, USA**
Pre-doctoral fellow, Advisor: Montserrat García-Closas *July 2017 - May 2019*

Johns Hopkins University, Department of Biostatistics **Baltimore, USA**
Research assistant, Advisor: Nilanjan Chatterjee *Oct 2015 - May 2019*

Johns Hopkins University, Department of Biostatistics **Baltimore, USA**
Research assistant, Advisor: Thomas A. Louis *July 2015 - Dec 2017*

Yale University, Department of Biostatistics **New Haven, USA**
Research assistant, Advisor: Hongyu Zhao *Jul 2013 - Sep 2013*

The University of Queensland, Queensland Brain Institute **Brisbane, Australia**
Research assistant, Advisor: Naomi Wray *Jul 2012 - Aug 2012*

GRANTS AND FELLOWSHIPS

1. Pre-doctoral fellowship, National Cancer Institute, Rockville, USA, July 2017- May 2019

HONORS AND AWARDS

Johns Hopkins University:

1. ASA Section on Statistics in Genomics and Genetics (SGG) **Distinguished Student Paper Award**, Denver, USA, Aug 2019

2. Delta Omega Scholarship, Baltimore, USA, May 2019
3. Delta Omega Public Health Honor Society, Alpha Chapter, Baltimore, USA, April 2019
4. Louis I. and Thomas D. Dublin Award, Baltimore, USA, March 2019
5. DCEG Fellows Award for Research Excellence (DFARE), National Cancer Institute, Rockville, USA, Nov 2018 (**only graduate student winner in 2018**)
6. Joseph Zeger Travel award to ENAR, Baltimore, USA, March 2018
7. Symposium on Advances in Genomics, Epidemiology, and Statistics (SAGES) Poster Award, University of Pennsylvania, Philadelphia, USA, June 2017

Zhejiang University:

1. Meritorious Winner in the American Mathematical Contest in Modeling, China, 2013
2. First Class Scholarship for Outstanding Merit, China, 2011
3. National Scholarship, Ministry of Education, China, 2011 (Top 2 %)
4. Excellent Student Honor, China, 2011
5. Scholarship for Distinguished Social Work, China, 2011

PREPRINTS

* represents joint first author

1. Ahearn, T.U.*, **Zhang, H.***, ..., Garcia-Closas, M. Chatterjee, N., Common breast cancer risk loci predispose to distinct tumor subtypes. *Submitted to Nature Communication*
2. Dikshit R. P., Mhatre S., Richmond R., ..., **Zhang H.**, ..., Patkar S., The role of gallstones and gallbladder cancer in India: a Mendelian randomization study (*submitted to The Lancet Oncology*)

PUBLICATIONS

* represents joint first author

1. Zhang, Y. D., Hurson, A. N., **Zhang, H.**, ..., Chanock, S. J., Chatterjee, N., Garcia-Closas, M. (2020). Assessment of polygenic architecture and risk prediction based on common variants across fourteen cancers. *Nature communications*, 11(1), 1-13.
2. **Zhang, H.***, Ahearn, T. U.*, ..., Easton, D., Chatterjee, N., Garcia-Closas, M. (2020). Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and

subtype-specific analyses. *Nature Genetics*, 1-10.

3. **Zhang, H.**, Zhao, N., Ahearn, T.U., Garcia-Closas, M., Chatterjee, N. (2020). A mixed-model approach for powerful testing of genetic associations with cancer risk incorporating tumor characteristics. *Biostatistics*.
4. Koutros, S., Kogevinas, M., Friesen, M. C., ..., **Zhang, H.**, ..., Tardon, A. (2020). Diesel exhaust and bladder cancer risk by pathologic stage and grade subtypes. *Environment International*, 135, 105346.
5. Zhao, N., **Zhang, H.**, Clark, J., Maity, A., Wu, M. (2018). Composite Kernel Machine Regression based on Likelihood Ratio Test with Application for Combined Genetic and Gene-environment Interaction Effect. *Biometrics*, 75(2), 625-637.
6. Hu, Q., Ding, Z., **Zhang, H.**, He, Y. (2018). The Superior Glenohumeral Joint Capsule Alone Does Not Prevent Superior Translation of the Humeral Head: An In Vitro Biomechanical Study. *Arthroscopy: The Journal of Arthroscopic Related Surgery*, 34(11), 2962-2970.
7. Derkach, A., **Zhang, H.**, Chatterjee, N. (2017). Power Analysis for Genetic Association Test (PAGEANT) provides insights to challenges for rare variant association studies. *Bioinformatics*, 34(9), 1506-1513.
8. Mhatre, S., Wang, Z., Nagrani, R., ..., **Zhang, H.**, ..., Dikshit, R., Chatterjee, N., Rajaraman, P. (2017). Common genetic variation and risk of gallbladder cancer in India: a case-control genome-wide association study. *The Lancet Oncology*, 18(4), 535-544.

PRESENTATIONS

Platform Talk:

1. Genome-wide Association Study (GWAS) Identifies 19 Novel Breast Cancer Loci From Analyses Accounting For Subtype Heterogeneity, *American Society of Human Genetics Annual Conference Platform Talk*, Orlando, USA, October 2017

Invited Talks:

1. Methods Methods for Developing Ancestry Specific Polygenic Risk Scores, *Breast Cancer Association Consortium (BCAC) 25nd meeting*, virtual meeting, Sep 2020
2. A Mixed-Model Approach for Powerful Testing of Genetic Associations with Cancer Risk Incorporating Tumor Characteristics, *Southwestern University of Finance and Economics*, Chengdu, China, January, 2020
3. A Mixed-Model Approach for Powerful Testing of Genetic Associations with Cancer Risk Incorporating Tumor Characteristics, *Shanghai Jiao Tong University*, Shanghai, China, December 2019

4. Subtypes Specific GWAS and Risk Prediction Modeling on Breast Cancer Subtypes, *Breast Cancer Association Consortium (BCAC) 23rd meeting*, Springdale, USA, April 2019
5. Standard Approach and Methods Accounting for Subtype Heterogeneity Identified 33 Novel Breast Cancer Loci, *Breast Cancer Association Consortium (BCAC) 22nd meeting*, Edinburgh, UK, June 2018
6. Testing for Genetic Associations in Case-control Studies Incorporating Multivariate Disease Characteristics, *National Cancer Institute, Division of Cancer Epidemiology & Genetics, Integrative Tumor Epidemiology Branch*, Rockville, USA, September 2017

Topic Contributed Talks:

1. Testing for Genetic Associations in Case-control Studies Incorporating Multivariate Disease Characteristics, *Joint Statistical Meetings*, Baltimore, USA, August 2017

Contributed Talks:

1. Robust Mendelian Randomization (MR) Analysis with Weak Instruments, *Joint Statistical Meetings*, virtual meeting, Aug 2020
2. A Mixed-Model Approach for Powerful Testing of Genetic Associations with Cancer Risk Incorporating Tumor Characteristics, *Joint Statistical Meetings*, Denver, USA, Aug 2019
3. Developing Polygenic Risk Prediction models for cancer subtypes incorporating multivariate disease characteristics, *Joint Statistical Meetings*, Vancouver, CA, July 2018
4. The Mixture Approach to Estimating a Weighted Population-averaged Value, *Eastern North American Regional meeting of the International Biometric Society*, Atlanta, USA, March 2018
5. The Mixture Approach to Estimating a Population-averaged Value, *Eastern North American Regional meeting of the International Biometric Society*, Washington, USA, March 2017

Posters:

1. A mixed-model approach for powerful testing of genetic associations with cancer risk incorporating tumor characteristics, *American Society of Human Genetics Annual Conference*, San Diego, USA, October 2018
2. Genome-wide association study (GWAS) identifies 19 novel breast cancer loci from analyses accounting for subtype heterogeneity, *Johns Hopkins Biostatistics Retreat*, Philadelphia, USA, April 2018
3. Genome-wide association study (GWAS) identifies 19 novel breast cancer loci from analyses accounting for subtype heterogeneity, *14TH Annual NIH Graduate Student Research Symposium*, Bethesda, USA, February 2018

4. Genome-wide association study (GWAS) identifies 19 novel breast cancer loci from analyses accounting for subtype heterogeneity, *Delta Omega Poster Competition*, Baltimore, USA, Feb 2018
5. Genome-wide association study (GWAS) identifies 19 novel breast cancer loci from analyses accounting for subtype heterogeneity, *The Maryland Genetics, Epidemiology, and Medicine Training Program (MD-GEM)*, Baltimore, USA, February 2018
6. Methods for Analyzing Genome-wide Association Studies, *Johns Hopkins University, Department of Biostatistics, 100 years celebration*, Baltimore, USA, July 2017
7. Testing for genetic association in case-control studies incorporating multivariate disease characteristics, *Symposium on Advances in Genomics, Epidemiology, and Statistics (SAGES)*, Philadelphia, USA, June 2017
8. Testing for genetic association in case-control studies incorporating multivariate disease characteristics, *The Maryland Genetics, Epidemiology, and Medicine Training Program (MD-GEM)*, Baltimore, USA, February 2017

TEACHING EXPERIENCE

Lab Instructor:

Responsibilities: teaching 3 to 4.5 hours labs per week, beta-test, proctor and grade exams

1. Statistical Methods in Public Health II-III (Biostat 622-623), Mater Level Class, taught by Dr. Marie Diener-West and Dr. Karen Bandeen-Roche, Baltimore, USA, Oct 2018 - Mar 2019
2. Statistical Methods in Public Health II-III (Biostat 622-623), Mater Level Class, taught by Dr. Marie Diener-West and Dr. Karen Bandeen-Roche, Baltimore, USA, Oct 2017 - Mar 2018

Teaching Assistant:

Responsibilities: hold office hour, grade homework and exams, and hold some classes

1. Statistical Methods in Public Health I-IV (Biostat 621-624), Mater Level Class, taught by Dr. Marie Diener-West, Dr. Karen Bandeen-Roche, Dr. John McGready and Dr. James Tonascia, Baltimore, USA, Sep 2016 - May 2017
2. Statistical Theory I-IV (Biostat 731-734), PhD Level Class, taught by Dr. Elizabeth L. Ogburn, Dr. Constantine Frangakis and Dr. Rosenblum Michael, Baltimore, USA, Sep 2015 - May 2016
3. Survey Methodology, Undergraduate Level Class, taught by Dr. Peng Zhang, Hangzhou, China, Feb 2014 - Jun 2014

SOFTWARE

R shiny app:

1. **PAGEANT**: Power Analysis for GEnetic AssociationN Tests
Software link: <https://andrewhaoyu.shinyapps.io/PAGEANT/>
Github link: <https://github.com/andrewhaoyu/PAGEANT>

R package:

1. **CKLRT**: Composite Kernel Machine Regression Based on Likelihood Ratio Test
CRAN link: <https://cran.r-project.org/web/packages/CKLRT/index.html>
Github link: <https://github.com/andrewhaoyu/CKLRT>
2. **TOP**: Two-stage polytomous logistic regression
Github link: <https://github.com/andrewhaoyu/TOP>

EDITORIAL ACTIVITIES

Journal Referee

1. Nature Communication
2. Plos Computational Biology

SOCIAL ACTIVITIES

1. Statistical genetics group website maintainer (<http://jhustatgen.org>), Johns Hopkins University, Department of Biostatistics, Baltimore, USA, August 2018 - now
2. Session chair, Statistical Advances in the Design and Analysis of Sequence-Based Genetic Association Studies, Joint Statistical Meeting, Baltimore, USA, August 2018
3. **Student representative**, Johns Hopkins University, Department of Biostatistics, Baltimore, USA, August 2017 - August 2018
 - o Hold the student lunches with topics including passing qualify exam and finding summer internship
 - o Organizes sessions to help students writing cover letter for summer internship
 - o Attend faculty monthly meetings
4. TA training day planning committee, Johns Hopkins University, Department of Biostatistics, Baltimore, USA, August 2018
5. **Johns Hopkins Biostatistics Mixer Organizer**, Joint Statistical Meeting, Vancouver, CA, August 2018
6. **Johns Hopkins Biostatistics Mixer Organizer**, Eastern North American Regional meeting,

Atlanta, USA, August 2018

7. TA training day planning committee, Johns Hopkins University, Department of Biostatistics, Baltimore, USA, August 2017
8. TA training day planning committee, Johns Hopkins University, Department of Biostatistics, Baltimore, USA, August 2016
9. **Volunteer service**, Sharadadevi Gramudyog Utpadak Mandali Ltd (SGS), **Baroda, India**, June 2011 - August 2011
 - o Researched micro-finance projects' impact on villagers' lives in Baroda teaming up with interns from 12 countries
 - o Surveyed 16 villages to collect data on sanitation, education and income levels

MENTORING

1. Sayan Das (Indian Statistical Institute), summer student, guidance on large scale simulation and parallel jobs on computing nodes, Baltimore, USA, June 2017 - August 2017

PROGRAMMING SKILLS

R (50k-60k lines), C (10k lines), Shell Script, RcppEigen, Stan, Photoshop, MATLAB, PLINK, QCTOOL, GTOOL, STATA, SPSS, E-View, Origin