

CURRICULUM VITAE: YAN LI

Notarization. I have read the following and certify that this *curriculum vitae* is a current and accurate statement of my professional record. Signature Yan Li Date 12/2020

1. PERSONAL INFORMATION

(a) UID, LAST NAME, FIRST NAME, CONTACT INFORMATION

UID: yli6; Last Name: Li; First Name: Yan

Contact Information:

1218 LeFrak Hall

University of Maryland

College Park, MD 20742-4015

Email: yli6@umd.edu

Phone number: 301-314-6570

(b) ACADEMIC APPOINTMENT AT UMD

Yan Li, Ph.D.

Professor

Joint Program in Survey Methodology

College of Behavioral and Social Sciences

and

Department of Epidemiology and Biostatistics

School of Public Health

University of Maryland at College Park, MD 20742

(c) OTHER ACADEMIC APPOINTMENTS WHILE AT UMD

Adjunct Professor, Survey Methodology Program, Institute for Social Research

University of Michigan, Ann Arbor, MI

(d) EDUCATIONAL BACKGROUND

Dec. 2006-Jul. 2009

PostDoc, Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute

Field of Study: Survey Statistics and Biostatistics

Mentor: Dr. Barry Graubard

Jan. 2003-Dec. 2006

Ph.D., University of Maryland at College Park

Major: Survey Methodology

Thesis Topic: Analysis of Complex Survey Data Using Model-Based and Model-Assisted Methods

Research Advisor: Prof. Partha Lahiri

Aug. 2001-Dec. 2002

M.S., University of Nebraska at Lincoln

Major: Statistics

Thesis Topic: Model-Based Small Area Estimation with an Application in Agriculture

Research Advisor: Prof. Partha Lahiri

Sep. 1997-Jul. 2000

M.S., China Agricultural University, China

Major: Animal Genetics and Breeding

Research Advisor: Prof. Changxin Wu (Academician of the Chinese Academy of Sciences)

Sep. 1993-Jul. 1997

B.S., Beijing Institute of Technology, China

Major: Computer Science

(e) OTHER EMPLOYMENT

Associate Professor, Joint Program in Survey Methodology, University of Maryland at College Park	College Park, MD, 08/2012 – 08/2019
Assistant Professor, Department of Mathematics, University of Texas at Arlington	Arlington, TX, 09/2009 – 08/2012
Visiting Assistant Professor, Department of Mathematics, University of Texas at Arlington	Arlington, TX, 09/2008 – 08/2009
Adjunct Faculty, Department of Clinical Science, University of Texas Southwestern Medical Center	Dallas, TX, 09/2010 – 01/2013
Mathematical Statistician, Office of Research and Methodology, National Center for Health Statistics	Hyattsville, MD, 07/2003 – 08/2005
Research and Teaching Assistant, Joint Program in Survey Methodology, University of Maryland at College Park	College Park, MD, 01/2003 – 07/2003
Mathematical Statistician, WESTAT	Rockville, MD, 08/2002-12/2002
Research and Teaching Assistant, Department of Mathematics and Statistics, University of Nebraska at Lincoln	Lincoln, NE, 08/2001-08/2002

2. **RESEARCH, SCHOLARLY AND CREATIVE ACTIVITIES**

(a) ARTICLES IN PEER REVIEWED JOURNALS

Summary: 52 journal papers (*paper with students)

Methodological Paper

- 1) C. Kern, Y. Li, L. Wang* (2020). Boosted Kernel Weighting – Using Statistical Learning to Improve Inference From Nonprobability Samples, *Journal of Survey Statistics and Methodology*, 10.1093/jssam/smaa028.
- 2) L. Wang*, B.I. Graubard, H.A. Katki, Y. Li (2020). Improving external validity of epidemiologic cohort analyses: a kernel weighting approach, *Journal of Royal Statistical Society A*, **183**, 1293-311..
- 3) L. Wang*, D.Y. Lin*, Y. Li (2020). Exploiting gene-environment Independence in haplotype-based inference for population-based case-control studies with complex sampling, *Statistics in Medicine* **39**, 57-69.
- 4) M. Yu, Y. Li, M. Qiu* (2019). Statistical inference of the relative concentration index for complex surveys. (Co-first Authors with M. Yu), *Statistics in Medicine*, **38**, 4083-95.
- 5) M. Yu, B. Liu, Y. Li, J. Zou, N. Breen (2019). Statistical inferences of extended concentration indices for directly standardized rates. *Statistics in Medicine*, **38**, 62-73.
- 6) Y. Li, M. Yu, J. Zhang* (2018). Statistical inference of health disparity indices for complex surveys. *American Journal of Epidemiology*, **187**, 2460-9.
- 7) Y. Li, P. Lahiri (2018), A simple adaptation of variable selection software for regression models to select variables in nested error regression models. *Sankhya B*. DO - 10.1007/s13571-018-0161-6.

- 8) Y. Li, T. Xiao*, D. Liao*, M.T. Lee (2018), Using threshold regression to analyze survival data from complex surveys: with application to mortality linked NHANES III phase II genetic data. *Statistics in Medicine*, **37**, 1162-77.
- 9) L.X. Wang*, B.I. Graubard, Y. Li (2016), A composite likelihood approach in testing for Hardy Weinberg Equilibrium using family-based genetic survey data. *Statistics in Medicine*, **35**, 5040-50. DOI: 10.1002/sim.7044.
- 10) D.Y. Lin*, L.X. Wang*, Y. Li (2016), Haplotype-based statistical inference for population-based case-control and cross-sectional studies with complex sample designs. *Journal of Survey Statistics and Methodology*, **4**, 188-214. DOI: 10.1093/jssam/smv040.
- 11) Y. Li, O. Panagiotou, A. Black, D. Liao*, S. Wacholder (2016), Multivariate piecewise exponential survival modeling. *Biometrics*, **72**, 546-53. DOI: 10.1111/biom.12435.
- 12) Y. Li, M. Safaeian, H. Robbins, B.I. Graubard (2015), Logistic analysis of epidemiologic and survey studies with augmentation sampling involving re-stratification and population expansion. *Biostatistics*, **16**, 169-78.
- 13) Y. Li, B.I. Graubard, P. Huang*, J.L. Gastwirth (2015), Extension of the Peter-Belson method to estimate health disparities among multiple groups using logistic regression with survey data. *Statistics in Medicine*, **34**, 595-612.
- 14) D. She*, H. Zhang, Y. Li, B.I. Graubard, Z. Li (2014), Family based association study with complex survey data, *Statistics and Its Interface*, **7**, 167-76.
- 15) Y. Li (2013), A comparison of tests for Hardy-Weinberg equilibrium in national genetic household surveys. *BMC Genetics*, **14**, 14.
- 16) Y. Li and B.I. Graubard (2012), Profile semi-parametric maximum likelihood estimation of gene-environment interaction using population-based case-control study with probability sampling. *Biostatistics*, **13**, 711-23.
- 17) Y. Li, M.H. Gail, D.L. Preston, B.I. Graubard, and J.H. Lubin (2012), Piecewise exponential survival times and analysis of case-cohort data. *Statistics in Medicine*, **31**, 1361-8.
- 18) H.A. Katki, Y. Li, and P.E. Castle (2012), Estimating the agreement and diagnostic accuracy of two diagnostic tests when one test is conducted on only a subsample of specimens. *Statistics in Medicine*, **31**, 436-48.
- 19) C.P. Han and Y. Li (2011), Regression analysis with block missing values and variables selection. *Pakistan Journal of Statistics and Operation Research*, **7**, 2.
- 20) Y. Li, Z. Li and B.I. Graubard (2011), Testing for Hardy Weinberg equilibrium in national household surveys that collect family-based genetic data. *Annals of Human Genetics*, **75**, 732-41.
- 21) Y. Li, B.I. Graubard, and R. DiGaetano (2010), Weighting methods for population-based case-control study. *Journal of Royal Statistical Society C*, **60**, 165–185.
- 22) N. Chatterjee and Y. Li (2010), Inference in semi-parametric regression models under partial questionnaire design and non-monotone missing data. *Journal of the American Statistical Association, Theory and Methods*, **105**, 787-797.
- 23) Y. Li, B.I. Graubard, and E.L. Korn (2010), Application of nonparametric quantile regression to body mass percentile curves from survey data. *Statistics in Medicine*, **29**, 558-572.

- 24) D. She, Y. Li, H. Zhang, B.I. Graubard, and Z. Li (2010), Trend test for genetic association population-based case-control study with complex survey data. *Biostatistics*, **11**, 48-56.
- 25) P. Lahiri and Y. Li (2009), A new alternative to the standard F test in regression analysis for clustered data. *Statistical Planning and Inference*, **139**, 3430-41.
- 26) Y. Li and B.I. Graubard (2009), Testing Hardy-Weinberg equilibrium and homogeneity of Hardy-Weinberg disequilibrium using complex survey data. *Biometrics*, **65**, 1096-104.
- 27) Y. Li (2008), Generalized regression estimators of a finite population total using the Box-Cox technique. *Survey Methodology*, **34**, 79-89.
- 28) Y. Li and P. Lahiri (2007), Robust model-based predictor of the finite population totals. *Journal of the American Statistical Association, Theory and Methods*, **102**, 664-73.

Collaborative Paper

- 29) M.A. Garza, S.C. Quinn, Y. Li, L. Assini-Meytin, E.T. Casper, C.S. Fryer, J.B. III, N.A. Brown, K.H. Kim, S.B. Thomas (2017). The influence of race and ethnicity on becoming a human subject: Factors associated with participation in research. *Contemporary Clinical Trials Communications*, **7**, 57-63.
- 30) L. Ke, R.S. Mason, E. Mpofu, J.L. Vingren, Y. Li, B.I. Graubard, K. Brock (2017). Hypertension and other cardiovascular risk factors are associated with vitamin D deficiency in an urban Chinese population: A short report. *Journal of Steroid Biochemistry and Molecular Biology*, **173**, 286-91.
- 31) F.M. Perna, K. Coa, R.P. Troiano, G.H. Lawman, C.Y. Wang, Y. Li, R.P. Moser, J.T. Ciccolo, B.A. Comstock, W.J. Kraemer (2016), U.S. Population muscular grip-strength estimates from the national health and nutrition examination survey (NHANES) 2011-2012. *Journal of Strength and Conditioning Research*, **30**, 867-74.
- 32) L. Ke, R.S. Mason, E. Mpofu, M. Dibley, Y. Li, et al. (2015), Vitamin D and parathyroid hormone status in a representative population living in Macau, China. *Journal of Steroid Biochemistry and Molecular Biology*, **148**, 261-8. PubMed PMID: 25636721.
- 33) L. Ke, J. Ho, J. Feng, E. Mpofu, M. Dibley, Y. Li, X. Feng, F. Van, W. Lau, K.E. Brock (2015) Prevalence, awareness, treatment and control of hypertension in Macau: results from a cross-sectional epidemiological study in Macau, China. *American Journal of Hypertension*, **28**: 159-165.
- 34) M.S. Shiels, H.A. Katki, N.D. Freedman, M. Purdue, N. Wentzensen, B. Trabert, C.M. Kitahara, M. Furr, Y. Li, T.J. Kemp, J.J. Goedert, C.M. Chang, E. Engels, N.E. Caporaso, L.A. Pinto, A. Hildesheim, A. Chaturvedi (2014), Cigarette smoking and variations in systemic immune and inflammation markers. *Journal of National Cancer Institute*, **106**, dju294.
- 35) A.W. Hsing, E. Yeboah, R. Biritwum, Y. Tettey, A.M. De Marzo, A. Adjei, G.J. Netto, K. Yu, Y. Li, A.P. Chokkalingam, L.W. Chu, D. Chia, A. Partin, I.M. Tompson, S.M. Quraishi, S. Niwa, R. Tarone, R.N. Hoover (2014), High prevalence of screen-detected prostate cancer in West Africans: implications for racial disparity of prostate cancer. *Journal of Urology*, **192**, 730-35.
- 36) K. A. Lang Kuhs, C. Porras, J. Schiller, A. C. Rodriguez, M. Schiffman, P. Gonzalez, S. Wacholder, A. Ghosh, Y. Li, D. Lowy, et al. (2014), Effect of different human papillomavirus serological and DNA criteria on vaccine efficacy estimates. *American Journal of Epidemiology*, **180**, 599-607.
- 37) H. Robbins, Y. Li, C. Porras, A. Ghosh, A.C. Rodriguez, M. Schiffman, S. Wacholder, et al. (2014), Glutathione S-transferase L1 multiplex serology as a measure of cumulative infection with human papillomavirus. *BMC Infectious Diseases*, **14**, 120.

- 38) L. Ke, J. Ho, E. Mpofu, M. Dibley, X. Feng, F. Van, S. Leong, W. Lau, P. Lueng, C. Kowk, Y. Li, R.S. Mason, K.E. Brock (2014), Modifiable risk factors including sunlight exposure and fish consumption are associated with risk of hypertension in a large representative population from Macau. *Journal of Steroid Biochemistry and Molecular Biology*, **144** Pt A: 152-5.
- 39) C. Schairer, Y. Li, P. Frawley, B.I. Graubard, R.D. Wellman, D.S.M. Buist, K. Kerlikowske, T.L. Onega, W.F. Anderson, and D.L. Miglioretti (2013), Risk factors for Inflammatory breast cancer and other Invasive breast cancers. *Journal of National Cancer Institute*, **105**, 1373-84.
- 40) Y. Pan, K. Caldwell, Y. Li, S. Caudill, M. Mortensen, A. Makhmudov, R. Jones (2013), Smoothed urinary iodine percentiles for U.S. population and pregnant women, national health and nutrition examination survey, 2001-2010. *European Thyroid Journal*, **2**, 127-34.
- 41) S.J. Nyante, B.I. Graubard, Y. Li, G.M. McQuillan, E.A. Platz, S. Rohrmann, G. Bradwin, D.A. McGlynn (2012), Trends in sex hormone concentrations in U.S. males: 1988-1991 to 1999-2004. *International Journal of Anthology*, **35**, 456-66.
- 42) A.P. Chokkalingam, E.D. Yeboah, A. DeMarzo, G. Netto, K. Yu, R.B. Biritwum, Y. Tettey, A. Adjei, S. Jadallah, E. Platz, Y. Li, L.W. Chu, D. Chia, S. Niwa, A. Partin, I.M. Thompson, Claus Roehrborn, Robert N. Hoover, A.W. Hsing (2012), Prevalence of BPH and lower urinary tract symptoms in West Africans. *Prostate Cancer and Prostatic Diseases*, **15**, 170-6.
- 43) L.G. Borrud, K.M. Flegal, D.S. Freedman, Y. Li, C.L. Ogden (2011), Smoothed percentage body fat percentiles: US children and adolescents, 1999-2004. *National Health Statistics Reports*, **43**, 1-7. (lightly reviewed)
- 44) D.M. van Bommel, Y. Li, J. McLean, M.H. Chang, N.F. Dowling, B.I. Graubard, P. Rajaraman (2011), Blood lead levels, ALAD gene polymorphisms, and mortality. *Epidemiology*, **22**, 273-8.
- 45) S. Goyal, Y.T. Kim, Y. Li and S.M. Iqbal (2010), Active and biomimetic nanofilters for selective protein separation. *Biomed. Microdevices*, **12**, 317-324.
- 46) C.L. Yu, Y. Li, D.M. Freedman, M.S. Linet, R.K. Kwok, B.H. Alexander, B.K. Armstrong, and T.R. Fears (2009), Assessment of lifetime cumulative sun exposure using a self-administered questionnaire: reliability of two approaches. *Cancer Epidemiology, Biomarkers & Prevention*, **18**, 464-71.
- 47) V.M. Chia, Y. Li, S.M. Quraishi, B.I. Graubard, J.D. Figueroa, J.P. Weber, S.J. Chanock, M.V. Rubertone, R.L. Erickson, K.A. McGlynn (2009), Effect modification of endocrine disruptors and testicular germ cell tumor risk by hormone-metabolizing genes. *International Journal of Andrology*, **33**, 588-96.
- 48) V.M. Chia, Y. Li, L.R. Goldin, B.I. Graubard, M.H. Greene, L. Korde, M.V. Rubertone, R.L. Erickson, K.A. McGlynn (2009), Risk of cancer in first- and second-degree relatives of testicular germ cell tumor cases and controls. *International Journal of Cancer*, **124**, 952-7.
- 49) L.A. Anderson, C. Lauria, N. Romano, E.E. Brown, D. Whitby, B.I. Graubard, Y. Li, A. Messina, L. Gaffa, F. Vitale, J.J. Goedert (2008), Risk factors for classical Kaposi Sarcoma in a population-based case-control study in Sicily. *Cancer Epidemiology, Biomarkers & Prevention*, **17**, 3435-43.
- 50) L.E. Kelemen, S.S. Wang, U. Lim, W. Cozen, M. Schenk, P. Hartge, Y. Li, N. Rothman, S. Davis, S.J. Chanock, M.H. Ward, J.R. Cerhan (2008), Vegetables and antioxidant-related nutrients, genetic susceptibility and non-Hodgkin lymphoma risk. *Cancer Causes and Control*, **19**, 491-503.
- 51) L.A. Anderson, Y. Li, B.I. Graubard, D. Whitby, G. Mbisa, S. Tan, J.J. Goedert, E.A. Engels (2008), Human herpesvirus 8 infection among children and adolescents in the United States. *Pediatric Infectious Diseases Journal*, **27**, 661-664.

- 52) M.P. Purdue, P. Hartge, S. Davis, J.R. Cerhan, J.S. Colt, W. Cozen, R.K. Severson, Y. Li, S.J. Chanock, N. Rothman, S.S. Wang (2007), Sun exposure, vitamin D receptor gene polymorphisms and risk of non-Hodgkin lymphoma. *Cancer Causes and Control*, **18**, 989-99.

PAPERS UNDER REVIEW

- 1) Wang L.*, B.I. Graubard, Hormuzd AK, Li Y. (2020). Efficient and Robust Propensity-Score-Based Methods for Population Inference using Epidemiologic Cohorts; <https://arxiv.org/abs/2011.14850>.
- 2) Wang L.*, Valliant R., Li Y. (2020). Adjusted Logistic Propensity Weighting Methods for Population Inference using Nonprobability Volunteer-Based Epidemiologic Cohorts; <https://arXiv:2007.02476v2>

BOOK REVIEW

- 1) Y. Li (2014), "When to Use What Research Design," *American Statistician*.

(b) PROCEEDINGS PAPERS AND TECHNICAL REPORTS

- 1) Y. Li and P.D. Williams (2005), A new multiple-bootstrap-datasets presentation method for confidentiality protection. *Proceeding of Survey Research Methods Section, American Statistical Association*, pp. 1306-33.
- 2) Y. Li, C. Lynch, and I. Shimizu (2004), Estimation of imputed variance for National Ambulatory Medical Care Survey. *Proceedings of Survey Research Methods Section, American Statistical Association*, pp. 3883-88.
- 3) Y. Li (2002), Estimation of crop yield for small areas – an application of EBLUP method. *Tech. Report*, University of Nebraska-Lincoln.
- 4) Y. Li, H.-W. Deng and R.R. Recker (2001), LOD score exclusion analyses for candidate QTLs using random population samples. *Proceedings of the 32nd Annual Midwest Student Biomedical Research Forum*, Page: P-19.
- 5) Y. Li and H.W. Deng (2001), Population admixture may mask, change or reverse genetic effects of genes underlying complex traits. *Second Annual Meeting of the Great Plain States Society for Molecular Biology and Genetics*, Abstract 31.

(c) CONFERENCES, WORKSHOPS, AND TALKS

- Abstract and Invited Talk: A Kernel Weighting Approach to Improve Population Representativeness for Estimating Prevalence of Risk-factors and Diseases, 08/2019, *Conference on Current Trends in Survey Statistics*, National University of Singapore, Singapore.
- Invited Talk: Using Surveys to Improve the Representativeness of Nonprobability Samples in Epidemiologic Studies, *National Institute of Statistical Sciences (NISS) workshop*, Washington DC, 03/2019
- Abstract and Invited Talk: A Kernel Weighting Approach to Improve Population Representativeness of Epidemiological Cohort in the Analysis, 06/2018, *Conference on Small Area Estimation and Other Topics of Current Interest in Surveys, Official Statistics, and General Statistics*, Shanghai, China.
- Abstract and Invited Talk: Genetic Analyses using Family-Based Survey Data, 05/2017, *Contemporary Theory and Practice of Survey Sampling: A Celebration of Research Contributions of J.N.K. Rao*, Kunming, China.

- Abstract and Invited Talk: Using Threshold Regression to Analyze Survival Data from Complex Surveys: With Application to Mortality Linked NHANES III Phase II Genetic Data, 12/2016, *The 10th ICSA International Conference*, Shanghai, China.
- Abstract and Invited Talk: Genetic Analyses Using Family-Based Survey Data, 11/15/2016, *National Center for Health Statistics*, Hyattsville, MD.
- Abstract and Speed Poster: Extension of the Peters-Belson method to estimate health disparities among multiple groups using logistic regression with survey data, 08/2016, *Joint Statistical Meeting*, Chicago, IL.
- Abstract and Contributed Talk: Variable selections for the nested error regression models with applications in small area estimation, 08/2016, *SAE Conference*, Maastricht, Netherlands.
- Abstract and Invited Talk: Genetic analyses using family-based survey data, 07/2016, *Academy of Mathematics and Systems Science, Beijing Academy of Sciences*, Beijing, China.
- Abstract and Invited Talk: Extension of the Peters-Belson method to estimate health disparities among multiple groups using logistic regression with survey data, 07/2016, *the Forth International Symposium on Biostatistics*, Shanghai, China.
- Abstract and Topic Contributed Talk: Genetic analyses using family-based survey data, 06/2016, *the 4th Institute of Mathematical Statistics Asia Pacific Rim Meeting*, Hong Kong, China.
- Abstract and Invited Talk: Extension of the Peter-Belson method to estimate health disparities among multiple groups using logistic regression with survey data, 03/2016, *ENAR*, Austin, TX.
- Abstract and Invited Talk: Genetic data analysis using family-based complex survey data, *the 4th Baltic-Nordic Conference on Survey Statistics*, 08/2015, Helsinki, Finland.
- Abstract and Topic Contributed Talk: A composite likelihood approach in testing for Hardy Weinberg equilibrium using family-based genetic survey data, 08/2015, *Joint Statistical Meeting*, Seattle, WA.
- Abstract and Invited Talk: Haplotype-based statistical inference for population-based case-control and cross-sectional studies with complex sampling, 12/2014, *International Conference by the Institute of Applied Statistics*, Colombo, Sri Lanka.
- Abstract and Invited Talk: Haplotype-based statistical inference for population-based case-control and cross-sectional studies with complex sampling, 05/2014, *Conference on Frontiers of Hierarchical Modeling in Observational Studies, Complex Surveys and Big Data*, College Park, MD.
- Abstract and Invited Talk: Application of nonparametric percentile regression of body mass index percentile curves from survey data, 08/2013, *The 59th World Statistics Congress*, Hong Kong, China.
- Abstract and Topic Contributed Talk: Testing Hardy-Weinberg equilibrium for national genetic household surveys, 03/2013, *ENAR*, Orlando, FL.
- Abstract and Invited Talk: Testing Hardy-Weinberg equilibrium for national genetic household surveys, 02/2013, *Washington Statistical Society*, Washington DC.
- Abstract and Topic Contributed Talk: Testing Hardy-Weinberg equilibrium for national genetic household surveys, 08/2012, *Joint Statistical Meeting*, San Diego, CA.

- Abstract and Invited Talk: Weighting methods and pseudo-semiparametric inference for population-based case-control studies with complex sampling, 02/2012, *Joint Program in Survey Methodology, University of Maryland at College Park, MD.*
- Abstract and Invited Talk: Weighting methods and pseudo-semiparametric inference for population-based case-control studies with complex sampling, 02/2012, *Biostatistics Division at School of Public Health, University of Minnesota, Minneapolis, MN.*
- Abstract and Invited Talk: Weighting methods and pseudo-semiparametric inference for population-based case-control studies with complex sampling, 02/2012, *Department of Statistics, George Washington University, Washington DC.*
- Abstract and Invited Talk: Weighting methods and pseudo-semiparametric inference for population-based case-control studies with complex sampling, 02/2012, *Department of Mathematical Sciences, University of Texas at Dallas, Dallas, TX.*
- Abstract and Invited Talk: Weighting methods and pseudo-semiparametric inference for population-based case-control studies with complex sampling, 02/2012, *The Biostatistics & Bioinformatics Branch, Division of Epidemiology, Statistics, and Prevention Research of the Eunice Kennedy Shriver, National Institute of Child Health and Human Development, NIH, Rockville, MD.*
- Abstract and Invited Talk: Weighting methods for population-based case-control studies with complex sampling, 02/2012, *University of Texas Chapter, Mathematical Association of America, Arlington, TX.*
- Abstract and Invited Talk: Weighting methods and pseudo-semiparametric inference for population-based case-control studies with complex sampling, 03/2012, *Department of Biostatistics and Computational Biology, Susan Smith Center for Women's Cancers, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA.*
- Abstract and Invited Talk: Semi-parametric pseudo-maximum-likelihood estimation exploiting gene-environment independence for population-based case-control studies with complex sampling, 07/2011, *IMS-China International Conference on Statistics and Probability, XiAn, China.*
- Abstract and Invited Talk: Semi-parametric pseudo-maximum-likelihood estimation exploiting gene-environment independence for population-based case-control studies with complex sampling, 06/2011, *International Chinese Statistical Association 2011 Applied Statistics Symposium, New York.*
- Abstract and Invited Talk: Semi-parametric pseudo-maximum-likelihood estimation exploiting gene-environment independence for population-based case-control studies with complex sampling, 10/2010, *Department of Biology, University of Texas at Arlington, Arlington, TX.*
- Abstract and Invited Talk: Semi-parametric pseudo-maximum-likelihood estimation exploiting gene-environment independence for population-based case-control studies with complex sampling, 08/2010, *Joint Statistical Meeting, Vancouver, British Columbia, Canada.*
- Abstract and Invited Talk: Weighting methods for population-based case-control studies with complex sampling, 08/2009, *Joint Statistical Meeting, Washington D.C.*
- Abstract and Contributed Talk: Application of nonparametric percentile regression of body mass index percentile curves from survey data, 03/2009, *Conference by Eastern North American Region/International Biometric Society, San Antonio, TX.*
- Abstract and Invited Talk: Testing Hardy-Weinberg equilibrium (HWE) and homogeneity of Hardy-Weinberg disequilibrium (HHWD) using complex survey data, 08/2008, *Joint Statistical Meeting, Denver, Colorado.*

- Abstract and Contributed Talk: Application of nonparametric percentile regression of body mass index percentile curves from survey data, 06/2008, *Western North American Region/Institute Mathematical Statistical meeting*, Davis, CA.
- Abstract and Invited Talk: Application of nonparametric percentile regression of body mass index percentile curves from survey data, 02/2008, *University of Texas Southwestern Medical Center*, Dallas, TX.
- Abstract and Invited Talk: Testing Hardy-Weinberg Equilibrium (HWE) and Homogeneity of Hardy-Weinberg Disequilibrium (HHWD) Using Complex Survey Data, 02/2008, *University of Texas*, Arlington, TX.
- Abstract and Invited Talk: Robust model-based and model-assisted predictors of the finite population total, 08/2006, *Joint Statistical Meeting*, Seattle, WA.
- Abstract and Contributed Talk: A new multiple-bootstrap-datasets presentation method for confidentiality protection, 08/2005, *Joint Statistical Meeting*, Minneapolis, MN.
- Abstract and Invited Talk: A new multiple-bootstrap-datasets presentation method for confidentiality protection, 07/2005, *National Center for Health Statistics*, Hyattsville, MD.
- Abstract and Contributed Talk: Estimation of imputed variance for National Ambulatory Medical Care Survey, 08/2004, *Joint Statistical Meeting*, Toronto, Canada.
- Abstract and Invited Talk: Estimation of imputed variance for National Ambulatory Medical Care Survey, 07/2004, *National Center for Health Statistics*, Hyattsville, MD.

(d) SPONSORED RESEARCH

Summary: 1. Total: about \$3.69M as a PI or Co-PI; my total share: about \$600K

2. Sole PI of three NIH R03 or equivalent grants, Co-PI of one NIH R01, and Co-PI of one DARPA.

PI	\$154,500	09/2020	–
“Improving Population Representativeness of the Inference from Non-Probability Sample Analysis,” NIH-R03 CA252782, National Institute of Health			
PI	\$104,489	08/2020	–
“Variable selection for propensity score models for estimating population means using samples subject to selection bias,” 20-IPA-20-06718, Centers for Disease Control and Prevention.			
PI	\$25,000	09/2019	–
“Improving External Validity of Epidemiologic Analyses,” 75N91019P00701, National Institute of Health			
PI	\$15,000	09/2017	–
“Algorithms for Inferences of Health Disparity Summary Measures Bases On Complex Surveys,” HHSN261201700511P, National Institute of Health			
PI	\$24,990	09/2016	–
“Estimation of Health Disparity Summary Measures for Complex Survey Samples,” HHSN261201600684P, National Institute of Health			
PI (with Co-PI Partha Lahiri at JPSM)	\$10,000	04/2014	–
“Conference on ‘Frontiers of Hierarchical Modeling in Observational Studies,”			
		05/2015	

Complex Surveys and Big Data', National Science Foundation

PI	\$140,325	09/2013	–
"Semiparametric Inference for Case-Control Studies with Complex Sampling," NIH-R03 CA1710649, National Institute of Health			
PI	\$133,741	01/2014	–
"Genetic Association Studies in National Household Surveys," Research Fellowship–American Statistical Association/National Center for Health Statistics			
PI (with Co-I Mei-Ling Lee in Department of Epidemiology and Biostatistics),	\$ 20,000	08/2013	–
"Incorporating Threshold Regression in Survey Methodology," SEED-University of Maryland at College Park			
PI	\$142,230	09/2011	–
"SNP-based pseudo-semiparametric inference for the case-control studies," NIH-U01 CA159424, National Institute of Health			
Co-PI (with PI Romero in Bioengineering at the University of Texas at Arlington)	\$2,195,054 (My Share: \$60,261)	11/2010	–
"Cellular and molecular contributions to signal instability in peripheral regenerative neurointerfaces," DARPA-BAA-10-32, Department of Energy, Defense Advanced Research Projects Agency			
Co-PI (with PI Chen in Biostatistics at the University of Pennsylvania)	\$1,000,000 (My Share: \$35,091)	04/2010	–
"Statistical Methods in Genetic Epidemiology Research," NIH-5R01ES016626, National Institute of Health			
PI	\$7,100	05/2009	–
"Statistical Analyses of Gene-Environment Interactions," Research Enhancement Program, The University of Texas at Arlington			

Research Proposal Under Review

PI \$147,853
 "Estimation of Absolute Risks and Mediation Effects from Time-to-Event Data Collected with Complex Sampling," NIH-CA219790

Collaborator (PI Mary Garza in Behavioral and Community Health at the University of Maryland at College Park) \$786,854
 "Elucidating the Role Family Health History Plays on CRC Risk and Screening," American Cancer Society #127666

Collaborator (PI Cynthia Baur in Health Literacy, School of Public Health at the University of Maryland at College Park)
 "Emergency Competitive Revisions for Social, Ethical, and Behavioral Implications (SEBI) Research on COVID-19 Testing among Underserved and/or Vulnerable Populations," NIH NOT-OD-20-119

3. TEACHING, MENTORING, AND ADVISING(a) COURSES TAUGHT

Summary: as an instructor, I have taught ten different courses at the University of Maryland at College Park, and nine courses at the University of Texas at Arlington.

- *Data Science for Social Science (undergraduate)*, **Instructor** (co-instructor Dr. Brian Kim). Class size – 25 students. Responsibilities: lectures/discussions, office hours, biweekly assignments.
- *Fundamentals of Inference (graduate)*, **Instructor** (co-instructor Dr. Michael Elliott and Dr. Trivellore Raghunathan in 2018, co-instructor Dr. Michael Elliott in 2019). Class size – 9 students (in 2018) and 24 students (in 2019). Responsibilities: lectures/discussions, office hours, biweekly homework, 1 midterm, 1 final.
- *Practical Tools for sampling and weighting (graduate)*, **Instructor**. Class size – 7 students. Responsibilities: lectures/discussions, office hours, weekly homework, 1 project, 1 final.
- *Federal Statistical System (undergraduate and graduate)*, **Instructor**. Class size – 15 students. Responsibilities: weekly homework, 2 essays.
- *Topics in Survey Methodology (graduate)*, **Instructor**. Class size – 8 students. Responsibilities: lectures, office hours, weekly homework, 2 essays and 1 oral presentation.
- *Statistical Methods for Analysis of Complex Samples in Public Health (graduate)*, **Instructor** (co-instructor Dr. Barry Graubard). Class size – 5 students. Responsibilities: lectures, office hours, biweekly homework, final.
- *Analysis of Complex Sample Data (graduate)*, **Instructor**. Class size – 17 students (in 2016) and 23 students (in 2017). Responsibilities: lectures, office hours, weekly homework, midterm, final.
- *Inference for Complex Surveys (graduate)*, **Instructor**. Class size – 8 students. Responsibilities: lectures, office hours, weekly homework, 2 exams.
- *Random/Nonrandom Experimental Design (graduate)*, **Instructor**. Class size – 15 students. Responsibilities: lectures, office hours, weekly homework, 2 exams.
- *Applied Sampling (graduate)*, **Instructor** (co-instructor Prof. Jim Lepkowski). Class size – 42 students. Responsibilities: lectures, office hours, weekly homework, project, 2 exams.
- *Statistical Inference (graduate)*, University of Texas at Arlington, **Instructor**. Class size – about 10 students. Responsibilities: lectures, office hours, 4 exams per semester.
- *Statistical Methods for Clinical Studies (graduate)*, University of Texas at Arlington, **Instructor**. Class size – 5 students. Responsibilities: lectures, office hours, 1 exam per semester.
- *Sample Surveys (graduate)*, University of Texas at Arlington, **Instructor**. Class size – about 10 students. Responsibilities: lectures, office hours, 3 exams per semester.
- *Mathematical Statistics I (graduate)*, University of Texas at Arlington, **Instructor**. Class size – 15 students. Responsibilities: lectures, office hours, 3 exams per semester.
- *Mathematical Statistics II (graduate)*, University of Texas at Arlington, **Instructor**. Class size – 6 students. Responsibilities: lectures, office hours, 3 exams per semester.
- *Statistical Theory for Research (graduate)*, University of Texas at Arlington, **Instructor**. Class size – 15 students. Responsibilities: lectures, office hours, 3 exams per semester.

- *Statistical Methods (undergraduate)*, University of Texas at Arlington, **Instructor**. Class size – 40 students. Responsibilities: lectures, office hours, 4 exams per semester.
- *Sampling Theory (graduate)*, University of Maryland at College Park, **Teaching assistant**. Class size – 15 students. Responsibilities: recitations, office hours, composing and grading quizzes, home works and exams.
- *Calculus I (undergraduate)*, University of Nebraska at Lincoln, **Teaching assistant**. Class size – 35 students. Responsibilities: recitations, office hours, composing and grading quizzes, homework and exams.

(b) ADVISING: RESEARCH DIRECTION

At University of Maryland, College Park

Supervised Doctoral and Master's Students

- Lingxiao Wang, Ph.D., 03/2020, Joint Program in Survey Methodology, UMD
Dissertation Title: "Bayesian Methods for Prediction of Survey Data Collection Parameters in Adaptive and Responsive Designs"
- Zhenyi Xue, Ph.D., 2015, Applied Mathematics & Statistics, and Scientific Computation, UMD (co-supervised with Prof. Partha Lahiri)
Dissertation Title: "Bayesian Methods to Estimation Inbreeding Coefficient for Single Nucleotide Polymorphism Collected from Complex Surveys"
- Daoying Lin, Ph.D., 2013, Department of Mathematics, University of Texas at Arlington
Dissertation Title: "Statistical Methods for Population-Based Case-Control Studies with Complex Sampling"
- Lingxiao Wang, M.S., 2015 Joint Program in Survey Methodology, UMD
Research Topic: A Composite Likelihood Approach in Testing for Hardy Weinberg Equilibrium using Family-Based Genetic Survey Data
- Pengyu Huang, M.S., 2014 Joint Program in Survey Methodology, UMD
Research Topic: Extension of the Peter-Belson method to estimate health disparities among multiple groups using logistic regression with survey data

Current Dissertation Exam Committee Membership

- Yixuan Pan, Member (Chair Lei Zhang, Civil and Environmental Engineering, UMD)
- Weiyi Zhou, DR (Chair Lei Zhang, Civil and Environmental Engineering, UMD)
- Alena Maze, Member (Chair Richard Valliant, Joint Program in Survey Methodology, UMD)

Completed Dissertation Exam Committee Membership

- William Waldron, Member (Chair Partha Lahiri, Joint Program in Survey Methodology, UMD), 10/2020
- Stephanie Coffey, Member (Chair Michael Elliott, Department of Biostatistics and Survey Research Center, University of Michigan at Ann Arbor), 09/2020
- Evan Olson, Member (Chair Hong Jiao, Department of Human Development and Quantitative Methodology, College of Education, UMD), 05/2020
- Yixin Ren, Dean's Representative (Chair Xin He, Department of Epidemiology and Biostatistics, School of Public Health; Co-advisor Paul Smith, Department of Mathematics, College of Computer, Mathematical and Natural Sciences, UMD), 11/2019
- Ji An, Dean's Representative (Chair Laura M. Stapleton, Department of Human Development and Quantitative Methodology, College of Education, UMD), 11/2019
- Afnan Al-Ibrahim, Member (Chair Robert Jackson, Department of Nutrition and Food Science, College of Agriculture and Natural Resources, UMD), 06/2018
- Ying Han, Member (Chair Partha Lahiri, Joint Program in Survey Methodology, UMD), 05/2018
- Dandan Liao, Dean's Representative (Chair Hong Jiao, Department of Human Development and Quantitative Methodology, College of Education, UMD), 03/2018

- Judith Law, Member (Chair Partha Lahiri, Joint Program in Survey Methodology, UMD), 01/2018
- Xuan Yao, Dean's Representative (Chair Eric Slud, Department of Mathematics, College of Computer, Mathematical and Natural Sciences, UMD), 01/2018
- Benjamin Reist, Member (Chair Richard Valliant, Joint Program in Survey Methodology, UMD), 06/2017

Comprehensive Exam Committee Membership

- Mark Wymer, Member (AMSC in Department of Mathematics, UMD), 10/2019
- Ali Rafei, Member (Program in Survey Methodology, University of Michigan-Ann Arbor), 07/2019
- Micha Fischer, Member (Program in Survey Methodology, University of Michigan-Ann Arbor), 05/2018
- Ji An, Member (Department of Human Development and Quantitative Methodology, College of Education), 01/2017
- Stephanie Coffey, Member (Joint Program in Survey Methodology, UMD), 05/2016
- Josh Langeland, Member (Joint Program in Survey Methodology, UMD), 05/2016
- Xia Li, Member (Department of Mathematics, UMD), 03/2016
- Sarah Vetting, Chair (Joint Program in Survey Methodology, UMD), 03/2015

Other Directed Research

- Meng Qiu, Master Student, Department of Human Development and Quantitative Methodology, College of Education, UMD, 2016-17 (coauthored in a manuscript submitted to a statistical journal)
- Dandan Liao, Master Student, Department of Human Development and Quantitative Methodology, College of Education, UMD, 2015-17 (coauthored in two papers appeared in *Statistics in Medicine* and *Biometrics*)
- Jonathan Zhang, Undergraduate, Computer Science, College of Computer, Mathematical & Natural Sciences, UMD, 2016-17 (coauthored in a manuscript submitted to a statistical journal)
- Neil Ghosh, Junior, Summer Research, Science, Math, and Computer Science House, Poolesville High School, Maryland, 05-08/2019

At University of Texas at Arlington

Supervised Master's Students at the University of Texas at Arlington, Texas

- William Kenney, M.S., 2012, Department of Mathematics
Thesis Title: "Genetic Association Test for National Household Surveys that Collect Family-Based Genetic Data"
- Daoying Lin, M.S., 2012, Department of Mathematics
Thesis Title: "Haplotype-Based Statistical Inference for Population-Based Case-Control Studies with Complex Sample Designs"
- Fred Tsai, M.S., 2011, Department of Mathematics
Thesis Title: "Comparisons among Hardy-Weinberg Equilibrium Tests for National Genetic Household Surveys"

Dissertation/Thesis Exam Committee Membership at the University of Texas at Arlington, Texas

- Thomas Seaquist, Dissertation committee (Chair Andrzej Korzeniowski in Mathematics)
- Katsuhiko Iwao, Dissertation committee member (Chair Shan Sun-Mitchell in Mathematics)
- Prince Albert Nfodzo, Dissertation committee (Chair Hyeok Choi in Civil Engineering)
- Juan Levesque, Dissertation committee (Chair James Grover in Biology)
- Darin Brezeale, Thesis committee (Chair Ren-Cang Li in Mathematics)
- Ashwin K. Satyal, Thesis committee (Chair Chien-Pai Han in Mathematics)
- Haowen Tang, Thesis committee (Chair Chien-Pai Han in Mathematics)

4. SERVICE

(a) EDITORSHIP, EDITORIAL BOARDS AND REVIEWING ACTIVITIES FOR JOURNALS

- i. Associate-Editor, *Sankhya B* (02/2014-12/2015)

- ii. Journal Reviewer for:
Journal of Survey Statistics and Methodology, Journal of American Statistical Association, Biometrics, Annals of Statistics, Statistics in Medicine, Medicine, Journal of Official Statistics, Human Heredity, American Journal of Epidemiology, Economics and Human Biology, International Journal of Data Mining and Bioinformatics, Briefings in Bioinformatics, Pakistan Journal of Statistics and Operation Research, BioMed Central Genetics, Statistical Applications in Genetics and Molecular Biology, Frontiers in Genetic Architecture, Statistical Methods in Medical Research.

(b) PROFESSIONAL SERVICE

i. OFFICES AND COMMITTEE MEMBERSHIPS IN PROFESSIONAL ORGANIZATIONS

- Member, Scientific Program Committee, Conference on Current Trends in Survey Statistics, National University of Singapore, Singapore, 2019
- Member, American Association for Public Opinion Research/American Statistical Association Data Falsification Task Force, 2016-2018
- Member, Design and Analysis Committee for the National Assessment of Educational Progress, Educational Testing Service and the National Center for Education Statistics, 2016-2017
- Program Chair, Survey Research and Methodology Section, American Statistical Association, 2016
- Program Chair Elect, Survey Research and Methodology Section, American Statistical Association, 2015
- Member, Statistical Partnerships Among Academe, Industry, and Government Committee, American Statistical Association, 2014-2016
- Member, Scientific Program Committee and Local Organization Committee, Conference on Frontiers of Hierarchical Modeling in Observational Studies, Complex Surveys and Big Data, 2013-2014
- Chair, Student/Young Investigator Travel Award and Paper Selection Committee, Conference on Frontiers of hierarchical Modeling in Observational Studies, Complex Surveys and Big Data, 2014.
- Member, American Statistical Association (ASA), 2004-present
- Member, Eastern North American Region (ENAR) / International Biometric Society, 2009-present

ii. OTHER NON-UNIVERSITY COMMITTEES, COMMISSIONS, PANELS, ETC.

- Chair, invited Session on *Data Integration in Bio and Health Statistics*, Conference on Current Trends in Survey Statistics, National University of Singapore, Singapore, 08/2019
- Organizer, National Institute of Statistical Sciences (NISS) workshop on *Improving the Representativeness of Nonprobability Samples*, Washington DC, 03/2019
- Organizer and Chair, Invited Session on *Using Surveys to Improve the Representativeness of Nonprobability Samples in Epidemiologic Studies*, Joint Statistical Meeting, 2018
- Judge, Poster Award Selection Committee, Survey Research Method Section, American Statistical Association, 2015
- Organizer, Topic-Contributed Session on *Calibration and Semiparametric Techniques in Surveys and Biostatistics*, Joint Statistical Meeting, 2015
- Chair, Invited/Plenary Session, Conference on Frontiers of Hierarchical Modeling in Observational Studies, Complex Surveys and Big Data, 2014
- Nominee, Biometrics Section Representative to ASA Council of Sections Representative 2013-15
- Chair, Invited Session on *Complex Survey Methodology and Application*, ENAR 2013
- Chair, Contributed Session on *Epidemiology Designs Based on Complex Survey Data*, Joint Statistical Meeting 2010

iii. OTHER UNPAID SERVICES TO LOCAL, STATE AND FEDERAL AGENCIES

- Special Volunteer, Division of Health and Nutrition Examination Surveys, National Center for Health Statistics, Center for Disease Control, 10/2014-06/2015
- Research Consultant, Biostatistics Branch, Division of Cancer and Epidemiology and Genetics, National Cancer institute, NIH, 04/2011-present

(c) CAMPUS SERVICE

i. DEPARTMENT

- Member, Center for Advances in Data and Measurement (CADM) executive director search committee, 09/2019-Present
- Member, Teaching assignment Committee, 08/2019-Present
- Member, JPSM Internal Program Review Committee, 09/2019-Present
- Director of Graduate Studies, 09/2014-Present
- Writer and Grader, Qualifying Exam, 09/2016-Present
- Member, *Graduate Admission Committee*, 01/2013-Present
- Organizer, *Research Seminar Series*, 09/2013-10/2017
- Member, Postdoc Search Committee, 04/2014
- Member, Graduate Affairs Committee, 08/2010-12/2010 & 08/2011-08/2012
- Leader, Preliminary-B Exam in Statistics, 08/2010-08/2012
- Grader, Preliminary-A Exam for Linear Algebra, 08/2010
- Facilitator, "Calculus Bowl 2011" competition among 28 Dallas/Fort Worth high schools, 2011

ii. COLLEGE OF BEHAVIOR AND SOCIAL SCIENCES

- Chair, Review Committee for the Charles A. Caramello Distinguished Dissertation Award, 02/2019
- Member, BSOS Dean's All S.T.A.R. Review/Selection Committee, 04/2018
- Member, Outstanding Graduate Advisor Award Committee, 03/2015
- Department Representative, College Council, 09/2013-12/2014

iii. UNIVERSITY

- *University Senate*, 05/2020-04/2023
- *University Senate*, 05/2014-04/2017

5. AWARDS AND HONORS

- 2016 Poster Award, Survey Research and Methods Section, Joint Statistical Meeting
- 2009 Travel Award to Eastern North American Region Workshop for Junior Investigators
- 2007 Division of Cancer Epidemiology and Genetics (DCEG) Fellows Award for Research Excellence
- 2006 One of the Best Six Submitted Papers for European Association of Methodology Award
- 2006 Washington Statistical Society Outstanding Graduate Student Award
- 2006 Travel Award to the Ninth Meeting of New Researchers in Statistics and Probability
- 2006 Winner of 2006 Joint Statistical Meeting Student Paper Competition
- 2005 Travel Award to Joint Statistical Meeting