

## Curriculum Vitae

Lei Sun

(Last updated May 24, 2018)

Department of Statistical Sciences	100 St. George Street
Faculty of Arts and Science	Toronto, ON M5S 3G3, Canada
Division of Biostatistics	P: (416) 978-4327, F: (416) 978-5133
Dalla Lana School of Public Health	sun@utstat.toronto.edu
University of Toronto	<a href="http://www.utstat.toronto.edu/sun">http://www.utstat.toronto.edu/sun</a>

### Degrees

- 08/2001 Ph.D. Statistics, University of Chicago, USA  
*Two statistical problems in human genetics: I. Detection of pedigree errors;  
II. Identification of polymorphisms that explain a linkage result*  
Supervisor: Mary Sara McPeck
- 07/1995 B.S. Mathematics, Fudan University, Shanghai, China

### Academic Positions

- 07/2014-present Full Professor Department of Statistical Sciences  
Faculty of Arts and Science  
Division of Biostatistics  
Dalla Lana School of Public Health  
University of Toronto  
(75% FTE in Statistics and 25% in Biostatistics)
- 07/2007-06/2014 Associate Professor Division of Biostatistics  
Dalla Lana School of Public Health  
University of Toronto
- 07/2005-06/2014 Adjunct Professor Department of Statistical Sciences  
Faculty of Arts and Science  
University of Toronto
- 09/2001-06/2007 Assistant Professor Division of Biostatistics  
Dalla Lana School of Public Health  
University of Toronto

### Interruptions/Delays

- 01/2007-08/2007 Maternity Leave (30 weeks)
- 04/2004-11/2004 Maternity Leave (30 weeks)

## Awards

- 2018 [NSERC Discovery Accelerator Supplements \(DAS\)](#) Program Award,  
*provides substantial and timely additional resources to accelerate progress and  
maximize the impact of superior research programs*
- 2017 [CRM-SSC prize in Statistics](#), award by the Centre de  
recherches mathématiques (CRM) and the Statistical Society of Canada (SSC)  
*in recognition of a statistical scientist's professional accomplishments in research  
during the first fifteen years after having received a doctorate*
- 2017 Dean's Excellence Award, Faculty of Arts and Science 5% Merit Pool  
for outstanding performance across the Faculty, University of Toronto
- 2015 Dean's Excellence Award, Faculty of Arts and Science 5% Merit Pool  
for outstanding performance across the Faculty, University of Toronto
- 2014 Dean's Excellence Award, Faculty of Arts and Science 5% Merit Pool  
for outstanding performance across the Faculty, University of Toronto
- 2006 Dean's Award, Faculty of Medicine's 5% Merit Pool for academic achievement  
in research, teaching and service, University of Toronto
- 2004 Dean's Award, Faculty of Medicine's 5% Merit Pool for academic achievement  
in research, teaching and service, University of Toronto
- 1996-2001 University of Chicago Fellowship
- 1996-1997 Paul Meier Fellowship, Department of Statistics, University of Chicago

## Research Grants

- 2018-2023 Natural Sciences and Research Council of Canada (NSERC)  
*Robust allele-based association analyses of complex genetic data*  
PI: **Lei Sun** (Individual)  
\$44,000/year; \$220,000 in total
- 2018-2021 Natural Sciences and Research Council of Canada (NSERC)  
Discovery Accelerator Supplements (DAS)  
*Robust allele-based association analyses of complex genetic data*  
PI: **Lei Sun** (Individual)  
\$40,000/year; \$120,000 in total
- 2016-2019 Cystic Fibrosis Canada (CFC)  
*Genetic epidemiology of cystic fibrosis*  
PI: Lisa Strug  
5 co-Is: Tanja Gonska, Felix Ratjen, Peter Durie, Johanna Rommens, **Lei Sun**  
\$120,571/year; \$361,713 in total

- 2014-2019 Canadian Institutes of Health Research (CIHR)  
*Statistical methods and computational tools for next-generation genetic studies of complex traits*  
 PI: **Lei Sun**  
 2 co-Is: Andrew Paterson, Lisa Strug  
 \$133,528/year; \$667,641 in total
- 2013-2018 Natural Sciences and Research Council of Canada (NSERC)  
*Joint analyses of multiple genetic variants and multiple trait phenotypes*  
 PI: **Lei Sun** (Individual)  
 \$23,000/year; \$115,000 in total
- 2012-2017 Canadian Institutes of Health Research (CIHR)  
*Modification of cystic fibrosis phenotypes in the pancreas and gut*  
 PI: Lisa Strug  
 3 co-Is: Peter Durie, Johanna Rommens, **Lei Sun**  
 \$214,079/year; \$1,070,395 in total
- 2012-2015 Juvenile Diabetes Research Foundation (JDRF)  
*Genetics of the decline in glomerular filtration rate in type 1 diabetes*  
 PI: Andrew Paterson  
 6 co-Is: Ian de Boer, Andrew Boright, Shelley Bull, Barbara Klein, Ronald Klein, **Lei Sun**  
 US \$485,827/year; US \$1,457,483 in total
- 2011-2012 McLaughlin Centre Accelerator Grant in Genomic Medicine  
*Developing novel hypothesis driven sequence analysis methodology towards improved therapeutics for cystic fibrosis*  
 PI: **Lei Sun**  
 3 co-Is: Peter Durie, Johanna Rommens, Lisa Strug  
 \$75,000 in total  
*One of the “10 Big Stories in Personalized Medicine” by the McLaughlin Centre*
- 2011-2012 McLaughlin Centre Accelerator Grant in Genomic Medicine  
*Identification of genes for severe early diabetic nephropathy*  
 PI: Andrew Paterson  
 3 co-Is: Andrew Boright, Shelley Bull, **Lei Sun**  
 \$75,000 in total
- 2008-2013 Natural Sciences and Research Council of Canada (NSERC)  
*Statistical methods for complex genetic data*  
 PI: **Lei Sun** (Individual)  
 \$14,000/year; \$70,000 in total
- 2007-2012 Canadian Institutes of Health Research (CIHR)  
*Design and analysis of genome-wide studies of complex diseases and traits*  
 co-PI: **Lei Sun**; 2 other co-PIs: Shelley Bull, Radu V. Craiu  
 \$108,606/year; \$543,030 in total

- 2006-2009 National Institute of Health (NIH)  
*Genome-wide association of common alleles with long-term diabetic complications*  
 PI: Andrew Paterson  
 5 co-Is: Andrew Boright, Shelley Bull, Patricia Cleary, John Lachin, **Lei Sun**  
 US \$1,391,927/year; US \$4,175,781 in total
- 2003-2006 Canadian Institutes of Health Research (CIHR)  
*Statistical methods to improve the reliability of results from genome-wide studies of complex disease and quantitative traits*  
 co-PI: **Lei Sun**; 1 other co-PI: Shelley Bull  
 \$87,000/year \$261,000 in total
- 2002-2007 Natural Sciences and Research Council of Canada (NSERC)  
*Mapping genetic variants for complex disease via statistical methods for positional cloning*  
 PI: **Lei Sun** (Individual)  
 \$14,500/year; \$72,500 in total
- 2001-2002 University of Toronto Connaught Start-up Grant  
 PI: **Lei Sun** (Individual)  
 \$10,000 in total

### **Training/Workshop/Equipment Grants**

- 2017-2022 Genome Canada - Genomics Technology Platforms:  
 Operations Support and Technology Development Funds  
 The Centre for Applied Genomics (TCAG)  
 co-Leaders: Stephen Scherer, Lisa Strug  
 co-Is: **Lei Sun** among over 10 co-Is  
 \$9,174,603 in total
- 2017-2020 Canadian Statistical Sciences Institute (CANSSI)  
 Health Science Collaborating Centre: Collaborating Centre for Statistical Omics  
 co-PIs: Shelley Bull, Laurent Biollais, Rafal Kustra, Lisa Strug, **Lei Sun**  
 \$10,000 in total
- 2014-2015 The Canadian Statistical Sciences Institute (CANSSI) - Workshop Grant  
*Statistical issues in biomarker and drug co-development*  
 Coordinator: Judy-Anne Chapman  
 Organizing committee members: **Lei Sun** among 14 members  
 \$17,145 in total

- 2010-2016 Canadian Institutes of Health Research (CIHR) Training Grant  
*STAGE (Strategic Training for Advanced Genetic Epidemiology):  
 An integrated program in statistical & epidemiological training for  
 genetics with a population health impact*  
 co-PIs and co-Directors: France Gagnon, Shelley Bull  
 3 other co-PIs: Steven Narod, Andrew Paterson, **Lei Sun**  
 \$294,969/year and 5,000 for equipment & 1,774,814 in total
- 2005-2006 Natural Sciences and Research Council of Canada (NSERC) Equipment Grant  
*Enhanced computing resources for statistical research*  
 PI: Radford Neal; co-Is: **Lei Sun** among over 10 co-Is  
 \$88,000 in total
- 2002-2003 Natural Sciences and Research Council of Canada (NSERC) Equipment Grant  
*High performance parallel computer server for statistics and biostatistics*  
 PI: James Stafford; co-Is: **Lei Sun** among over 10 co-Is  
 \$44,728 in total

## Research Contributions ([Google Citations](#))

Trainees underlined, Name<sup>#</sup> for equal contributions and **Sun L\*** for senior or co-senior corresponding authorship

### Refereed Book Chapters

5. **Sun L\*** (2017). [Detecting pedigree relationship errors](#). In *Statistical Human Genetics: Methods and Protocols, 2nd Edition*, Elston R (Editor). Springer.
4. **Sun L\***, Dimitromanolakis A, Chen WM (2017). [Identifying cryptic relationships](#). In *Statistical Human Genetics: Methods and Protocols, 2nd Edition*, Elston R (Editor). Springer.
3. Craiu RV, **Sun L\*** (2014). [Bayesian methods in Fisher's statistical genetics world](#). In *Statistics in Action: A Canadian Perspective*, Lawless JF (Editor).
2. **Sun L\*** (2012). [Detecting pedigree relationship errors](#). In *Statistical Human Genetics: Methods and Protocols*, Elston R, Satagopan J and Sun S (Editors), Human Press, Inc. Springer, pp.25-46.
1. **Sun L\***, Dimitromanolakis A (2012). [Identifying cryptic relationships](#). In *Statistical Human Genetics: Methods and Protocols*, Elston R, Satagopan J and Sun S (Editors), Human Press, Inc. Springer, pp.47-58.

## Refereed Journal Publications

62. Gonçalves VF<sup>#</sup>, Cappi C<sup>#</sup>, Hagen CM, Sequeira A, Vawter MP, [Derkach A](#), Zai CC, Sequeira A, Hedley PL, Bybjerg-Grauholm J, Pouget JG, Cuperfain AB, Sullivan PF, Christiansen M, Kennedy JL<sup>\*</sup>, **Sun L**<sup>\*</sup> (2018). [A comprehensive analysis of nuclear-encoded mitochondrial genes in schizophrenia](#). *Biological Psychiatry* 83(9):780-789.
61. [Panjwani N](#)<sup>#</sup>, [Xiao B](#)<sup>#</sup>, Xu L, [Gong J](#), Keenan K, Lin F, [He G](#), Baskurt Z, [Kim S](#), [Zhang L](#), [Esmaeili M](#), Blackman S, Scherer SW, Corvol H, Drumm M, Knowles M, Cutting G, Rommens JM, **Sun L**, Strug LJ (2018). [Improving imputation in disease-relevant regions: lessons from cystic fibrosis](#). *npj Genomic Medicine* 3:8;doi:10.1038/s41525-018-0047-6.
60. [Soave D](#), **Sun L**<sup>\*</sup> (2017). [A generalized Levene's scale test for variance heterogeneity in the presence of sample correlation and group uncertainty](#). *Biometrics* 73(3):960-971.
59. Yoo YJ, **Sun L**, Poirier J, Paterson AD, Bull SB (2017). [Multiple-linear-combination \(MLC\) regression tests for common variants adapted to linkage disequilibrium structure](#). *Genetic Epidemiology* 41(2):108-121.
58. Strug LJ, Gonska T, [He G](#), Keenan K, Ip W, Boelle PY, Lin F, [Panjwani N](#), [Gong J](#), Li W, [Soave D](#), [Xiao B](#), Tullis E, Rabin H, Parkins MD, Price A, Zuberbuhler PC, Corvol H, Ratjen F, **Sun L**, Bear CE, Rommens JM (2016). [Cystic fibrosis gene modifier SLC26A9 modulates airway response to CFTR-directed therapies](#). *Human Molecular Genetics* 25 (20): 4590-4600.
57. [Xu L](#), Craiu RV<sup>\*</sup>, **Sun L**<sup>\*</sup>, Paterson AD (2016). [Parameter expanded algorithms for Bayesian latent variable modeling of genetic pleiotropy data](#). *Journal of Computational and Graphical Statistics* 25(2):405-425.
56. [Derkach A](#), Lawless JF<sup>\*</sup>, **Sun L**<sup>\*</sup> (2015). [Score tests for association under response-dependent sampling designs for expensive covariates](#). *Biometrika* 102(4):988-994.
55. Corvol H, Blackman S, Boelle PV, Gallins P, Pace R, Stonebraker J, Accurso F, Clement A, Collaco J, Dang H, Dang A, Franca A, Gong J, Guillot L, Keenan K, Li W, Lin F, Patrone M, Raraigh K, **Sun L**, Zhou YH, O'Neal W, Sontag M, Levy H, Durie P, Rommens J, Drumm M, Wright F, Strug L, Cutting G, Knowles M (2015). [Genome-wide association meta-analysis identifies five modifier loci of lung disease severity in cystic fibrosis](#). *Nature Communications* 6. doi:10.1038/ncomms9382
54. [Poirier JG](#), [Faye LL](#), [Dimitromanolakis A](#), Paterson AD, **Sun L**, Bull SB (2015). [Resampling to address the winner's curse in genetic association analysis of time to event](#). *Genetic Epidemiology* 39(7):518-528.

53. Soave D, Corvol H, Panjwani N, Gong J, Wei Li, Boelle PV, Durie P, Paterson AD, Rommens JM, Strug LJ\*, **Sun L\*** (2015). [A joint location-scale test improves power to detect associated SNPs, gene-sets and pathways.](#) *The American Journal of Human Genetics* 97(1):125-138.
- Selected for the inaugural “Trainee Paper Spotlight” that highlights outstanding publications and research done by trainees, by the American Society of Human Genetics’ Training and Development Committee.*
- Recommended by Faculty of 1000 In F1000Prime.*
52. Miller MR, Soave D, Li W, Gong J, Pace RG, Boelle PV, Cutting GR, Drumm ML, Knowles MR, **Sun L**, Rommens JM, Accurso F, Durie PR, Corvol H, Levy H, Sontag MK, Strug LJ (2015). [Variants in solute carrier SLC26A9 modify prenatal exocrine pancreatic damage in cystic fibrosis.](#) *Journal of Pediatrics* 166(5):1152-1157.
51. Hosseini SM, Boright AP, **Sun L**, Canty AJ, Bull SB, Klein BE, Klein R, the DCCT/EDIC Research Group, Paterson AD (2015). [The association of previously reported polymorphisms for microvascular complications in a meta-analysis of diabetic retinopathy.](#) *Human Genetics* 134(2):247-257.
50. Soave D, Miller M, Keenan K, Li W, Gong J, Ip W, Accurso F, **Sun L**, Rommens JM, Sontag M, Durie PR, Strug LJ (2014). [Evidence for a causal relationship between early exocrine pancreatic disease and cystic fibrosis-related diabetes: a Mendelian randomization study.](#) *Diabetes* 63(6):2114-2119.
49. Gonçalves VF, Zai CC, Tiwari AK, Derkach A, Meltzer HY, Lieberman JA, Mueller DJ, **Sun L\***, Kennedy JL\* (2014). [A hypothesis driven association study of 28 nuclear-encoded mitochondrial genes with antipsychotic-induced weight gain in schizophrenia.](#) *Neuropsychopharmacology* 39:1347-1354.
48. Derkach A, Lawless J\*, **Sun L\*** (2014). [Pooled association tests for rare genetic variants: a review and some new results.](#) *Statistical Science* 29(2):302-321.
47. Blue EM, **Sun L\***, Tintle NL, Wijsman EM (2014). [Value of Mendelian laws of segregation in families: data quality control, imputation and beyond.](#) *Genetic Epidemiology* 38(S1):21-28.
46. Xu L, Craiu RV, Derkach A, Paterson AD, **Sun L\*** (2014). [Using a Bayesian latent variable approach to detect pleiotropy in the GAW18 Data.](#) *BMC Proceedings* 8(S1):S77, 1-5.
45. **Sun L\***, Dimitromanolakis A (2014). [PREST-plus identifies pedigree errors and cryptic relatedness in the GAW18 sample using genome-wide SNP data.](#) *BMC Proceedings* 8(S1):S23, 1-6.

44. [Derkach A](#), Lawless J, Merico D, Paterson AD, **Sun L\*** (2014). [Evaluation of gene-based association tests for analyzing rare variants using Genetic Analysis Workshop 18 data](#). *BMC Proceedings* 8(S1):S9, 1-6.
43. Bickeboller H, Bailey JN, Beyene J, Cantor RM, Cordell HJ, Culverhouse RC, Engelman CD, Fardo DW, Ghosh S, Křnig IR, Bermejo JL, Melton PE, Santorico SA, Satten GA, **Sun L**, Tintle NL, Ziegler A, MacCluer JW, Almasy L (2014). [Genetic Analysis Workshop 18: Methods and strategies for analyzing human sequence and phenotype data in members of extended pedigrees](#). *BMC Proceedings* 8(S1):1-4.
42. [Li W](#), [Soave D](#), [Miller MR](#), Keenan K, Lin F, [Gong J](#), [Chiang T](#), Stephenson AL, Durie P, Rommens J, **Sun L\***, Strug LJ (2014). [Unraveling the complex genetic model for Cystic Fibrosis: pleiotropic effects of modifier genes on early CF-related morbidities](#). *Human Genetics* 133(2):151-161.
41. Yoo YJ, **Sun L**, Bull SB (2013). [Gene-based multiple regression association testing for combined examination of common and low frequency variants in quantitative trait analysis](#). *Frontiers in Genetics* 4(233):1-17.
40. Blackman S, Commander C, Watson C, Arcara K, Strug L, Stonebraker J, Wright F, Rommens J, **Sun L**, Pace R, Norris S, Durie P, Drumm M, Knowles M, Cutting G (2013). [Genetic modifiers of cystic fibrosis-related diabetes](#). *Diabetes* 62(10):3627-3635.
39. [Faye LL](#), Machiela MJ, Kraft P, Bull SB, **Sun L\*** (2013). [Re-ranking sequencing variants in the post-GWAS era](#). *PLoS Genetics* 9(8):e1003609:1-16.
38. [Acar E](#), **Sun L\*** (2013). [A generalized Kruskal-Wallis test incorporating group uncertainty with application to genetic association studies](#). *Biometrics* 69(2):427-435.
37. [Derkach A](#), Lawless J\*, **Sun L\*** (2013). [Robust and powerful tests for rare variants using Fisher's method to combine evidence of association from two or more complementary tests](#). *Genetic Epidemiology* 37(1):110-121.  
*In the Genetic Epidemiology Publisher's September 2014 report, this work was among the top 10 most downloaded and cited articles published in 2013.*
36. [Gonçaves VF](#), Tiwari AK, de Luca V, Kong SL, Zai C, Tampakeras M, Mackenzie B, **Sun L**, Kennedy JL (2012). [DRD4 VNTR polymorphism and age at onset of severe mental illnesses](#). *Neuroscience Letters* 519(1):9-13.
35. **Sun L#**, Rommens JM#, Corvol H, [Li W](#), [Li X](#), Chiang T, Lin F, Dorfman R, Busson PF, Parekh RV, Zelenika D, Blackman S, Corey M, Doshi V, Henderson L, Naughton K, O'Neal WK, Pace RG, Stonebraker JR, Wood SD, Wright FA, Zielenski J, Clement A, Drumm ML, Bořlle PY, Cutting GR, Knowles MR, Durie PR, Strug LJ (2012). [Multiple apical plasma membrane constituents are associated with susceptibility to meconium ileus in individuals with cystic fibrosis](#). *Nature*



*Genetics* 44(5):562-569.

Featured as one of the “10 Big Stories in Personalized Medicine” in the McLaughlin Centre Biennial Report, and in the Hospital for Sick Children [News Release](#).

34. [Mirea L](#), Infante-Rivard C, **Sun L**, Bull SB (2012). [Strategies for genetic association analyses combining unrelated case-control individuals and family trios](#). *American Journal of Epidemiology* 176(1):70-79.
33. Wright F, Strug LJ, Doshi VK, Commander CW, Blackman SM, **Sun L**, Berthiaume Y, Cutler D, Cojocaru A, Collaco JM, Corey M, Dorfman R, Goddard K, Green D, Kent Jr JW, Lange EM, Lee S, Li W, Luo J, Mayhew GM, Naughton KM, Pace RG, ParÓ P, Rommens J, Sandford A, Stonebraker JR, Sun W, Taylor C, Vanscoy LL, Zou F, Blangero J, Zielenski J, O’Neal WK, Drumm ML, Durie PR, Knowles MR, Cutting GR (2011). [Genome-wide association and linkage identify modifier loci of lung disease severity in cystic fibrosis at 11p13 and 20q13.2](#). *Nature Genetics* 43(6):539-548.
32. [Faye LL](#), **Sun L\***, [Dimitromanolakis A](#), Bull SB\* (2011). [A flexible genome-wide bootstrap method that accounts for ranking- and threshold-selection bias in GWAS interpretation and replication study design](#). *Statistics in Medicine* 30(15):1898-1912.
31. **Sun L\*** (2011). [On the efficiency of genome-wide scans: a multiple hypothesis testing perspective](#). U.P.B. Sci. Bull., Series A., 73(1):19-26.
30. **Sun L\***, [Dimitromanolakis A](#), [Faye L](#), Paterson AD, [Waggott D](#), The DCCT/EDIC Research Group, Bull SB (2011). [BR-squared: a practical solution to the winner’s curse in genome-wide scans](#). *Human Genetics* 129(5):545-552.
29. [Dorfman R](#), Taylor C, Lin F, **Sun L**, Sandford A, Pare P, Berthiaume Y, Corey M, Durie P, Zielenski J, on behalf of the members of the Canadian Consortium for CF Genetic Studies (2011). [Modulatory effect of the SLC9A3 gene on susceptibility to infections and pulmonary function in children with cystic fibrosis](#). *Pediatric Pulmonology* 46(4):385-392.
28. [Li W](#), **Sun L**, Corey M, Zou F, Lee S, [Cojocaru AL](#), Taylor C, Blackman SM, Stephenson A, Sandford AJ, Dorfman R, [Drumm ML](#), Cutting GR, Knowles MR, Durie P, Wright FA, Strug LJ (2011). [Understanding the population structure of North American patients with Cystic Fibrosis](#). *Clinical Genetics* 79(2):136-146.
27. [Xu L](#), Craiu RV, **Sun L\*** (2011). [Bayesian methods to overcome the winner’s curse in genetic studies](#). *Annals of Applied Statistics* 5(1):201-231.
26. [Mirea L](#), **Sun L\***, Stafford JE, Bull SB (2010). Using evidence for population stratification bias in combined individual- and family-level genetic association analyses of quantitative traits. *Genetic Epidemiology* 34:502-511.

25. Paterson AD, Waggott D, Boright AP, Hosseini SM, Shen E, Sylvestre MP, Wong I, Bharaj B, Cleary PA, Lachin JM, MAGIC, Below JE, Nicolae D, Cox NJ, Canty AJ, **Sun L**, Bull SB, and the DCCT/EDIC Research Group (2010). A genome-wide association study identifies a novel major locus for glycemic control in type 1 diabetes, as measured by both HbA1c and glucose. *Diabetes* 59:539-549.
24. Yoo YJ, Bull SB, Paterson AD, Waggott D, DCCT/EDIC Research Group, **Sun L\*** (2010). Were genome-wide linkage studies a waste of time? Exploiting candidate regions within genome-wide association studies. *Genetic Epidemiology* 34:107-118.
23. Paterson AD, Lopes-Virella MF, Waggott D, Boright AP, Hosseini SM, Carter RE, Shen E, Mirea L, Bharaj B, **Sun L**, Bull SB, and the DCCT/EDIC Research Group (2009). Genome-wide association identifies the ABO blood group as a major locus associated with serum levels of soluble E-Selectin. *Arteriosclerosis, Thrombosis, and Vascular Biology* 29:1958-1967.
22. Dorfman R, Li W, **Sun L**, Lin F, Wang Y, Sandford S, Pare PD, McKay K, Kayserova H, Macek M, Bal J, Sands D, Tiddens H, Castro S, Sontag M, Accurso FJ, Blackman S, Cutting GR, Tsui LC, Corey M, Durie P, Zielenski J, Strug LJ (2009). Modifier gene study of meconium ileus in cystic fibrosis: statistical considerations and gene mapping results. *Human Genetics* 126:763-778.
21. Yoo YJ, Pinnaduwaage D, Waggott D, Bull SB, **Sun L\*** (2009). Genome-wide association analyses of North American Rheumatoid Arthritis Consortium and Framingham Heart Study data utilizing genome-wide linkage results. *BMC Proceedings* 3:S103.
20. Asimit J, Yoo YJ, Waggott D, **Sun L**, Bull SB (2009). Region-based analysis in genome-wide association study of Framingham Heart Study blood lipid phenotypes. *BMC Proceedings* 3:S127.
19. Craiu RV, **Sun L\*** (2008). Choosing the lesser evil: trade-off between false discovery rate and non-discovery rate. *Statistica Sinica* 18:861-879.
18. Lee SSF, **Sun L\***, Kustra R, Bull SB (2008). EM-random forest and new measures of variable importance for multi-Locus quantitative trait linkage analysis. *Bioinformatics* 24:1603-1610.
17. Dorfman R, Sandford A, Taylor C, Huang B, Frangolias D, Wang Y, Sang R, Pereira L, **Sun L**, Berthiaume Y, Tsui LC, Pare PD, Durie P, Corey M, Zielenski J (2008). Complex two-gene modulation of lung disease severity in children with cystic fibrosis. *Journal of Clinical Investigation* 118:1040-1049.
16. Al-Kateb H, Boright AP, Xie X, Mirea L, Sutradhar R, Mowjoodi A, Bharaj B, Liu M, Buckska JM, Arends VL, Steffes MW, Cleary PA, Sun W, Lachin JM, Thorner PS,

- Ho M, McKnight AJ, Maxwell PA, Savage DA, Kidd KK, Kidd JR, Speed WC, Orchard TJ, Miller RG, **Sun L**, Bull SB, Paterson AD and the DCCT/EDIC research group (2008). Multiple superoxide dismutase 1/splicing factor serine alanine 15 variants are associated with the development and progression of diabetic nephropathy. *Diabetes* 57:218-228.
15. Al-Kateb H, Mirea L, Xie X, **Sun L**, Liu M, Chen H, Bull SB, Boright AP, Paterson AD, The DCCT/EDIC Research Group (2007). Multiple variants in Vascular Endothelial Growth Factor (VEGF) are risk factors for time to severe retinopathy in type 1 diabetes: The DCCT/EDIC genetics studies. *Diabetes* 56:2161-2168.
  14. Huang B, Rangrej J, Paterson AD, **Sun L\*** (2007). The multiplicity problem in linkage analysis of gene expression data - the power of differentiating *cis*- and *trans*-acting regulators. *BMC Proceedings* 1:S142.
  13. Greenwood C, Rangrej J, **Sun L\*** (2007). Optimal selection of markers for validation from genome-wide association studies. *Genetic Epidemiology* 31:396-407.
  12. Wu LY, **Sun L\***, Bull SB (2006). Locus-specific heritability estimation via the bootstrap in linkage scans for quantitative trait loci. *Human Heredity* 62:84-96.
  11. **Sun L\***, Craiu RV, Paterson AD and Bull SB (2006). Stratified false discovery control for large-scale hypothesis testing with application to genome-wide association studies. *Genetic Epidemiology* 30:519-530.
  10. Wu LY, Lee SSF, Shi HS, **Sun L**, Bull SB (2005). Resampling methods to reduce the selection bias in genetic effect estimation in genome-wide scans. *Genetic Analysis Workshop 14: Microsatellite and single-nucleotide polymorphism. BMC Genetics* 6:S24.
  9. Biernacka J, **Sun L\***, Bull SB (2005). Tests for the presence of two linked disease susceptibility genes. *Genetic Epidemiology* 29:389-401.
  8. **Sun L\***, Bull SB (2005). Reduction of selection bias in genome-wide genetic studies by resampling. *Genetic Epidemiology* 28:352-367.
  7. Biernacka J, **Sun L\***, Bull SB (2005). Simultaneous localization of two linked disease susceptibility genes. *Genetic Epidemiology* 28:33-47.
  6. Paterson A, **Sun L**, Liu XQ (2003). Transmission ratio distortion in families from the Framingham Heart Study. *Genetic Analysis Workshop 13: Analysis of longitudinal family data for complex diseases and related risk factors. BMC Genetics* 4:S48.
  5. Strug LJ, **Sun L**, Corey M (2003). The genetics of cross-sectional and longitudinal BMI. *Genetic Analysis Workshop 13: Analysis of longitudinal family data for complex diseases and related risk factors. BMC Genetics* 4:S14.

4. **Sun L\***, Wilder K, McPeck MS (2002). Enhanced pedigree error detection. *Human Heredity* 54:99-110.
3. **Sun L\***, Cox NJ, McPeck MS (2002). A statistical method for identification of polymorphisms that explain a linkage result. *American Journal of Human Genetics* 70:399-411.
2. **Sun L\***, Abney M, McPeck MS (2001). Detection of misspecified relationships in inbred and outbred pedigrees. *Genetic Analysis Workshop 12: Analysis of complex genetic traits: Applications to asthma and simulated data. Genetic Epidemiology* 21:S36-S41.
1. McPeck MS, **Sun L\*** (2000). Statistical tests for detection of misspecified relationships by use of genome-screen data. *American Journal of Human Genetics* 66:1076-1094.

### Publicly-Released Research Software

*An important component of this research program is the implementation of developed and tested methodology as user-friendly and open-resource software and programs as part of the Research Dissemination and Knowledge Transfer and Exchange. All programs are freely available at <http://www.utstat.toronto.edu/sun/>*

11. Dimitromanolakis A, Paterson A, **Sun L\*** (2017). *TRUFFLE*, a scalable learning of shared ancestry from high-throughput genetic data.
10. Soave D, **Sun L\*** (2017). *gJLS*, a generalized Joint Location-Scale association testing framework that tests the null hypothesis of equal mean and equal variance between genotypes, allowing for sample correlation and group membership uncertainty.
9. Soave D, Strug LJ, **Sun L\*** (2015). *JLS*, a Joint Location-Scale association testing framework that detects both main and interaction effects.
8. Acar E, **Sun L\*** (2011). *gKW*, a Generalized Kruskal-Wallis test that incorporates group uncertainty.
7. Dimitromanolakis A, **Sun L\*** (2010). *PREST-plus*, for detecting pedigree errors and cryptic relationships.
6. Dimitromanolakis A, Faye L, Bull SB, **Sun L\*** (2009). *BR-squared*, Bias-Reduced estimates via Bootstrap Resampling for overcoming the winner's curse due to selective inference
5. Yoo YJ, **Sun L\*** (2008). *sFDR*, a stratified False Discovery Rate control for multiple hypothesis testing.

4. Lee SSF, **Sun L**, Kustra R, Bull SB (2007). *EMRF*, a EM-Random Forest approach for multi-locus quantitative trait linkage analysis.
3. Wu LY, **Sun L\***, Bull SB (2006). *BR-squared - Linkage* for calculation of Bias-Reduced Bootstrap Resampling-based estimates of locus-specific gene-effect size in linkage analyses
2. Wen W, **Sun L\*** (2004). *STEPC*, STatistical Explanation of Positional Cloning for identification of polymorphisms that explain a linkage result.
1. **Sun L\***, Wilder K, McPeck MS (2000). *PREST*, Pedigree RElationship Statistical Test and ALTERTEST for detecting pedigree errors and estimation of relationships using genome-wide marker data.

## Patents

1. Patent application for “Modifiers of CFTR-directed therapy”, WIPO (PCT) Patent Application No. 16840481.2 - 1111 PCA/CA2016051044, filed September 2, 2016. Inventors are
  - Johanna Rommens, Hospital for Sick Children, Toronto, Ontario, Canada
  - Lisa Strug, Hospital for Sick Children, Toronto, Ontario, Canada
  - Lei Sun, University of Toronto, Toronto, Ontario, Canada

## Invited Talks/Lectures

- June 4, 2018, Statistical Society of Canada (SSC) Annual Meeting, Montreal, Quebec, Canada. (Panel discussion on NSERC, CIHR and Beyond, organized by New Investigators Committee)
- June 4, 2018, Statistical Society of Canada (SSC) Annual Meeting, Montreal, Quebec, Canada. (Session on Analyses of Big/Large Health Data: Challenges and Strategies, organized by Committee on Women in Statistics)
- May 2, 2018, Organization for the Study of Sex Differences (OSSD) Annual Meeting, Atlanta, Georgia, USA.
- April 5, 2018, Department of Statistics and Actuarial Science, University of Waterloo.
- January 26, 2018, CRM-SSC Prize 2017 Colloque, the Centre de recherches mathématiques (CRM), Montreal, Canada.
- December 1, 2017, Department of Mathematics and Statistics, McGill University, Montreal, Canada.

- June 14, 2017, CRM-SSC Prize in Statistics Address, Statistical Society of Canada (SSC) Annual Meeting, Winnipeg, Canada.
- June 6, 2017, Institute of Big Data, Fudan University, Shanghai, China.
- May 31, 2017, Institute of Statistics, Nankai University, Tianjin, China.
- May 26, 2017, Department of Statistics and Financial Mathematics, Nanjing University of Science and Technology, Nanjing, China.
- May 24, 2017, 50<sup>th</sup> Anniversary Seminar Series, Department of Statistics and Actuarial Science, University of Hong Kong, Hong Kong.
- April 12, 2017, Department of Public Health Sciences, University of Chicago, Chicago, USA.
- March 30, 2017, Department of Statistics, University of Manitoba, Winnipeg, Canada.
- August 2, 2016, Joint Statistical Meetings (JSM), Chicago, USA.
- June 13, 2016, ICSA Applied Statistics Symposium, Atlanta, USA.
- May 30, 2016, The 2nd Conference on New Advances in Big Data, Nanjing, China.
- May 25, 2016, Institute of Statistics, Nankai University, Tianjin, China.
- December 3, 2015, Department of Biostatistics, Columbia University, New York, USA.
- April 29, 2015, Department of Economics, University Ca' Foscari of Venice, Venezia, Italy.
- April 21, 2015, School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, Kent, UK.
- March 13, 2015, Department of Statistical Science, University of Roma, Rome, Italy.
- February 5, 2015, The French National Institute for Agricultural Research (INRA), Toulouse, France.
- August 4, 2014, Joint Statistical Meetings (JSM), Boston, USA
- May 27, 2014, Statistical Society of Canada (SSC) Annual Meeting, Toronto, Canada
- March 25, 2014, Department of Statistics, University of British Columbia, Canada
- March 21, 2014, Department of Statistics and Actuarial Science, Simon Fraser University, Canada

- March 17, 2014, International Biometric Society Eastern North American Region Regional Meeting (ENAR), Baltimore, USA
- September 25, 2013, Department of Statistical Science, Cornell University, USA
- June 4, 2013, Advanced Topics in Genome-Wide Association Studies, Ontario Institute for Cancer Research (OICR), Toronto, Canada
- April 10, 2013, Department of Statistics and Biostatistics, Rutgers University, USA
- April 12, 2012, Department of Statistics and Actuarial Science, University of Waterloo, Canada
- May 29, 2012, Advanced Topics in Genome-Wide Association Studies, Ontario Institute for Cancer Research (OICR), Toronto, Canada
- February 10, 2012, Department of Mathematics and Statistics, Acadia University, Canada
- November 14, 2011, Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, USA
- July 9, 2011, Institute of Mathematical Statistics - China International Conference on Statistics and Probability, XiAn, China
- May 12, 2011, The 6th Canadian Genetic Epidemiology and Statistical Genetics Workshop, Toronto, Canada
- September 24, 2010. The 1st Bruce Kaufman Symposium on Immunoregulation and inflammatory bowel disease, Montreal, Canada
- July 20, 2010, Banff International Research Station (BIRS) 5-day Workshop - Statistical Genomics in Biomedical Research, Banff, Canada
- January 8, 2010, Cancer Epidemiology Discussion Group, Samuel Lunenfeld Research Institute Mount Sinai Hospital, Toronto, Canada
- August 3, 2009, Joint Statistical Meetings (JSM), Washington DC, USA
- July 4, 2009, Institute of Mathematical Statistics - China International Conference on Statistics and Probability, Weihai, China
- April 3, 2009, Department of Biostatistics & Medical Informatics, University of Wisconsin, Madison, USA
- March 27, 2009, CRM-ISM-GERAD Statistics Colloque (Universite de Montreal, McGill University, Concordia University, and Universite du Quebec a Montreal), Canada

- September 30, 2008, Department of Mathematics and Statistics, U. of Guelph, Canada
- April 17, 2008, Department of Statistics, University of British Columbia, Canada
- April 16, 2008, Department of Statistics and Actuarial Science, Simon Fraser University, Canada
- October 11, 2007, Genetics & Genomic Biology, HSC Research Institute, Toronto, Canada
- June 13, 2007, Statistical Society of Canada (SSC) Annual Meeting, St. John's, Newfoundland, Canada
- July 2006, SAMSI Workshop on Multiplicity and Reproducibility in Scientific Studies, Research Triangle Park, USA
- May 2006, International Workshop on Applied Probability (IWAP), University of Connecticut, USA
- April 2006, Center for Statistical Genetics, University of Michigan at Ann Arbor, USA
- April 2006, Department of Statistics, University of Chicago, USA
- March 2006, The 1st Canadian Genetic Epidemiology and Statistical Genetics Workshop, Toronto, USA
- February 2006, Department of Mathematics and Statistics, Laval University, Canada
- December 2005, Department of Statistics, University of Toronto, Canada
- December 2004, Department of Statistics, Colorado State University, USA
- September 2003, Department of Mathematics and Statistics, York University, Canada
- March 2003, Department of Statistics and Actuarial Science, University of Waterloo, Canada
- February 2002, Genetics & Genomic Biology, HSC Research Institute, Toronto, Canada
- June 2001, Applied Statistics Symposium, International Chinese Statistical Association (ICSA), Chicago, USA
- April 2001, Department of Public Health Sciences, University of Toronto, Canada
- February 2001, Department of Biostatistics, University of North Carolina at Chapel Hill, USA



- February 2001, Department of Statistics, Ohio State University, USA

## Academic/Professional Activities and Services - International

### Editorial Board

10/2012-present	Editorial Board	Genetic Epidemiology
09/2011-present	Review Editor	Frontiers in Statistical Genetics and Methodology
05/2015	Guest Associate Editor	PLoS Genetics
07/2007-08/2012	Associate Editor	Statistical Applications in Genetics and Molecular Biology

### Scientific Advisory Board

- 09/2010-08/2015, CIHR Net Grant entitled “Immunomodulation of Regulatory Mechanisms in Mucosal Immunity: A multi-disciplinary bench-to-bedside approach to the study and treatment of IBD”. McGill University Health Centre, University of Montreal.

### Organizer and Member

- 2019, Co-Organizer, 5-day Workshop at the Banff International Research Station (BIRS) for Mathematical Innovation and Discovery; “The Role of Genomics and Metagenomics in Human Health: Recent Developments in Statistical and Computational Methods”.
- 2017-2018, Member, Scientific Committee of the 7<sup>th</sup> Canadian Human and Statistical Genetics meeting (CHSGM).
- 2018, Organizer and Chair, Invited session on “Recent Development in Statistical Genetics and Genomics”, Statistical Society of Canada (SSC) annual meeting, McGill University, Montreal, Canada.
- 2017, Chair of the Poster Session and Member of the Organizing Committee, “40 years of Statistical Sciences at the University of Toronto”. The Fields Institute, Toronto, Canada.
- 2016-2018, Member, Publication Committee, the International Genetic Epidemiology Society (IGES).
- 2014, Organizing Committee Member, 2-day Workshop at the Fields Institute; “Statistical issues in biomarker and drug co-development”.
- 2014, Co-Organizer, 5-day Workshop at the Banff International Research Station (BIRS) for Mathematical Innovation and Discovery; “Emerging Statistical Challenges and Methods For Analysis of Massive Genomic Data in Complex Human Disease Studies”.

- 2014, Chair, Session on New Estimation Methods, Statistical Society of Canada (SSC) Annual Meeting, Toronto, Canada.
- 2013, Member, Program Committee, IEEE International Conference on Bioinformatics and Biomedicine (BIBM).
- 2013, Member, Program Committee, Eastern North American Region/International Biometric Society meeting (ENAR).
- 2012, Group Leader, Genetic Analysis Workshop (GAW) 18.
- 2012, Member, Program Committee, IEEE International Conference on Bioinformatics and Biomedicine (BIBM).
- 2011, Co-Organizer, Invited session on “Statistical Genetics”, STATISTICS CANADA/IMST.
- 2010, Organizer and Chair, Invited session on “Genomics”, Institute of Mathematical Statistics (IMS)
- 2008, Member, Education Committee, International Genetic Epidemiology Society (IGES)
- 2007, Member, Education Committee, International Genetic Epidemiology Society (IGES)
- 2007, Member, Publication Committee, International Genetic Epidemiology Society (IGES)
- 2005, Organizer and Chair, Invited session on “Statistical analysis and modeling of complex traits”, Joint Statistical Meeting (JSM)
- 2004, Organizer, Invited session on “Statistical genetics - modeling interaction and multi-locus analyses”, Eastern North American Region/International Biometric Society meeting (ENAR)

## Journal Referee

American Journal of Human Genetics (AJHG), Annals of Applied Statistics (AOAS), Annals of Human Genetics, Bioinformatics, Biometrics, BioTechniques, BMC Bioinformatics, BMC Genetics, proceedings of the Genetic Analysis Workshop, BMC Medical Genetics, BMC Medical Research Methodology, Canadian Journal of Statistics (CJS), European Journal of Human Genetics (EJHG), Frontiers in Genetics, Genetics, Genetic Epidemiology, Human Genetics, Human Heredity, IEEE BIBM special issues, Journal of the American Statistical Association (JASA), Obesity, PLoS Genetics, PLoS One, Proceedings of the National Academy of Sciences (PNAS), Statistical Applications in Genetics and Molecular Biology, Statistics in Medicine, Statistica Sinica

## **Grant Review**

Canadian Statistical Sciences Institute (CANSSI), Canadian Breast Cancer Foundation (CBCF), Canadian Institutes of Health Research (CIHR), German Research Foundation (Deutsche Forschungsgemeinschaft), Israel Science Foundation (ISF), Natural Sciences and Engineering Research Council of Canada (NSERC), Internal grant reviewer for Hospital for Sick Children (HSC) Research Institute, Toronto

## **Other International Services**

- As external reviewer for various Full-Professor, Associate-Professor and Tenure promotions.

## **Professional Activities and Services - Local**

### **Department of Statistical Sciences, FAS, University of Toronto**

- 2016-present, Chair, Research Committee
- 2016-present, Member, Executive Committee
- 2015-present, Member, Social Committee
- 2015-present, Member, Hiring/Search Committee
- 2015-present, Member, Graduate Studies Committee
- 2015-present, Member, Data Science Committee
- 2015-present, Member, Promotion Committee
- 2016, Member, PTR Committee
- 2015-2016, Co-Chair, Social Committee
- 2013-2014, Member, Hiring/Search Committee
- 2010-2011, Member, Hiring/Search Committee
- 2008-2009, Member, Hiring/Search Committee

### **Dalla Lana School of Public Health, University of Toronto**

- 2017-present, Member, MPH-Data Science Committee, Division of Biostatistics

- 2010-present, Member, Program Steering Committee, CIHR “STAGE (Strategic Training in Genetic Epidemiology): An integrated program in statistical & epidemiological training for genetics with a population health impact”
- 2010-present, Member, Admission Committee, CIHR STAGE Training Program
- 2002-present, Member, Admission Committee, Division of Biostatistics
- 2002-present, Member, Comprehensive Exam Committee, Division of Biostatistics
- 2013-2014, Member, Hiring/Search Committee, Division of Biostatistics
- 2013, Member, Executive Committee, School Council
- 2013, Chair, Comprehensive Exam Committee, Division of Biostatistics
- 2008-2012, Chair, Admission Committee, Division of Biostatistics
- 2011, Chair, Student Presentation Award Committee, The Inaugural Biostatistics Research Day
- 2008, Co-Chair, Comprehensive Exam Committee, Division of Biostatistics
- 2006-2007, Member, Curriculum Committee
- 2003-2007, Member, Space Committee
- 2001-2007, Associate Member, Graduate Faculty
- 2005-2006, Chair, Comprehensive Exam Committee, Division of Biostatistics
- 2002-2003, Member, Advisory Committee

### **Other Local Services**

- 2017-present, Member, Program Steering Committee, Canadian Statistical Sciences Institute (CANSSI) Health Science Collaborating Centre for Statistical Omics (University of Toronto, McMaster University, and a large number of research hospitals including Lunenfeld-Tanenbaum Research Institute, Hospital for Sick Children).
- 2017, 2016, Reviewer, Data-Science concentration of the MScAC cohort, Department of Computer Science, University of Toronto
- 2017, 2016, Member, Connaught International Scholarship Committee, University of Toronto
- 2016, Chair, PhD Exam Committee, Department of Biochemistry, University of Toronto

- 2015, Host and Resource Person, Immunology Tonight Session: Biostatistics in Medical Research. Department of Immunology, Faculty of Medicine, University of Toronto.
- 2013, Chair, PhD Exam Committee, Department of Medical Biophysics, University of Toronto
- 2011, Member, OGS Fellowship Review Committee, University of Toronto
- 2008, Member, OGS Fellowship Review Committee, University of Toronto

## Teaching

### Graduate Courses

- STA 4515 - Multiple Hypothesis Testing and Its Applications  
Department of Statistical Sciences, University of Toronto  
Fall 2017; Fall 2016; Winter 2016  
*Role: proposed and designed the course and as the sole instructor*
- CHL 5224 - Modern Statistical Genetics  
(Required course for STAGE trainees)  
Division of Biostatistics, DLSPH, University of Toronto  
Fall 2016; Winter 2016  
*Role: proposed and designed the course and as the sole instructor*
- STA 2080 - Fundamentals of Statistical Genetics  
Department of Statistical Sciences, University of Toronto  
Winter 2018; Winter 2017; Fall 2015  
*Role: proposed and designed the course and as the sole instructor*
- CHL 7001 - Statistical Models on Complex Human Genetic Diseases  
Division of Biostatistics, DLSPH, University of Toronto  
Summer 2016; Summer 2013; Summer 2012  
*Role: Guest Lecturer*
- CHL 7001 - An Introduction to the Likelihood Paradigm  
Division of Biostatistics, DLSPH, University of Toronto  
Winter 2014  
*Role: co-designed the course and co-instructor; 2 other co-instructors: Dr. Lisa Strug, Hospital For Sick Children and DLSPH and Dr. Laurent Briollais, Samuel Lunenfeld Research Institute and DLSPH*

- CHL 5224 - Statistical Genetics  
(Required course for STAGE trainees)  
Division of Biostatistics, DLSPH, University of Toronto  
Winter 2016; Fall 2013; Winter 2013; Fall 2006; Fall 2005; Winter 2004; Winter 2003  
*Role: proposed and designed the course and as the sole instructor*  
Winter 2012; Winter 2011; Winter 2010; Winter 2009; Fall 2008; Fall 2007  
*Role: proposed and designed the course and as the primary instructor; co-instructor: Dr. Wei Xu, Princess Margaret Hospital and DLSPH*
- CHL 5210 - Categorical Data Analysis  
(Required course for Biostatistics graduate students)  
Division of Biostatistics, DLSPH, University of Toronto  
Fall 2013; Fall 2012; Fall 2011; Fall 2010; Fall 2009  
*Role: designed the course and as the primary instructor; co-instructor: Dr. Laurent Briollais, Samuel Lunenfeld Research Institute and DLSPH*  
Winter 2004  
*Role: designed the course and as the primary instructor; co-instructor: Dr. Joseph Beyene, Hospital for Sick Children and DLSPH*
- CHL 5207/5208 - Lab in Statistical Design & Analysis  
(Required course for Biostatistics graduate students)  
Division of Biostatistics, DLSPH, University of Toronto  
2016-2017; 2011-2012; 2010-2011  
*Role: Practicum Supervisor*
- CHL 7001 - Statistical Methods for Genetics and Genomics  
(Required course for STAGE trainees)  
Division of Biostatistics, DLSPH, University of Toronto  
2011-2012; 2010-2011  
*Role: one of the three Faculty Discussion Leaders*
- STA 4315 - Computational Methods in Statistical Genetics  
Department of Statistics, Faculty of Arts and Science, University of Toronto  
Winter 2009; Winter 2006; Winter 2005  
*Role: co-proposed and co-designed the course and as co-instructor; co-instructor: Prof. Radu Craiu, Department of Statistics*
- CHL 7001 - Statistical Methods in Genetic Epidemiology  
Division of Biostatistics, Department of PHS, University of Toronto  
Winter 2005  
*Role: proposed and designed the course and as the sole instructor*
- STAT 946 - Statistical Genetics  
(Innovative teaching via video conferencing)  
Department of Statistics and Actuarial Science, University of Waterloo

Winter 2004

*Role: proposed and designed the course, proposed the idea and as the sole instructor*

- CHL 5250 - Special Topics in Biostatistics Division of Biostatistics, Department of PHS, University of Toronto

Winter 2003; Fall 2002

*Role: designed the course and as the sole instructor*

- CHL 7001 - Statistical Methods in Genetics

Division of Biostatistics, Department of PHS, University of Toronto

Winter 2002

*Role: proposed and designed the course and as the sole instructor*

## **Undergraduate Courses**

- STA480 - Fundamentals of Statistical Genetics

Department of Statistical Sciences, University of Toronto

Winter 2018; Winter 2017; Fall 2015

*Role: proposed and designed the course and as the sole instructor*

- STAT220 - Statistical Methods and Their Applications

Department of Statistics, University of Chicago

Spring 2000; Spring 1999; Spring 1998

*Role: as the sole instructor with some contribution to the course design; mostly relying on the existing syllabus*

## **Training of Highly Qualified Personnel (HQP)**

### **Post-Doctoral Fellows (PDF)**

- YanYan Zhao, PDF (09/2017-; Primary Supervisor). Statistical Genetics.
- Marc Woodbury-Smith, PDF in Genetics (10/2011-08/2013; Statistical Genetics Mentor for STAGE trainees). “Genetic studies of autism spectrum disorder”.
- Vanessa F. Gonçalves, PDF in Genetics (03/2011-08/2013; Statistical Genetics Mentor for STAGE trainees). Genetic studies of behavioural phenotypes.
- Yun Joo Yoo, PDF in Statistical Genetics (10/2007-06/2009; Co-Primary Supervisor). “Design and analysis of genome-wide studies of complex diseases and traits”.
- Longyang Wu, PDF in Statistical Genetics (02/2003-08/2006; Co-Primary Supervisor). “Statistical methods to improve the reliability of results from genome-wide studies of complex disease and quantitative traits.”

## PhD Students

- Yanbo Tang, PhD in Statistics (05/2018-; Primary Supervisor). “On p-values obtained from alternatives”.
- Yu-Chung (Jerry) Lin, PhD in Biostatistics (09/2017-; Co-Supervisor). Statistical Genetics.
- Fan Wang, PhD in Statistics (06/2017-; Primary Supervisor). Statistical Genetics.
- Lin Zhang, PhD in Statistics (06/2016-; Primary Supervisor). Statistical Genetics.
- Wei Deng, PhD in Statistics (10/2014-; Primary Supervisor). “Scale tests for X-chromosome variants”.
- Emery Goossens, PhD in Statistics (08/2014-; Primary Supervisor). “Subset analysis accounting for post-selection bias”.
- Bo Chen, PhD in Statistics (09/2013-; Co-Primary Supervisor). “Statistical insights on the association analysis of X chromosome: X-inactivation, heterogeneity and interaction effects”.
- Linda Kachuri, PhD in Epidemiology (10/2014-03/2018; Statistical Genetics Mentor for STAGE trainees). Comprehensive investigation of telomere length and its genetic determinants in cancer risk and clinical outcomes.
- David Soave, PhD in Biostatistics (07/2012-11/2016; Co-Primary Supervisor). “Statistical methods for marker discovery and disease risk prediction amid unobserved genetic interactions”.  
*The paper on joint location-scale association test was [Featured in the “Trainee Paper Spotlight”](#) that highlights outstanding publications and research done by trainees, by the American Society of Human Genetics’ Training and Development Committee. [Nominated for the CAGS/ProQuest-UMI Distinguished Dissertation Award](#) by DL-SPH.*
- Andriy Derkach, PhD in Statistics (06/2010-06/2014; Co-Primary Supervisor). “Statistical methodologies for genetic association studies with rare variants”.  
*[Recipient of the Department of Statistics 2012-2013 Doctoral Award](#), given to a PhD Student for excellence in research.  
[One of the three finalists for the 2012 Williams Award](#) of the International Genetic Epidemiology Society.*
- Elif Acar, PhD in Statistics (04/2009-11/2011; Primary Supervisor). “A generalized Kruskal-Wallis test incorporating group uncertainty with application to genetic association studies.”



- Laura Faye, PhD in Biostatistics (09/2007-06/2013; Co-Primary Supervisor). “Identification of candidate causal variants and estimation of genetic associations in GWAS and post-GWAS studies.”  
*Winner of the 2012 Williams Award of the International Genetic Epidemiology Society.*
- Lizhen Xu, PhD in Statistics (09/2006-11/2012; Co-Supervisor). “Bayesian methods for genetic association studies.”  
*Recipient of the Department of Statistics 2011-2012 Doctoral Award, given to a PhD Student for excellence in research.*
- Sophia Lee, PhD in Biostatistics (09/2003-11/2007; Co-Primary Supervisor). “Random forests for multi-locus quantitative trait linkage analysis”.
- Lucia Mirea, PhD in Biostatistics (09/2005-03/2011; Co-Supervisor). “Approaches incorporating evidence for population stratification bias in genetic association analyses combining individual and family data”.
- Joanna Biernacka, PhD in Biostatistics (09/2001-11/2004; Co-Primary Supervisor). “Statistical methods for studying two linked disease genes.”

## Research Assistants/ Associates

*The Master programs of both biostatistics and statistics are mostly course-based programs, so with a few exceptions, the supervision of MSc students in this section is research project-based unless specified otherwise.*

- Apostolos Dimitromanolakis, Research Associate (01/2017-; Primary Supervisor). “Exploratory data analysis and data visualization”.
- Ting Zhang, (thesis-based) MSc in Biostatistics (12/2016-; Primary Supervisor). “Statistical genetics”.
- Scott Mastromatteo, MSc in Applied Computer Science with concentration in Data Science (10/2016-; Primary Supervisor). “Ranking enriched motifs from Chip-seq data”.
- Rose Garrett, MSc in Biostatistics (09/2016-09/2017; Primary Supervisor). “Statistical methods for gene expression data of *C.elegans*”.
- Emery Goossens, MSc in Statistics (01/2014-08/2014; Primary Supervisor). “Genetic model selection using a non-nested linear model comparison framework”.
- Lauren Erdman, MSc in Biostatistics (12/2013-10/2014; Primary Supervisor). “Combining association evidence from different data resources”.

- Sayem Borhan, MSc in Biostatistics (11/2012-08/2013; Primary Supervisor). “X-chromosome analysis”.
- Apostolos Dimitromanolakis, Research Assistant (06/2012-03/2013; Primary Supervisor). “PREST-plus identifies pedigree errors and cryptic relatedness in the GAW18 sample using genome-wide SNP data”.
- Jingxiong (Amelia) Xu, MSc in Biostatistics (01/2012-04/2012; Primary Supervisor). “Multiple hypothesis testing - robust methods incorporating prior information”.
- Paul Popadiuk, MSc in Biostatistics (09/2011-04/2012; Primary Supervisor). “Interpreting the statistical evidence for association analysis of SNPs in the context of genotype uncertainty”.
- Julia Taleban, Programmer (10/2010-09/2012; Co-Supervisor). “A general procedure to address the winner’s curse in genetic association studies: bias-reduction in analysis of time-to-event traits”.
- Haiyan Jiang, MSc in Biostatistics (10/2010-06/2011; Co-Primary Supervisor). “Genome-wide linkage and association study of APCR trait”.
- Apostolos Dimitromanolakis, Research Associate (09/2010-09/2011; Co-Primary Supervisor). “Genetic studies of venous thromboembolism”.
- Junming Yang, MSc in Biostatistics (02/2010-11/2010; Primary Supervisor). “Joint analysis of multiple correlated phenotypes and SNPs”.
- Apostolos Dimitromanolakis, MSc in Biostatistics (11/2008-11/2010; Primary Supervisor). “Developing PREST-plus for detection of pedigree errors and cryptic relatedness using high-throughput genotype data”.
- Andreea Cojocaru, Research Assistant (10/2008-09/2009; Co-Supervisor). “Association analysis for Meconium Ileus incorporating phenotype misclassification.”
- Apostolos Dimitromanolakis, Programmer (10/2008-10/2009; Co-Primary Supervisor). “Implementation of bootstrap bias reduction methods for genome-wide association data.”
- Theodore (Ted) Chiang, Programmer (02/2008-07/2008; Primary Supervisor). “Upgrading PREST for fast and efficient detection of pedigree errors and cryptic relatedness using high-throughput genotype data.”
- Lam Opal Huang, MSc in Biostatistics (09/2007-09/2008; Primary Supervisor). “Some statistical issues in genome-wide association of common alleles with long-term diabetic complications.”

- Baisong Huang, PhD in Biostatistics (05/2006-06/2007; Primary Supervisor). Genetic Analysis Workshop (GAW) 15: “The multiplicity problem in linkage analysis of gene expression data - the power of differentiating *cis* and *trans* regulators.”
- Jag Rangrej, Research Assistant (01/2006-04/2007; Co-Supervisor). “Optimal selection of markers for validation from genome-wide association studies”
- Xiaoquan (William) Wen, Programmer (05/2002-11/2004; Primary Supervisor). Methods implementation as software STEPC - STatistical Explanation for Positional Cloning.

## Undergraduate Students

*The Dalla Lana School of Public Health does not have an undergraduate program, so the undergraduate supervision started after moving to the Department of Statistical Sciences (75% FTE), Faculty of Arts and Science in 2014.*

- Ryan Ma, Undergraduate Student interested in Statistics, President’s Scholar of Excellence Award (PSEP) Mentorship Match. (11/2016-12/2017; Academic Advisor)
- Rose Garrett, Undergraduate Student majored in Statistics and Ecology & Evolutionary Biology (04/2016-08/2016; Primary Supervisor). “Statistical methods for genetic studies of complex human traits”. Undergraduate Student Research Award, Natural Sciences and Engineering Research Council of Canada (NSERC).
- Xiaotong Wang, Undergraduate Student majored in Statistic and Immunology (04/2016-07/2016; Primary Supervisor). “A statistical analysis of association between physiological trait and human sexual orientation”.

## Statistical Mentoring in Weekly Interdisciplinary Lab Meetings (since 2010)

*Trainees in this section are MSc and PhD level graduate students, and PDFs or Research Assistants/Associates of my collaborators. I provide statistical mentoring during the weekly interdisciplinary lab meetings (1.5-2 hr/week/lab) and manuscripts/conference abstracts preparation periods.*

- Dr. Lisa Strug’s Lab - Genetic studies of Cystic Fibrosis modifier genes.
  - Scott Mastromatteo, MSc level Computer Scientist (01/2018-)
  - Lehang Zhong, MSc level Biostatistician (09/2017-)
  - Bowei Xiao, MSc level Biostatistician (09/2014-)
  - Gengming He, MSc level Biostatistician (09/2014-)
  - Naim Panjwani, MSc level Biostatistician (06/2014-)

- Jiafen Gong, PhD level Biostatistician (06/2012-)
- Weili (Liz) Li, MSc level Biostatistics (09/2009-11/2015)
- Melissa Miller, PDF in Genetic Epidemiology (01/2012-10/2014)
- Theodore (Ted) Chiang, Bioinformatician (10/2007-03/2012)
- Dong Su, MSc level Biostatistician (09/2011-08/2012)
- Xin Li, MSc level Biostatistician (09/2010-10/2011)
- Dr. Andrew Paterson’s Lab - Genetic studies of type 1 diabetic complications.
  - Sareh Keshavarzi, PDF in Genetics (09/2016-)
  - Jingjing Cao, MSc level Biostatistician (09/2016-)
  - Delnaz Roshandel PDF in Genetics (09/2014-)
  - Linda Hiraki, PDF in Genetics (07/2013-06/2017)
  - Mohsen S. Hosseini, PDF in Genetics (10/2010-12/2015)
  - Karen M. Eny, PDF in Genetics (10/2010-10/2014)
  - Hoi Wong, MSc level Biostatistician (05/2012-10/2014)
  - Zhijian (Charlie) Chen, PDF in Statistical Genetics (07/2010-09/2013)

### **Committee Member**

- Mirlinda Kachuri, PhD in Epidemiology (07/2013-03/2018). “Investigation of genetic profiles in chromosome 5p15.33 and telomere length in lung cancer risk and clinical outcomes”.
- Weili (Liz) Li, PhD in Biostatistics (09/2010-11/2015). “Pure likelihood-based methods for genetic association studies”.
- Merav Yarkoni-Abitbul, PhD in Dentistry (01/2008-09/2014). “Expression profiling of chronic pain genes in mice using a whole genome approach”.
- Yan Lu, MSc in Dentistry (05/2005-06/2006). “Polymorphisms in Catechol-O-Methyl-transferase gene affect chronic post-mastectomy pain syndrome”.
- Juan Pablo Lewinger, PhD in Statistics (09/2001-02/2004). “Family-based nonparametric tests of linkage and association”.

### **Thesis Examination Committee Member**

- Daniel Pu, PhD in Statistics, Department of Mathematics and Statistics, York University. Thesis title: “Dependency network in multivariate analysis” (External thesis examiner, May 2009).