

Curriculum Vitae

Pengcheng Lu

Contact Information

Office Address: Department of Biostatistics  
Vanderbilt University School of Medicine  
1161 21st Avenue South, S-2323 MCN  
Nashville, Tennessee 37232  
Phone: 615-322-6283(Office), 615-343-4924(Fax)  
Email: pengcheng.lu@vanderbilt.edu

Education

08/2004 M.S. in Statistics  
Department of Statistics, Iowa State University, USA  
06/1993 B.S. in Biology  
Department of Biology, Huazhong Normal University, China

Master Degree Thesis

06/2004  
**Title:** Modeling approaches for simulation-based spatial point and count data  
**Advisor:** Dr. Mark Kaiser

Previous Employments

11/2004-11/2007  
**Title:** Microarray Statistician(Assistant Scientist)  
Center for Plant Genomics, Iowa State University Plant Science Institute

**Duties:**

- (1) Provided general statistics consulting strategies to biologists at the Center for Plant Genomics
- (2) Researched and recommended appropriate methodologies for two-color cDNA microarrays, provided analysis plans to experimental designs and data analysis to Investigator's various projects.
- (3) Conducted statistical analysis and data visualization for microarray data; Identified differentially expressed genes by fitting linear mixed models and applying appropriate tests within SAS; Interpreted analytical results and prepared manuscripts for publication
- (4) Managed microarray data using R to meet MIAMI standards; Deposited microarray data to GEO database
- (5) Performed statistical analysis for non-microarray biological research data

08/2004-10/2004

**Title:** SAS Consultant

The 1ST Consulting LLC, Iowa

**Duties:**

- (1) Colaborated on SAS programming projects with local companies
- (2) Taught SAS programming training classes held by Iowa SAS User Group and the 1ST Consulting LLC
- (3) Assisted in recruiting SAS programmers for The 1ST Consulting LLC and its partner companies

## Current Position

11/2007-Present

**Title:** Biostatistician II

Department of Biostatistics, Vanderbilt University School of Medicine

**Current Percent Efforts:**

- (1) Vanderbilt-Ingram Cancer Center (41%, PI's: David Carbon, Robert Coffey, Pietenpol, Zaika, etc.)
- (2) Department of Ophthalmology (34%, PI: Penn John)
- (3) Genetics and Endocrine Pathways Linking Obesity with Prostate Cancer (10%, PI: Fowke Jay)
- (4) Silvio Conte Neuroscience Research Center (15%, PI: Blakely Randy)

**Current Key Functions:**

- (1) Provide statistical expertise in analyzing both high-dimensional data and general medical research data
- (2) Meet with Investigators and Supervisors, form analysis plans, provide data analysis strategies for various projects
- (3) Conduct data verification, correction, manipulation, data re-sampling, simulating and missing data imputation, visualize data with high quality graphics
- (4) Perform data analysis following reproducible research principles, using R, Sweave and Latex. Methods and types of analyses include descriptive statistics, power and sample size calculation, parametric and nonparametric testing, logistic regression, Cox proportional hazard regression, Proportional Odds model, trend testing, Linear and non-linear mixed model, parameter selection, model validation, and others.
- (5) Present and interpret statistical results at the research meetings. Participate in writing abstracts, posters, grants, and manuscripts.
- (6) Perform statistical analyses for microarray data using existing BioConductor packages and/or write R scripts to identify differentially expressed genes, perform gene set enrichment analysis, perform Pathway analysis, classification and clustering analyses, form SOP for daily job; Fit statistical models to gene expression data combined with survival outcome data for cancer research projects, adjusting p values for multiple testing. Monitor and control microarray data quality.
- (7) Query microarray data from GEO database using R system to execute meta analysis and/or objective analysis for publication or grant applications; Provide consulting and supervision related to GEO db.
- (8) Explore, research innovative statistical methodologies for challenging projects, provide cutting-edge statistical method support to investigators, develop R package for multiple comparison methodology research.

## Research Projects

### **Department of Ophthalmology**

1. Age-related Macular Degeneration Research

(Dr. Anita Agarwal)

Age-related macular degeneration (AMD) is the most common cause of legal blindness in people over the age of 55 in the US. The cause of AMD has not been determined, but it is clearly a complex disease with multiple environmental and genetic risk factors. This research focuses on evaluating the associations between AMD and the risk factors with environmental influences.

Provided statistical analysis for counts and events data, generated graphics. An abstract was submitted to ARVO in 2008

2. AMD Study

(Drs. Kasra A. Rezaei, Jiyang Cai)

Conduct group comparison between AMD and control samples. Fit logistic regression and calculate power and sample size. Fit other models to multiple data sets, prepare graphics and provide statistics consulting. Submit an abstract to ARVO in 2008 and the American Academy of Ophthalmology Meeting in 2008, Submitted an abstract to Retina Congress 2009.

3. Intermittent Exotropia Study

(Dr. Jeffery Colburn)

Test the association of surgical success and surgical procedures (blrc and r&r), fit logistic regression model to find significant predictors. Perform survival analysis for testing the relationship between follow up time and predictors. Linear trend test and Chi square test applied too.

Submitted an abstract to ARVO 2008.

Analyzed the incidence rate difference and ratio of the outcome Amblyopia and StrabismusHgBA1C. Applied survival analysis methods to the survival data set. Provided statistical consulting for his presentation in 2009.

4. Study the relationship between Asymptomatic Retinal Arteriolar Emboli and All-Cause Mortality in Type II Diabetic Patients

(Dr. Wright B. Lauten)

Revealed the relationship between patient's status and the prior CVA levels. Performed survival data analysis and prepared graphics for presentation and publish. Provided statistical interpretation and consulting to the analysis methods and results.

Submitted an abstract to the ARVO 2008. A manuscript is in reparation.

Provided statistics consulting and help with his new datasets in 2009.

5. Consultant group EMR and Freehand Comparison Study

(Dr. Keshini Parbhu)

Compare the billing levels between the Freehand group and EMR group. Perform trend test for the relationship between billing levels and groups. Apply proportional odds model by adjusting for covariates. Submitted an abstract to ARVO 2008.

6. HbA1c Study

(Dr. Lawrence Merin)

Performed descriptive data analysis, provided statistical consulting.

7. The Study of Intraocular Pressure Changes over Time Following Surgical Interventions  
(Dr. Tarantola Ryan, Dr. Joos and Dr. Agarwal)  
Performed statistical data analysis and data visualization, prepared graphics for publish. A paper is under preparation.
8. A Study of HgBA1C and Progress Degrees  
(Dr. David Reichstein)  
Performed statistical data analysis and data visualization, provided statistics consulting.  
Submitted an abstract to ARVO 2008.  
Conducted statistical analysis to new data set. Applied chi-squared test and Fisher's exact test to various count data comparisons in 2009.
9. Compare the Effects between an Investigational Device and the Habitual Glasses by Studying the Work Distance and LogMAR Measurements  
(Dr. Jeffery Sonsino)  
Provided statistics consulting on experiment design and analysis plan, performed power analysis and sample size calculation, conducted data analysis and data visualization. Submitted an abstract to ARVO 2008.
10. A study of Efficacy of Online Resources for the Pediatric Ophthalmology Population  
(Dr. Desiree Ong)  
Conducted data analysis and data visualization on a survey data set, provided statistics consulting to her presentation in 2009.
11. Comparison of the Grading of Cataracts between Five Experts  
(Dr. Marc Moore)  
Conducted Kappa tests to updated grading data set. Provide statistics consulting to his presentation in 2009.
12. A Survival Study of Eye Pressure  
(Dr. Maziar Lalezary)  
Conducted survival analysis to the data set with sample size 35. Provide statistics consulting and prepared graphics for his presentation in 2009.

### **Institute for Medicine and Public Health**

1. Genetics and Endocrine Pathways Linking Obesity with Prostate Cancer  
(PI: Dr. Jay Fowke)  
Research team appointed Biostatistician, will provide statistical analysis supports to the project.

### **Silvio O. Conte Neuroscience Research Center**

1. Rostral and Caudal Serotonin Neuron Gene Expression Profiles  
(PI: Evan Deneris, Blakely Randy)  
Analyzed gene expression data, identify significantly changed genes among samples from different regions of brain with serotonin neuron and non-serotonin neuron. Perform gene set enrichment analysis and clustering analysis, generate graphics for publish. Provide statistical consulting and input statistical thoughts to manuscript. A manuscript is in preparation.

2. Association study between genes and total life levodopa among Parkinson disease samples  
(PI: Christina Konradi)  
Check microarray data quality, provide quantity and visualization results. Compare the normalization method of quantile in BioConductor and dChip method. Perform differentially expressed genes identification and gene set enrichment analysis. Provide statistical consulting.
3. Heat Shock siRNA Experiment  
(Dr. Bing Zhang)  
Identify differentially expressed genes by applying limma package and perform one-way ANOVA analysis. Perform clustering analysis and compare clustering outcomes from different heat map functions/packages.
4. Post-exercise Heart Rate Recovery of HEP Mice Experiment  
(Brett English)  
Propose a conditional linear regression model to time-dependent covariates, generate graphics and provide statistics consulting.

## **Vanderbilt-Ingram Cancer Center**

1. Circulating Gut- or Skin-Homing Regulatory T Cells (Tregs) Predict Whether Acute Graft-versus-Host Disease (aGVHD) Occurs in Gut or Skin Following Allogeneic Stem Cell Transplantation  
(Dr. Jagasia Madan, Brian Engelhardt)  
Provided microarray gene expression data analysis and model fitting strategies, interpreted results and prepared abstract for conference and manuscript, graphics for publishing. Submitted an abstract to the 50<sup>th</sup> ASH (American Society of Hematology) Annual Meeting and Exposition in 2008. A paper is in preparation.
2. LTED Study  
(Dr. Todd Miller)  
Performed gene expression data analysis and got winners for case (LTED) and control comparison. Conducted clustering analysis for interested genes.
3. A Specific Gene ALCAM Expression Analysis  
(Dr. Andries Zijlstra)  
Provided gene expression data analysis for experiments from GEO database and verify winners with Zijlstra Lab's interested genes. A grant proposal has been submitted.
4. Metastasis Gene Expression Profile Predicts Recurrence-Related Death in Human Colorectal Carcinoma  
(Dr. Joshua Smith)  
Performed gene expression data analysis and microarray data validation for this project. Validated the winner found from cell lines by applying re-sampling and bootstrap methods, checked the array qualities and verified statistical discoveries. Conducted survival data analysis, prepared graphics and statistical summary tables for manuscript. Provide statistics consulting. A manuscript is in preparation.
5. Rectal Microarray with Survival Study  
(Drs. Joshua Smith, Edwards Christina)  
Formed analysis plan, performed data analysis and get winners for paired groups, conducted expression data analysis with survival data, generated graphics and provided statistics consulting.

6. Integral Activity of the P53 Family and Its Role in the Progression of Colon Tumors  
(Dr. Alex Zaika)  
Provided statistics consulting, formed analysis plan. Conducted tissue array data analysis and generated graphics as requested.
7. GEO Meta Data Sets Analysis  
(Dr. Andrew Yi and Sarki)  
Provided statistics consulting. Conducted microarray data analysis for multiple experiments in GEO database. Studied two interested genes.
8. SPORE in GI Cancer  
(Dr. Robert Coffey)  
Will provide statistics support.
9. SPORE in Breast Cancer  
(Dr. Arteaga Carlos)  
Will provide statistics support.

### Awards and Honors(Selected)

12/2008	Employee Recognition Award for Job Performance with Cancer Center's Microarray Projects
11/2008	The 2 <sup>nd</sup> Place of Award Poster at the Department Poster Contest 2008 Poster Title: What Can Biostatisticians Do for Your Graphics? Poster Authors: JoAnn Alvarez, Pengcheng Lu and Li Wang
03/2008	Employee Recognition Award for CQI Project and Job Performance with Ophthalmology projects and Conte Center's Sorotony Microarray projects
09/2004	Award for Volunteer Services at the Mid-West SAS Users Meeting 2004 in Chicago, Illinois
03/2004	Scholarship of Professional Training in SAS Programming(3-Day Class) in Des Moines, Iowa Awarded by the Iowa SAS Users Group

### Professional Memberships

03/2003-Present	American Statistical Association, Member
01/2003-11/2007	Iowa SAS Users Group, Member

### Statistical Softwares

SAS(SAS Certified Base Programmer by SAS Institute Inc. in 2004), R, S-Plus, BioConductor, Stata, SPSS, PS, nQuery Advisor, WFCCM, GSEA, WebGestalt, Ingenuity

### Reproducible Research Skill

R, Sweave and Latex

## Intramural Activities

07/2008-Present	High-Dimensional Data Center (HDDC) at the Division of Cancer Biostatistics Vanderbilt University, Member
01/2008-Present	Participate in Master Biostatistician Interviewing and Recruiting
12/2007-Present	Medical Center Community Survey Biostatistics Department Community Survey CQI Project, Leader
11/2007-Present	Department Book Club, Member
11/2007-Present	Participate in Department Biostatistics Clinics (Omics and others)
11/2007-Present	Participate in Department Weekly R Clinics

## Presentations

08/2009	GCRC Weekly Biostatistics Workshop Co-Lecturer at Vanderbilt University Topic: An Introduction to Choosing Higher Quality Graphics
05/2009	VICC Cancer Biostatistics Center Monthly Biostatistics Workshop Co-Lecturer at Vanderbilt University Topic: Data Mining at GEO Database and Beyond
05/2006	Microarray Course Related Seminar at Iowa State University Topic: FWER and Bayesian Method for Multiple Comparisons in Microarray Data Analysis

## Continuing Education: Course Work/Short Course

11/2008	Clinical Trials by Dr. Yu Shyr at Vanderbilt University, Nashville, TN
09/2008	Biostatistics I by Daniel Byrne at Vanderbilt University, Nashville, TN
04/2008	Biostatistics II by Dr. William Dupont at Vanderbilt University, Nashville, TN
03/2008	Design and Analysis of Genome-wide Association Study at ENAR Conference 2008 in Washington DC
03/2008	Receiver Operating Characteristics Curves for the Uninitiated at ENAR Conference 2008 in Washington DC
03/2008	Attended various sessions at ENAR Conference 2008 in Washington DC
01/2008	Regression Modeling Strategies by Dr. Frank Harrell at Vanderbilt University, Nashville, TN
08/2007	Attended various sessions at UseR! 2007 Conference at Iowa State University, Ames, Iowa
07/2007	Attended various sessions at JSM 2007 Conference in Salt Lake City, Utah
05/2007	Web Course: Advanced SAS Programming by SAS Institute Inc.
04/2004	SAS SQL Procedures by SAS Institute Inc. and Iowa SAS Users Group at Des Moines, Iowa

## Continuing Education: Workshop/Seminar

01/2008-Present	VICC monthly Biostatistics workshop at Vanderbilt University
11/2007-Present	Biostatistics Department weekly seminar at Vanderbilt University

11/2007-Present	Biostatistics Department monthly master continuing education workshop
11/2007-Present	GCRC weekly workshop (Biostatistics related) at Vanderbilt University
11/2007-Present	Kennedy Center workshop (Biostatistics related) at Vanderbilt University
11/2007-Present	Biomedical Informatics weekly seminar (selected) at Vanderbilt University
07/2008	Role of R in Statistical Teaching and Learning (I, II, III) by Dr. Joseph Lu at the Cancer Biostatistics Center Vanderbilt University
05/2008	Agilent's Genomics Road Show by Agilent Technologies in Nashville, TN
05/2006	NCBI Field Training at Iowa State University, Ames, Iowa
11/2004-11/2007	Biological Statistics weekly seminar at Iowa State University
09/2004	SAS Macro Programming by Mid-West SAS Users Group Meeting in Chicago, Illinois
01/2002-11/2007	Statistics Department weekly seminar at Iowa State University
01/2003-07/2007	IOWA SAS Users Group's monthly meeting in Des Moines, Iowa
01/2002-07/2004	Ecological and Environmental Statistics weekly meeting at Iowa State University

## Publications In Print

1. Ke-Feng Lei, Yan-Fang Wang, Xiao-Qun Zhu, **Peng-Cheng Lu**, Bing-Sheng Sun, Hu-Liang Jia, Ning Ren, Qing-Huai Ye, Hui-Chuan Sun, Lu Wang, Zhao-You Tang, Lun-Xiu Qin, "Identification of MSRA gene on chromosome 8p as a candidate metastasis suppressor for human hepatitis B virus-positive hepatocellular carcinoma", *BMC Cancer* 2007, **7**:172

## Publications Submitted/In Preparation

1. J. Joshua Smith, M.D., Fei Wu, M.D., Ph.D., Natasha G. Deane, Ph.D., Bing Zhang, Ph.D., Nipun B. Merchant, M.D., Aixiang Jiang, M.S., **Pengcheng Lu, M.S.**, J. Chad Johnson, M.D., M.S., Carl Schmidt, M.D., M.S.C.I., Steven Eschrich, Ph.D., Christian Kis, M.S., Shawn Levy, Ph.D., Kay Washington, M.D., Ph.D., Martin Heslin, M.D., Robert J. Coffey, M.D., Timothy J. Yeatman, M.D., Yu Shyr, Ph.D. and R. Daniel Beauchamp, M.D., "An Experimentally-derived Metastasis Gene Expression Profile Predicts Recurrence-Related Death in Human Colorectal Carcinoma", will be submitted to *Nature Genetics* soon.
2. Aixiang Jiang, **Pengcheng Lu**, William Wu, Yu Shyr. "Local False Discovery Rate Estimation Based on Bootstrap Null Distribution of Control Group Samples", in preparation.
3. Brian G. Engelhardt, Madan H. Jagasia, Michael T. Rock, Nikki L. Bratcher, John P. Greer, Aixiang Jiang, Adetola A. Kassim, **Pengcheng Lu**, Friedrich G. Schuening, Sandra M. Yoder, James E. Crowe, Jr. "Regulatory T cell expression of cutaneous lymphocyte antigen at engraftment determines organ-specific acute graft-versus-host disease outcomes following allogeneic stem cell transplantation", in preparation.
4. Ryan M. Tarantola, Chun Li, **Pengcheng Lu**, Anita Agarwal, Karen M. Joos, "Long-term Results of Endoscopic-Assisted Pars Plana Vitrectomy for Glaucoma Tube Shunt Placement", in preparation.
5. Christi Wylie, Lily Wang, Bing Zhang, **Pengcheng Lu**, Jay Snoddy, Evan Deneris, "A library of CNS serotonin neuron enriched markers", in preparation.



6. Wright Lauten, Kevin Wise, Chun Li, **Pengcheng Lu**, Rehan Ahamed, Lawrence M. Merin, Amy S. Chomsky, “Restrospective Review: Relationship between Asymptomatic Retinal Arteriolar Emboli and All-Cause Mortality in Type II Diabetic Patients”, in preparation.
7. Kazuhiro Ohtsu, Lu Gao, **Pengcheng Lu**, Dan Nettleton, Patrick S. Schnable, “A Global Analysis of Gene Expression in Histological Layers of the Shoot Apical Meristem of Maize”, in preparation.
8. Sudhansu Dash, Michael Miller, **Pengcheng Lu**, Dan Nettleton, Patrick S. Schnable, “Maize Genes Involved in Nitrate Uptake and Metabolism”, in preparation.

## Technical Report

1. Aixiang Jiang, **Pengcheng Lu**, William Wu, Yu Shyr, “Local False Discovery Rate Estimation Based on Bootstrap Null Distribution of Control Group Samples”, VICC(Vanderbilt-Ingrem Cancer Center) Research Shared Resources.

## Acknowledgements

1. Kazuhiro Ohtsu, Marianne B. Smith, Scott J. Emrich, Lisa B. Borsuk, Ruilian Zhou, Tianle Chen, Xiaolan Zhang, Marja C.P. Timmermans, Jon Beck, Brent Buckner, Diane Janick-Buckner, Dan Nettleton, Michael J. Scanlon and Patrick Schnable, “Global gene expression analysis of the shoot apical meristem of maize (*Zea mays* L.)”, *The Plant Journal* 2007 52, 391-404.
2. Ruth A. Swanson-Wagner, Yi Jia, Rhonda DeCook, Lisa A. Borsuk, Dan Lettleton and Patrick Schnable, “All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents”, *PNAS* 2006. 103; 6805-6810.

## Posters and Abstracts

1. Sarkar, Ananda K., Madi, Shahinez, **Lu, Pengcheng**, Scanlon, Michael J., Schnable, Patrick S., Timmermans, Marja CP., “Organ Polarity in Maize is Regulated by Small RNAs”, *Proceedings of 49<sup>th</sup> Annual Maize Genetics Conference*, St. Charles, Illinois, 2007
2. Cao, Jun, Emrich, Scott J., **Lu, Pengcheng**, Borsuk, Lisa A., Schnable, Patrick S., “Programmed Cell Death Genes Are Differentially Expressed in Tapetal Cells of cms-T Maize During Fertility Restoration”, *Proceedings of 49<sup>th</sup> Annual Maize Genetics Conference*, St. Charles, Illinois, 2007
3. Ohtsu, Kazuhiro1; Smith, Marianne B.; Borsuk, Lisa A.; **Lu, Pengcheng**; Emrich, Scott J.; Zhou, Ruilian; Chen, Tianle; Zhang, Xiaolan; Jin, Hailing; Chen, Hsin D.; Brooks, Lee; Beck, Jon; Buckner, Brent; Janick-Buckner, Diane; Timmermans, Marja C. P.; Scanlon, Michael J.; Nettleton, Dan; Schnable, Patrick S.; “Global Expression Analyses of Genes Involved in Meristem Organization and Leaf Initiation”, *Proceedings of 49<sup>th</sup> Annual Maize Genetics Conference*, St. Charles, Illinois, 2007
4. Brian Engelhardt, MD, Madan Jagasia, MD, MS, Michael T Rock, PhD, Nikki L Bratcher, BS, John P. Greer, MD, Aixiang Jiang, MS, Adetola A. Kassim, MD, MS, **Pengcheng Lu, MS**, Friedrich

Schuening, MD and James E Crowe Jr., MD. "Circulating Gut- or Skin-Homing Regulatory T Cells (Tregs) Predict Whether Acute Graft-Versus-Host Disease(aGVHD) Occurs in Gut or Skin Following Allogeneic Stem Cell Transplantation", *50<sup>th</sup> ASH Annual Meeting and Exposition*, 2008

5. Janice C. Law, MD; Kasra Rezaei, MD; **Pengcheng Lu, BS**; Jiyang Cai, MD, PhD; Paul Sternberg, Jr., MD; Pilot Study on Adverse Events During Procedural Sedation in Overweight and Nonoverweight Children in the Pediatric Emergency Department, *Annual Pediatric Academic Society Meeting 2008*, May 2008. *2008 Joint Meeting of the American Academy of Ophthalmology and European Society of Ophthalmology (SOE)*, November 2008.
6. Bing Zhang, Lily Wang, **Pengcheng Lu**, Dexter Duncan, Jay Snoddy. "Bioinformatics and Biostatistics Core Silvio O. Conte Neuroscience Research Center at Vanderbilt". *2008 Conte Symposium* at Vanderbilt University, 2008.
7. JoAnn Alvarez, **Pengcheng Lu** and Li Wang, "What Can the Biostatisticians Do for your Graphics?", *Department Poster Contest 2008*, October 2008.
8. Ruth A. Swanson-Wagner, Yi Jia, Rhonda DeCook, Tieming Ji, Sarah K. Hargreaves, Mitzi Wilkening, **Pengcheng Lu**, Xuefeng Zhao, Dan Nettleton, Patrick S. Schnable, "Wide-Spread Differential Gene Expression In Reciprocal Crosses Of Maize", *Plant and Animal Genome XVII Conference*, January 2009