

## **SUSMITA DATTA**

Department of Bioinformatics & Biostatistics  
School of Public Health and Information Sciences  
University of Louisville  
481 East Gray Street  
Louisville, KY 40292

### **EDUCATION**

1995 PhD University of Georgia (Statistics)  
1990 MS University of Georgia (Applied Statistics)  
1986 BS University of Calcutta, India (Physics)

### **POSITIONS**

#### **2005-present University of Louisville, Louisville, KY**

2010(June)-present University Scholar  
2010(January)-present Professor (Tenured), Department of Bioinformatics & Biostatistics  
2005-2009 Associate Professor (Tenured), Department of Bioinformatics & Biostatistics

#### **1997-2005 Georgia State University, Atlanta, GA**

2002-2005 Associate Professor (Tenured), Department of Mathematics and Statistics  
2002-2005 Associate Professor, Department of Biology (joint appointment)  
1997-2002 Assistant Professor, Department of Mathematics and Statistics

#### **1995-1997 Emory University, Atlanta, GA**

1995-1997 NRSA Post- Doctoral Fellow, Department of Biostatistics

#### **2000-2001 University of Georgia, Athens, GA**

2000(Fall)-2001(Summer) Visiting Assistant Professor, Department of Genetics.

### **RESEARCH INTERESTS**

Methodological: Bioinformatics, Clustering and Classification, Genomics, Proteomics, Infectious Disease Modeling, Non-linear Regression modeling for Systems Biology, Statistical Issues in Population Biology, Statistical Genetics, Systems Biology, Survival Analysis and Multi state models.

Clinical: Birth Defects Research

### **FUNDED RESEARCH GRANTS AND CONTRACTS**

#### **PI level**

1. PI (no co-PI) R15, 25% effort, National Institute of Health, NCI R15 CA133844-01. Development of Statistical Methods for Analyzing Proteomic Cancer Data, 2009-2012.

2. PI (no co-PI) NSF, 17% effort, National Science Foundation, Statistics Program (DMS), Standard Grant, Statistical peak detection, adaptive classification and protein-protein network construction using mass spectra, DMS-0805559, 2008-2011.
3. PI U of L subcontract (E. Voit, PI, Georgia Tech.) 10% effort, The Trehalose Cycle as Paradigm, National Science Foundation, 2006-2009.
4. PI 30% effort, Pilot Project, Proteomics Based Approach for Early Detection of Fetal Alcohol Syndrome, P20-RR/DE17702, NIH COBRE (PI, R. Green), 2006-2007.
5. PI (no co-PI) NSF (DMS) Statistical Analysis of Microarray Gene Expression Data, \$127,671, 2000-2002.
6. PI (no co-PI), Research Experience for Undergraduates in Fungal Genomics and Computational Biology: GSU VPRSP grant, \$18,666, Summer 2001.
7. PI (no co PI) A Pilot Project for Developing Statistical Tools for Bioinformatics. GSU faculty initiation grant, \$5000, 2000-2001.

#### **Non-PI level**

8. Co-I R01 NIDDK/NIH 10% effort biomarker validation and dosing algorithms for anemia management in ESRD 12/01/2011-11/30/2016.
9. Co-I R01, 8% effort, NIDDK/NIH DNA Sequences Impact Estrogen and Anti-Estrogen Activities, NIH R01 DK053220-10A1, (Carolyn M. Klinge, PI) 2008-2013.
10. Co-I R01, 8% effort, NCI/NIH, Estrogen Signaling on mitochondrial function R01 DK053220-10A1 (PI Klinge, Carolyn M.), 2/1/2009-1/31/14
11. Co-I, R01 U of Louisville Subcontract, 5% effort, NIH, Environmental Interactions with the Genome and Epigenome. 2008-2010.
12. (Role: Co-I, Mentor), NCRI 3P20RR016481- 09S1 KY IDEA Networks of Biomedical Research Excellence, RC2 proposal. (Cooper, PI), 4/1/2009-3/31/2011
13. Co-director of Biostatistics Consortium for the Research Development Core 15% effort submitted CTSI proposal, 2010.
14. Biostatistics Advisor on Bioinformatics core, 10% effort submitted CTSI proposal, 2010.
15. Co-I/Biostatistician, 10% effort, NIH P01 AA017103-01, Alcohol, Liver Disease and Alcohol Nutrient Interactions, (Craig McClain, PI), 2008-2013.
16. Biostatistics Group Leader, Bioinformatics, Biostatistics and Computational Biology Core, Center for Environmental Genomics and Integrative Biology, P30ES014443-01A1 (K. Ramos, PI), 10% effort, NIEHS-NIH, 2007-2011.
17. Co-I (M. J. Kennedy, PI, Louisville) 5% effort, Aminoglycoside Urinary Proteomics, 2007-2009.
18. Biostatistician (J. Klein, PI, Louisville) 10% effort, Pediatric Clinical Proteomics Center, Department of Energy, 2005-2008.
19. Co-I (P. Epstein, PI, Louisville) 10% effort, NIH R01, Podocytes and Oxidative stress in diabetic Kidney, 2006-2007.

20. Co-PI (K. B. Grant, PI) Brains and Behavior Seed Grant, GSU, \$25414, 2005-2006.
21. Co-PI (I. Weber, PI) Research Program Enhancement Award, GSU, student support, \$36000, 2004-2007.
22. Investigator, Student & Travel support for five years, \$75000, Georgia Cancer Coalition (Michael Eriksen PI), 2004-2009.
23. Statistician (2 months of summer salary), BimCore, Emory University, summer 2004.
24. Co PI (M. Brinton, PI) Biomedical Computing Center Seed Grant, GSU, \$13467, summer 2004.
25. Consultant on a NSF funded project in Structural Biology (B. C. Wang, PI), University of Georgia, \$8087, summer 2003.
26. Co-PI (J. Arnold, PI) Genomics and Computational Biology: A REU Site, National Science Foundation, Joint Program between UGA, GA State and Clarke Atlanta University, \$210,000, 2003-2005.
27. Co-PI (G. Chen, PI) Tech Fee Grant , GSU, \$58002, 2003.
28. Co PI (D. Vidacovic, PI) Instructional Improvement Grant, GSU Center for Teaching & Learning, \$5000, 2000-2001.
29. Co PI (E. Dubinsky, PI) IPCURT Project Course and Curriculum Development, National Science Foundation, \$100,000, 1998 - 1999.

## PUBLICATIONS

### Five Most Cited Papers

(Google Scholar, accessed on 2/13/2012)

1. **Datta, S.** and Datta, S. *Bioinformatics*, 2003, #20 ,Citation Count = 240
2. Gerrit J. Schut, G., J., Brehm, S., D., Datta, S. and Adams, M. W., *Journal of Bacteriology*, July 2003, #22, Citation Count = 107
3. **Datta, S.** and Datta, S. *BMC bioinformatics*, 2006 , # 33, Citation Count = 80
4. Longini, I., Datta, S. and Halloran, E., *Journal of Acquired Immune Deficiency Syndromes & Human Retrovirology*, 1996, Citation Count = 63
5. Weinberg, M. W., Schut, G., J., Brehm, S., D., Datta, S. and Adams, M. W., *Journal of Bacteriology*, 2005, #27, Citation Count= 51

### Refereed papers and book chapters:

#### 2012

Chakraborty, S., Datta, S. and **Datta, S.** Surrogate variable analysis using partial least squares (SVA-PLS) in gene expression studies. *Bioinformatics*, to appear (Available under Advanced Access, doi: 10.1093/bioinformatics/bts022).

Mostajabi, F., Datta, S., **Datta, S.** Predicting Patient Survival from Proteomic Profiles using Mass Spectrometry Data: An Empirical Study. *Communications in Statistics - Simulation and Computation*, to appear.

Li, X., Klinge, C. M., Datta, S. Novel and Alternative Bioinformatics Approaches to Understand miRNA-mRNA Interactome in Cancer Research. *Systems Biology in Cancer Research and Drug Discovery* (Azmi, A. ed) Springer, to appear.

Datta, S. and Pihur, V. Feature selection and machine learning with mass spectrometry data (Revised), In *Clinical Proteomics: Methods, Applications and Tools, Mass spectrometry Data Analysis in Proteomics (Second Edition)* (Matthiesen, R. ed.), Humana Press, to appear.

**Datta, S.**, Shah, J. and Datta, S. A multi-objective ensemble regression method (MERM) for high dimensional data mining. Preprint.

Litchfield, L. M., Riggs, K. A., Emberts, C. G., Hockenberry, A. M., McConda, D. B, Oliver, L. D., Fox, J. M., Hu, C., Cai, J., Pierce, W. M., Jr., Ivanova, M. M., Bates, P. J., Martin, R. G. C., Appana, S. N., **Datta, S.**, Kulesza P., and Klinge, C. M. Regulation of COUP-TFII transcriptional activity by interaction with nucleolin in human breast cancer cells and tumors. Preprint.

Stallons, L. J., Kalbfleisch, T. Cambon, A., **Datta, S.**, Kucherlapati, R., Kunkel, T. A. and WG McGregor. DNA polymerase iota alters PLK signaling for mitotic entry and exit after UV. Preprint.

## 2011

Ndukum, J., Fonseca, L. L., Santos, H., Voit, E., O. And **Datta, S.** (2011). Statistical Inference Methods for Sparse Biological Time Series Data, *BMC Systems Biology*, **5**, 57. PMC3114728

Li, X, Gill, R., Cooper, N.,G.,F., Yoo, J., K., and **Datta, S.** (2011). Modeling microRNA-mRNA Interactions Using PLS Regression in Human Colon Cancer, *BMC Medical Genomics*, **4**, 44. PMC3123543 (**Highly Accessed**)

Ndukum, J., Atlas, M., **Datta, S.** (2011). pkDACCLASS: open source software for analyzing MALDI-TOF, *Bioinformatics*, **6**, 45-47. PMC3064853

Manavalan, T. T., Teng, Y., Bhimani, S., Appana, S. N., Datta, S., Kalbfleisch, T. S., Li, Y., and Klinge, C. M. (2011). Differential expression of microRNA expression in tamoxifen-sensitive MCF-7 versus tamoxifen-resistant LY2 human breast cancer cells, *Cancer Letters*, **313**, 26-43.

Pihur, V., Datta, S., **Datta, S.** (2011). Meta analysis of Chronic Fatigue Syndrome through integration of clinical, gene expression, SNP and proteomic data, *Bioinformatics*, **6**, 120-124. PMC3089886

## 2010

**Datta, S.**, Pihur, V. and Datta, S. (2010). An adaptive optimal ensemble classifier via bagging and rank aggregation with applications to high dimensional data, *BMC Bioinformatics*, **11**, 427.

Gill, R., Datta, S. and **Datta, S.** (2010). A statistical framework for differential network analysis from microarray data using partial least squares, *BMC Bioinformatics*, **11**, 95. PMC2838870

**Datta, S.**, Datta, S., Kim, S., Chakraborty, S. and Gill, R. S Statistical Analyses of Next Generation Sequence Data: A Partial Overview (2010). *Journal of Proteomics & Bioinformatics*, **3**, 183-190. PMC2989618.

## 2009

Pihur, V., Brock, G., **Datta, S.** and Datta, S. Cluster validation for microarray data: An appraisal. (2009). In *Multivariate Statistical Methods*, ( A. SenGupta, ed), ISI Platinum Jubilee series, Vol 5, World Scientific Press, pp. 79-94.

**Datta, S.** and Datta, S. (2009). Computational biology touches all bases. *Genome Biology*, **10**, 303.

Pihur, V., **Datta, S.** and Datta, S. (2009). RankAggreg, an R package for weighted rank aggregation. *BMC Bioinformatics*, **10**, 62. (**Highly Accessed**)

Yoo, J. K., Becky S. Patterson, B. S. and **Datta, S.** (2009). OLS-based predictor test in single index model to predict transcription rate by histone acetylation level, *Statistics & Probability Letters*, **79**, 2109-2114.

Atlas, M. and **Datta, S.** Monoisotopic Peak Detection for Mass spectrometry Data (2009). *Journal of Proteomics and Bioinformatics*, **2**, 202-216.

**Datta, S.** and Pihur, V. (2009). Feature selection and machine learning with mass spectrometry data, R. Matthiesen, ed., In *Clinical Proteomics: Methods, Applications and Tools*, Humana Press, (Matthiesen, R. ed.), pp. 205-229.

## 2008

Brock, G., Pihur, V., **Datta, S.** and Datta, S. (2008). clValid, an R package for cluster validation. *Journal of Statistical Software*, **25**, 4.

Pihur, V., Datta, S. and **Datta, S.** (2008). Finding cancer genes through meta-analysis of microarray experiments: Rank aggregation via the cross entropy algorithm. *Genomics*, **92**, 400-403.

Pihur, V., Datta, S. and **Datta, S.** (2008). Reconstruction of genetic association networks from microarray data: A partial least squares approach. *Bioinformatics*, **24**, 561-568.

**Datta, S.**, Turner, D., Singh, R., Ruset, B., Pierce, W. M., and Knudsen, T. B. (2008). Fetal alcohol syndrome in mice detected through proteomics screening of the amniotic fluid. *Birth Defects Research Part A: Clinical and Molecular Teratology*, **82**, 177-186.

## 2007

**Datta, S.**, Le-Rademacher, J. and Datta, S. (2007). Predicting patient survival from microarray data by accelerated failure time modeling using partial least squares and LASSO, *Biometrics*, **63**, 259-271.

**Datta, S.**, Datta, S., Parrish, R. S. and Thompson, C. M. (2007). Microarray data analysis. In *Computational Methods in Biomedical Research*, R. Khatree and D. Naik, eds., Chapman & Hall/CRC Biostatistics Series, Volume 24, 1-43.

Boratyn, G. M., **Datta, S.** and Datta, S. (2007). Incorporation of biological knowledge into distance for clustering genes. *Bioinformatics*, **1**, 396-405.

Pihur, V., **Datta, S.** and Datta, S. (2007). Weighted rank aggregation of cluster validation measures: A Monte Carlo cross-entropy approach. *Bioinformatics*, **23**, 1607-1615.

Pihur, V., **Datta, S.** and Datta, S. (2007). Understanding Chronic Fatigue Syndrome (CFS) from CAMDA data: A systems biology approach. In *Proceedings of CAMDA 2007*, full paper, online @ <http://camda.bioinfo.cipf.es/camda07/agenda/detailed.html>.

## 2006

**Datta, S.** and de Padilla, L.M. (2006). Feature selection and machine learning with mass spectrometry data for distinguishing cancer and non-cancer samples, *Statistical Methodology* (Special Issue on *Bioinformatics*), **3**, 79-92.

**Datta, S.** and Datta, S. (2006). Validation measures for clustering algorithms incorporating biological information, *IEEE Proceedings of International Multi-Symposiums on Computer and Computational Sciences* (IMSCCS'06), (J. Ni, J. Dongarra, Y. Zheng, G. Gu, G. Wolfgang and H. Jin, Eds.), **1**, 131-135.

**Datta, S.** and Datta, S. (2006). Evaluation of clustering algorithms for gene expression data, *BMC Bioinformatics*, **7** (Suppl 4), S17.

**Datta, S.** and Datta, S. (2006). Methods for evaluating clustering algorithms for gene expression data using a reference set of functional classes, *BMC Bioinformatics*, **7**, 397. (Highly Accessed)

Boratyn, G. M., **Datta, S.** and Datta, S. (2006). Biologically supervised hierarchical clustering algorithms for gene expression data, In *Proceedings of the 28th IEEE EMBS Annual International Conference*, New York City, USA, 5515-5518.

## 2005

**Datta, S.** and Datta, S. (2005). Empirical Bayes screening (EBS) of many p-values with applications to microarray studies, *Bioinformatics*, **21**, 1987-1994.

Weinberg, M. V., Schut, G. J., Brehm, S., **Datta, S.**, and Adams, M. W. W. (2005). A hyperthermophilic cold shock response: the archaeon *Pyrococcus furiosus* synthesizes novel membrane-bound glycoproteins at a sub-optimal growth temperature. *Journal of Bacteriology*, **187**, 336-348.

Datta, S. (2005). Statistics in Genetics, In *Encyclopedia of Statistical Sciences*, Second edition, Wiley, New York.

Datta, S. (2005). Statistics in Microarray Analysis, In *Encyclopedia of Statistical Sciences*, Second edition, Wiley, New York.

Datta, S. (2005). Statistics in Vaccine Studies, In *Encyclopedia of Statistical Sciences*, Second edition, Wiley, New York.

## 2004

**Datta, S.**, Satten, G. A., Benos, D. J., Xia, J., Heslin, M., and Datta, S. (2004). An empirical Bayes adjustment to increase the sensitivity of detecting differentially expressed genes in microarray experiments, *Bioinformatics*, **20**, 235-242.

**Datta, S.** and Datta, S. (2004). An empirical Bayes adjustment to multiple p-values for the detection of differentially expressed genes in microarray experiments. In *Bioinformatics 2004, Conferences in Research and Practice in Information Technology - Second Asia-Pacific Bioinformatics Conference*, **29**, Y-P. P. Chen, Ed., 155-159, Australian Computer Society, Sydney.

Warrenfeltz, Z., Pavlik, S., **Datta, S.**, Kraemer, E., Benedict, B. McDonald, J. F. (2004). Gene expression profiling of epithelial ovarian tumors correlated with malignant potential. *Molecular Cancer*, **3**, 27.

## 2003

Datta, S. (2003). Statistical techniques for microarray data: A partial overview, *Communications in Statistics-Theory and Methods*, **32**, 263-280.

**Datta, S.** and Datta, S. (2003) Comparisons and validation of statistical clustering techniques for microarray gene expression data, *Bioinformatics*, **19**, 459-466.

Arnold, J., Schuttler, H.-B., Logan, D., Griffith, J., Arpinar, B. **Datta, S.**, Kochut, K. J., Kraemer, E., Miller, J. A., Sheth, A., Aleman-Meza, B., Doss, J., Harris, L. and Nyong, A. (2003). Metabolomics, In *Handbook of Industrial Mycology, Chapter 23*. Marcel-Dekker, New York, NY.

Schut, J., G., Brehm, S., **Datta, S.**, and Adams, M. W. W. (2003). Whole Genome DNA microarray of a hyperthermophile and an archaeon: *Pyrococcus furiosus* grown on peptides and carbohydrate, *Journal of Bacteriology*, **185**, 3935-3947.

## 2002

**Datta, S.** and Arnold, J. (2002). Some comparisons of clustering and classification techniques applied to transcriptional profiling data. In *Advances in Statistics, Combinatorics and Related Areas*, Eds.: C. Gulati, Y-X. Lin, S. Mishra, and J. Rayner, World Scientific, 63-74.

## 2001

Datta, S. (2001). Estimation of selection parameters using multi-generation cytonuclear data, *Biometrical Journal*, **43**, 219-233.

Datta, S. (2001). Exploring relationships in gene expressions: A partial least squares approach, *Gene Expression*, **9**, 257-264.

Datta, S. (2001). Testing neutrality of mtDNA using multigeneration cytonuclear data, Selected Proceedings of the Symposium on Inference for Stochastic Processes, Eds.: I. V. Basawa, C. C. Heyde and R. L. Taylor, *IMS Lecture Notes - Monograph Series*, **37**, 173-184, IMS, Beachwood, OH.

## 2000

Datta, S. (2000). Some statistical aspects of cytonuclear disequilibria. In ***Statistics in Molecular Biology and Genetics***, Ed: Francoise Seillier-Moiseiwitsch, IMS Lecture Notes-Monograph Series, **33**, 21-37.

Datta, S., Satten, G. A. and **Datta, S.** (2000). Nonparametric estimation for the three stage irreversible illness-death model, ***Biometrics***, **56**, 841-847.

Datta, S. (2000). Some statistical issues involving multi-generation cytonuclear data, In ***Advances on Methodological and Applied Aspects of Probability and Statistics***, N. Balakrishnan, Ed., Gordon and Breach, 525-546.

**Datta, S.**, Satten, G. A. and Datta, S. (2000). Estimation of stage occupation probabilities in multistage models, In ***Advances on Theoretical and Methodological Aspects of Probability and Statistics***, N. Balakrishnan, Ed., Gordon and Breach, 493-506.

Datta, S. (2000). Book Review: Statistics in Human Genetics by Pak Sham. ***Statistics in Medicine***, **19**,1384-1385.

## 1999

Scribner, K. T., **Datta, S.**, Arnold, J., and Avise, J. C. (1999). Empirical evaluation of cytonuclear models incorporating genetic drift and tests for neutrality of mtDNA variants: data from experimental *Gambusia* hybrid zones, ***Genetica***, **105**, 101-108.

**Datta, S.**, Halloran, E. M. and Longini, I. M. (1999). Efficiency of estimating vaccine efficacy for susceptibility and infectiousness: randomization by individual versus household, ***Biometrics***, **55**, 792-798.

## 1998

**Datta, S.** and Arnold, J. (1998). Dynamics of cytonuclear disequilibria in subdivided populations, ***Journal of Theoretical Biology***, **192**, 99-111.

## 1997

**Datta, S.**, Longini, I. M., and Halloran, E. (1997). Measuring vaccine efficacy for different HIV vaccine trials, ***Statistics in Medicine***, **17**, 185-200.

## 1996

**Datta, S.**, Fu, Y. X., Arnold, J. (1996). Dynamics and equilibrium behavior of cytonuclear disequilibria under genetic drift, mutation, and migration, ***Theoretical Population Biology***, **50**, 298-324.

**Datta, S.** and Arnold, J. (1996). Diagnostics and a statistical test of neutrality hypothesis using the dynamics of cytonuclear disequilibria, ***Biometrics***, **52**, 1042-1054.

**Datta, S.**, Rand, D. M., and Arnold, J. (1996). A statistical test of a neutral model using the dynamics of cytonuclear disequilibria, ***Genetics***, **144**, 1985-1992.

Longini, I. M., **Datta, S.**, and Halloran, E. (1996). Measuring vaccine efficacy for both susceptibility to infection in infectiousness for prophylactic HIV-1 vaccines, *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*, **13**, 440-447.

#### **Non-Refereed Publications:**

**2005**

**Datta, S.** and Datta, S. (2006). Validation of statistical clustering using biological information, *Proceedings of INTERFACE 2005* (CD-ROM).

1999

Datta, S. (1999). Hypotheses testing for different selection models using multi-generation cytonuclear data, *Proceedings of American Statistical Association, Biometrics Section*, 157-161, Alexandria, USA.

#### **EDITORIAL AND REFERRING SERVICE:**

- 2010- Editorial Board Member, Briefings in Bioinformatics, (Impact Factor 9.283)
- 2008- Associate Editor, BMC Research Notes
- 2007- Associate Editor, Bioinformation
- 2006 Editorial Board Member, Bioinformation
- 2007- Special Issue Editor (Gene Expression Analysis), Bioinformation
- 2007- Associate Editor, Statistical Methodology
- 2007- Associate Editor, Statistics & Probability Letters

#### **Ad-hoc Member/Reviewer:**

- 2012 National Institute of Health, ZRG1 BST-M (02) **Chair of** Special Emphasis Study Section of Topics in Bioengineering Sciences and Technologies.
- 2011 Book proposal review (2011) Title: Advances in Statistical Bioinformatics: Models and Integrative Inference for High-Throughput Data. Publisher: Cambridge University Press.
- 2008 National Science Foundation, Biology Program, Mail Reviewer, May 2008.
- 2003 National Institute of Health, Bio-defense Study Section, April 2003.
- 2003 (Invited) Emtech Bio Scientific Advisory Board members and Seed Grant Reviewers: Georgia Tech, Atlanta, October 2002.
- 2002 (Invited) Emtech Bio Scientific Advisory Board members and Seed Grant Reviewers: Georgia Tech, Atlanta, October 2002.
- 2001 Member, Advisory Panel for MRI Program, National Science Foundation, 2001-2002.
- 2000-2005 Mathematical Review
- 2000 Book review for Statistics and Medicine, 2000.
- 2000- Paper reviewer of multiple ISCB and CAMDA conferences as program Committee members, several years.

#### **Referee work:**

Biometrics, Journal of American Statistical association, Journal of Statistical Planning and Inference, Statistics and Its Interface, Statistics in Medicine, Journal of Applied Statistics, Computational Statistics & Data Analysis, Journal of Multivariate Statistics, Scandinavian Journal of Statistics, Communications in Statistics, Statistics and Interface, Computational

Statistics and Data Analysis, Statistical Applications in Genetics and Molecular Biology, Mathematical Biosciences, Bioinformatics, BMC Bioinformatics, BMC Research Notes, Briefings in Bioinformatics, Journal of Bioinformatics and Computational Biology, BMC Systems Biology, BMC Health Services Research, Proceedings of National Academy of Sciences, Nucleic Acids Research, Proteomics, Biotechnology, Genomics, Journal of Proteome Research, Human Reproduction, Pattern Recognition, International Journal of Data mining and Bioinformatics, Computational Biology and Bioinformatics, Physica A: Statistical Mechanics and Its Applications, Int. J. Developmental Neuroscience, Computer Methods and Programs in Biomedicine, Chemometrics, Computer Methods and Programs in Biomedicine, Environmental Health Perspectives. Moreover, refereed many manuscripts for the conference submissions to Intelligence Systems for Molecular Biology (ISMB) and European Conference on Computational Biology (ECCB) and Critical Assessment of Massive Data Analysis (CAMDA) as a program committee member **multiple times**.

#### **MEMBERSHIPS AT PROFESSIONAL SOCIETIES:**

2011-       Caucus for Women in Statistics (CWS)  
 2006-       American Association for the Advancement of Science (AAAS)  
 2007-       **Elected member of** International Statistical Institute (ISI)  
 2000-       International Biometric Society (ENAR)  
 1999-       International Society for Computational Biology (ISCB)  
 1997-       International Indian Statistical Association (Life) (IISA)  
 1997-       Forum for Interdisciplinary Mathematics (Life) (FCM)  
 1995-       American Statistical Association (ASA)  
 1995-       Institute of Mathematical Statistics (IMS)

#### **SERVICE TO STATISTICAL AND COMPUTATIONAL BIOLOGY ORGANIZATIONS**

2014       Ex-Officio member of ASA committee of Women in Statistics  
 2011-2014 Appointed member of Executive Committee for the Conference of Women in Statistics 2014  
 2011-2013 Elected member of ASA Scientific & Public Affairs Advisory Committee  
 2012-2012 Appointed member of the COWIS-Gertrude M. Cox Scholarship Subcommittee of ASA.  
 2012       CAMDA 2012 Conference Scientific Committee, Long Beach, California.  
 2012-13   Elected President of Caucus for Women in Statistics  
 2011-2011 Elected Representative-at-Large Caucus for Women in Statistics  
 2011       Breakfast Round Table Discussion Coordinator for the Caucus for Women in Statistics, JSM 2011.  
 2011       Co-Chair of the 'Statistical Significance' Poster Competition organized by the ASA Scientific & Public Affairs Advisory Committee, JSM 2011.  
 2011       Program Committee Member of 19th Annual International Conference on Intelligent Systems for Molecular Biology ISMB and ECCB10, The 10th European Conference on Computational Biology, Vienna, July, 15-16, 2011.  
 2010       Program Committee Member ECCB10, the 9th European Conference on Computational Biology, Ghent, Belgium, Sept 26-29, 2010.  
 2009       CAMDA 2009 Conference Scientific Committee, Chicago, IL, U.S.A., October 5-6, 2009.  
 2008       CAMDA 2008 Conference Scientific Committee, Vienna, Austria, December 2008.  
 2008       Program Committee member, Frontiers of Probability and Statistical Science, Connecticut-Storrs, May 2008.  
 2007       CAMDA 2007 Conference Scientific Committee, Valencia, Spain, December 2007  
 2007       Program Committee Member, ISMB 2007, Vienna, Austria, July 2007.  
 2005       Program Committee Member, ISMB 2005, Michigan, July 2005.

2003 Executive Board Member and President of Young Professional Statisticians, International Indian Statistical Association IISA.

## ORGANIZATION OF CONFERENCE SESSIONS

2008 Invited session organizer at **JSM 2008**, A New Paradigm of Statistical Data Analysis: Omics Data, Denver, August 2008.

2007 Program Committee Member, ISMB 2007, Vienna, Austria, July 2007.

2007 Chair, Invited session at **JSM 2007**, Inference for Multistate Data under Complex Censoring Structures, Salt Lake City, July - August, 2007.

2006 Invited session organizer, Statistics in Genomics and Proteomics, International Biometric Society Conference IBC 2006, Montreal, Canada, July, 2006.

2004 Invited session organizer Statistics in Genomics, **JSM Toronto, August 2004**.

2004 Invited session organizer, Genetic Data Analysis, International Conference on Statistics in Health Sciences, Nantes, France, June 2004.

2002 Co-organizer, student paper competition for the IISA conference, Athens, GA, May 2004.

2003 Organized (and chaired) an invited session titled "Recent Contributions in Bioinformatics" at **JSM San Francisco, August, 2003**.

2003 Executive Board Member and President of Young Professional Statisticians, IISA.

2002 Organized an invited session titled Survival Skills for Young Statisticians at the IISA International Conference on Statistics, Probability and Related Areas, DeKalb, Illinois, June 2002.

2002 Organized an invited session on Bioinformatics at SCRA 2002-FIM IX: Ninth International Conference of Forum for Interdisciplinary Mathematics on Statistics Combinatorics and Related Areas, Department of Statistics and Department of Mathematics: University of Allahabad, Allahabad, UP 211 002, India, December 21-23, 2002.

2001 Invited Session Organizer, ENAR, 2001 Joint Statistical Meeting, 2001, Atlanta, Georgia.

2001 Organized an invited session on Statistics in Bioinformatics at the International Conference on Statistics, Combinatorics and Related Areas and the Eighth International Conference of the Forum for Interdisciplinary Mathematics, Wollongong, Australia, December 2001.

2001 Chair, (invited session) Bioinformatics: Statistical Perspectives and Controversies' at International Conference on Statistics, Combinatorics and Related Areas and the Eighth International Conference of the Forum for Interdisciplinary Mathematics December 2001.

2001 Invited Session Organizer, ENAR, 2001 Joint Statistical Meeting, 2001, Atlanta, Georgia.

1999 Session Chair, Statistical Genetics, ENAR Spring Meeting, 1999, Atlanta, Georgia.

1999 Local Organizing Committee, ENAR Spring Meeting, 1999, Atlanta, Georgia.

1999 Session Chair, Applications of State-Space Modeling in the Science, Special Contributed Session, **Joint Statistical Meeting, 1999, Baltimore, Maryland**.

## SERVICE TO ACADEMIC INSTITUTIONS

### @ University of Louisville

2011- Bioinformatics Executive committee member (university level), University of Louisville

- 2010-2008 Primary member of University Faculty Grievance Committee
- 2008 Chair of Bioinformatics Faculty Recruitment Committee
- 2007 Member, Biostatics Director Search Committee of Brown Cancer Center
- 2006-2008 Member, Biostatistics Faculty Recruitment Committees
- 2006-11 Biostatistics Group Leader for the Center for Genomics and Integrated Biology, University of Louisville
- 2006-11 Biostatistics Group Leader for the Alcohol Center at University of Louisville
- 2005-2009 Member of Bioinformatics Task force (university level), University of Louisville
- 2005-2011 Chair of Graduate Admissions Committee (department level)
- 2006- Bioinformatics PhD Program Development Committee (university level)
- 2006- Chair of Bioinformatics Emphasis development for the PhD program in Biostatistics (department level).

**@ Georgia State University**

- 2000-2005 Bioinformatics Executive Committee Member, School of Arts and Sciences GA State University
- 1998-2004 Member of Faculty Recruitment Committees
- 1998 Chair of Biostatistics Program Development Committee
- 1999 Chair of Bioinformatics Program Development Committee

**External Reviewer for Tenure and Promotion Evaluation**

- 2010 Department of Statistics, Purdue University
- 2009 Department of Statistics, University of Missouri, Columbia
- 2007 Department of Biostatistics, Medical College of Georgia

**Short Academic Visits**

- 2011 School of Public Health, University of Tampere, June 2011.
- 2010 Department of Statistics, Southwest Jiaotong University, Chengdu, China, May-June, 2010.
- 2004 Department of Statistics and Applied Probability, National University of Singapore, December 2004.
- 1999 Math Stat Division, Indian Statistical Institute, Calcutta, July 1999.

**HONORS AND AWARDS:**

- 2011-2013 Elected member of ASA Scientific and Public Affairs Advisory Committee
- 2013 Elected President for the Caucus for Women in Statistics
- 2010-2013 Elected Representative-at-large for Women in Statistics
- 2010 (July) Appointed as a University Scholar for outstanding research, scholarship and creative activities
- 2010 Biography inclusion in 2010 (64 th Edition) of Marquis Who's Who in America and then in Who's Who in the world, in Science and Engineering and Who's Who in women.
- 2007 Elected member of International Statistical Institute, October 2007
- 2007 Nominated for Provost's Award for exemplary advising, University of Louisville
- 2003 Appeared in Fox News Atlanta, Comment on State Lottery System
- 2003 Featured Research faculty in College of Arts and Sciences, Georgia State University, Feb., 2003.
- 2002 Co-recipient of the CURO Excellence in Undergraduate Research Mentoring Award from University of Georgia, April 2002.

2001-2003	Executive Board Member and Elected President of Young Professional Statisticians, International Indian Statistical Association (IISA) 2001-2003
2002	Press coverage (Atlanta Business Chronicle), April, 2002
2002	Press coverage (Georgia State Magazine), Fall, 2002
2001	NCI travel award for "Workshops for Junior Biostatisticians, 2001 ENAR", Charlotte, N. Carolina.
2000	Outstanding Junior Faculty Award nomination, Georgia State University, Atlanta, Georgia, April 2000
2000	Phi Kappa Phi honor society, Georgia State University Chapter, acceptance, April 2000
1998	NSF Travel Award for IBC98, Cape Town, South Africa, December 1998.
1998	NSF Travel Award for Pathways to the Future workshop, Dallas, Texas, August 1998
1995	Student paper award in SRCOS/ASA summer conference, Melbourne, Florida, June 1995
1994	Best Theoretical Student Award, Department of Statistics, University of Georgia, Athens, Georgia, 1994

## **INVITED PRESENTATIONS AT NATIONAL OR INTERNATIONAL CONFERENCES**

### **2011**

"Nonparametric regression using partial least squares dimension reduction in multistate models" Statistical Concepts and Methods for the Modern World, Colombo, Sri Lanka, December 28-30, 2011.

"Nonparametric regression using partial least squares dimension reduction in multistate models" 4th International Conference of the ERCIM WG on Computing & Statistics, University of London, UK, December 17-19, 2011.

"Predicting Patient Survival from Proteomic Profile using MALDI-TOF Mass Spectrometry Data", Applied Stochastic Models and Data Analysis (ASMDA 2011), Rome, Italy, June 7 - 10, 2011.

### **2010**

"Rank Aggregation and its Use in Bioinformatics Problems", The First International Conference on Theory and Application of Statistics, Dhaka, Bangladesh, 26-28 December 2010.

"Rank Aggregation and its Use in Bioinformatics Problems". LinStat 2010, Tomar, Portugal, July 27, 2010.

"Monoisotopic Peak Detection and Disease Classification for Mass Spectrometry Data". ENAR Spring Meeting, New Orleans, March 22, 2010.

"Improved Automated Monoisotopic Peak Detection Method for Mass Spectrometry Data", UT-ORNL-KBRIN Bioinformatics Summit, Lake Barkley State Resort Park, Cadiz, KY. March 19, 2010.

"Predicting Survival from High Dimensional Data", International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management, February 10, 2010, Beer Sheva, Israel.

### **2009**

"Improved Automated Monoisotopic Peak Detection Method for Mass Spectrometry Data", Seventh International Triennial Calcutta Symposium, Kolkata, Dec. 28, 2009, Kolkata, India.

"Improved Automated Monoisotopic Peak Detection Method for Mass Spectrometry Data", International Conference on Frontiers of Interface between Statistics and Sciences, Hyderabad, India, 31 December 2009.

"A Statistical Framework for Differential Network Analysis (DNA) from Microarray Data Using Partial Least Squares", FACSS, October 18-22, 2009, Louisville, KY.

"Monoisotopic Peak Detection and Detecting Protein-Protein Interaction", First Institute of Mathematical Statistics Asia Pacific Rim Meeting, Seoul, June 28-July 1, 2009.

## **2008**

"Reverse Engineering To Construct Protein-Protein Interaction Network", JSM August 1-7, 2008, Denver, Colorado.

"Construction of Genetic Association Networks", International Indian Statistical Association Conference: Frontiers of Probability and Statistical Science, May 22-26, 2008, University of Connecticut-Storrs.

"Determination of Optimal Clustering Algorithm by Weighted Rank Aggregation: Cross Entropy Algorithm", UT-ORNL-KBRIN Bioinformatics Summit 2008, March 28, 2008, Cadiz, KY.

## **2007**

International Conference on Statistics, Probability and Related Areas by IISA, January 2-5, 2007, Cochin, India.

## **2006**

International Conference on Multivariate Statistical Methods, Dec 28-29, 2006, Kolkata, India.

"Combining Functional Information in Validation of Statistical Clustering", International Multi-Symposiums on Computer and Computational Sciences (IMSCCS|06), June 20-24, 2006, Zhejiang University, Hangzhou, China.

"Clustering Microarray Data", UT-ORNL-KBRIN Bioinformatics Summit 2006, April 21-23, 2006, Cadiz, Kentucky.

## **2005**

"Feature Selection in Mass Spectrometry Data for Cancer Classification", SCMA 2005 / FIM XII, International Conference on Statistics, Combinatorics, Mathematics and Applications: 12th Annual Conference of the Forum for Interdisciplinary Mathematics, December 2-4, 2005, Auburn University, AL, USA.

"Selecting an Appropriate Clustering Algorithm for Analyzing Microarray Data", Joint Annual Meeting of the Interface and the Classification Society of North America, June 8, 2005 - June 12, 2005, Washington University School of Medicine, St. Louis, Missouri.

International Conference on Future of Statistical Theory, Practice and Education, December 29, 2004 - January 1, 2005, Hyderabad, India.

## **2004**

"Empirical Bayes Screening (EBS) of Many P-values with Applications to Microarray Studies", Mathematical Biosciences Institute, Ohio State University. Invited presenter in the workshop titled "Genomics, Proteomics, and Bioinformatics", October 13th 2004.

Eleventh International Conference on Interdisciplinary Mathematical and Statistical Techniques, SCRA 2004, December 27-29, 2004, Lucknow, India.

"Parametric and Nonparametric Empirical Bayes Adjustments to Multiple P-values for the Detection of Differentially Expressed Genes in Microarray Experiments", JSM August 7-12, 2004, Toronto, Canada.

"Empirical Bayes Screening of Many P-values with Applications to Microarray Studies", Microarray Data Analysis Conference arranged by Infocast Inc. Conference arranged, June 28-29, 2004, Rockville, MD, USA.

"Empirical Bayes Screening of Many P-values with Applications to Microarray Studies", International Conference on Statistics in Health Sciences, June 23-25, 2004, Nantes, France.

"Empirical Bayes Analyses of Multiple p-values For the Detection of Differentially Expressed Genes in Microarray Experiments", IISA 2004 Meeting, May 7-9, 2004, Athens, Georgia, USA.

The Second Asia Pacific Bioinformatics Conference 18-22 Jan, 2004, Dunedin, New Zealand (full paper).

## **2002**

"Comparisons of Clustering Algorithms for Grouping Genes Based of Expression Profiles", SCRA 2002-FIM IX: Ninth International Conference of Forum for Interdisciplinary Mathematics on Statistics Combinatorics and Related Areas, Department of Statistics and Department of Mathematics: University of Allahabad, Allahabad, UP 211 002, India, December 21-23, 2002.

"Statistical Techniques to Analyze Microarray Data: A Partial Overview", Applying Bioinformatics, from Genes to Systems: The University System of Georgia Research Symposium, Georgia State University, Atlanta, GA, October 3-4, 2002.

"Clustering Algorithms for Microarray Data: Overview and Comparative Studies", International Conference on Current Advances and Trends in Nonparametric Statistics, Crete, Greece, July 2002.

"Clustering Algorithms for Microarray Data: Overview and Comparative Studies", ISA International Conference on Statistics, Probability and Related Areas, Dekalb, Illinois, June 2002.

## **2001**

"Use of Partial Least Squares in Microarray Data", International Conference on Statistics, Combinatorics and Related Areas And The Eighth International Conference of the Forum for Interdisciplinary Mathematics, Wollongong, Australia, December 2001.

"Microarray Data and Bioinformatics: A Statistical Future?", IISA-JSM-2000-2001 Conference, New Delhi, India, January 2001.

## **2000**

"Microarray Data and Bioinformatics: A Statistical Future?", Triennial Calcutta Symposium, Calcutta, India, December 2000.

"Estimation of Selection Parameters Using Multigeneration Cytonuclear Data", Symposium on Inference for Stochastic Processes, Athens, Georgia. Estimation of Selection Parameters University of Georgia, Athens, May 2000.

## **1999**

Sixth International Conference of Forum for Interdisciplinary Mathematics on Statistics, Combinatorics and Related Areas, Mobile, Alabama, December 1999.

## **1998**

IISA International Conference on Statistics: IISA98, Hamilton, Canada, October 1998.

## **1997**

Pathways to the Future workshop, Dallas, Texas, June 1997.

AMS - IMS - SIAM Joint Summer Research Conference on Statistics in Molecular Biology, Seattle, Washington, June 1997.

## **1995**

Second IMS new researchers' meeting, Kingston, Canada, July 1995.

SRCOS/ASA summer conference, Melbourne, FL, Student Paper Award, June 1995.

## **INVITED COLLOQUIA AND SEMINARS**

### **2011**

"Novel Ensembling Technique for Mining Complex High Dimensional Data", Steklov Mathematical Institute of Academy of Sciences, St. Petersburg, Russia, June 17, 2011.

"Novel Ensembling Technique for Mining Complex High Dimensional Data", School of Public Health, University of Tampere, Tampere, Finland, June 13, 2011.

"Novel Ensembling Technique for Mining Complex High Dimensional Data", School of Public Health and information Sciences, University of Louisville, May 26, 2011.

Presentation of a Chair's Vision, School of Public Health and Information Sciences, May 26 2011.

### **2010**

"High Dimensional Data: A New Paradigm of Biomedical Research". Research Incubatio Meeting, SPHIS, University of Louisville, Louisville, Kentucky, 10 March, 2010.

### **2008**

"Bioinformatics in Cancer Research: A Hype or Hope?" JGB Cancer Center, University of Louisville, March 2008.

Department of Quantitative Health Sciences, Cleveland Clinic, January 2008.

## **2007**

Department of Statistics, Bioinformatics Seminar, Texas A&M University, November 2007.

National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, April 2007.

Department of Epidemiology, Michigan State University, East Lansing, March 2007.

## **2006**

Department of Statistics, University of Kentucky, Lexington, October 2006.

## **2005**

ASA, Kentucky Chapter, December, 2005.

"Statistical Techniques to Analyze Microarray Data: A Partial Overview", BiComB Seminar, SPHIS, Emory University, Atlanta, GA.

"Analysis of MALDI-TOF Data", Proteomics Center, University of Louisville, Louisville, November 2005.

## **2004**

Department of Statistics and Applied Probability, National University of Singapore, December 2004.

Department of Biostatistics and Bioinformatics, University of Louisville, December 2004.

Department of Biostatistics, Univ. of Minnesota, April 2004.

## **2003**

Department of Mathematics, Univ. of North Carolina at Charlotte, April 2003.

Department of Mathematics, Georgia Institute of Technology, March 2003.

Department of Mathematics and Statistics, Georgia State University, March 2003.

## **2001**

A one day professional training course on Poisson regression, Centers for Disease Control and Prevention, Atlanta, GA, February 2001.

Department of Biostatistics, University of Alabama, Birmingham, AL, August 2001.

Department of Statistics, University of Georgia, Athens, GA, April 2001.

## **2000**

Bioinformatics group, Jonathan Arnold Lab, Department of Genetics, University of Georgia, Athens, GA, December 2000.

A two day professional training course on Smoothing using S+, Centers for Disease Control and Prevention, Atlanta, GA, November 2000.

#### **1999**

Computer Science Unit and Mathematical Science Unit, Indian Statistical Institute, Calcutta, India, June 1999.

Department of Mathematics and Statistics, University of North Carolina, Charlotte, April 1999.

Department of Statistics, Purdue University, April 1999.

#### **1997**

Department of Mathematics and Computer Science, Georgia State University, March 1997.

#### **1996**

Department of Genetics, University of Georgia, December, 1996.

#### **1995**

Population Biology Group, Emory University, November, 1995.

Department of Biostatistics, Emory University, October, 1995.

Department of Mathematics & Statistics, University of North Florida, April 1995.

### **REFEREED TALKS AT PROFESSIONAL MEETINGS**

#### **2011**

"Predicting Patient Survival from Proteomic Profile using MALDI-TOF Mass Spectrometry Data in Non-small Cell Lung Cancer Patients", RECOMB Satellite Conference on Computational Proteomics, March 11-13, 2011, University of California, San Diego, La Jolla, California.

#### **2009**

"Next Generation Sequencing: Statistical Challenges and Opportunities", CAMDA 2009, October 5-6, 2009, IL, Chicago.

"Improved Monoisotopic Peak Detection in Mass Spectrometry Data", MCP 2009, Tokyo, Japan, March 27, 2009.

#### **2008**

"A Statistical Framework for Differential Network Analysis (DNA) from Microarray Data using Partial Least Squares", Rocky '08, 6th Annual Rocky Mountain Bioinformatics Conference, Snowmass, CO, December 5-8, 2008.

#### **2007**

"Fetal Alcohol Syndrome Detected through Proteomics Screening of the Amniotic Fluid in High-Risk (B6J) and Low-Risk (B6N) C57BL/6 Mice", Teratology Society 47th Annual Meeting, Pittsburg, PA, June 2007.

## **CONTRIBUTED TALKS AT PROFESSIONAL MEETINGS**

### **2004**

ENAR Spring Meeting, Pittsburgh, March 2004.

### **2003**

Digital Biology: The Emerging Paradigm, NIH/BISTI, Bethesda, Maryland, November 2003.

### **1999**

Joint Statistical Meetings, Baltimore, Maryland, August 1999.

ENAR Spring Meeting, Atlanta, Georgia, March 1999.

### **1998**

19th International Biometric Conference: IBC-98, Cape Town, South Africa, December 1998.

### **1996**

18th International Biometric Conference: IBC-96, Amsterdam, The Netherlands, June 1996.

IMS/ENAR meeting, Richmond, Virginia, March 1996.

### **1995**

ASA Winter Conference, Raleigh, North Carolina, January 1995.

## **REFEREED POSTER PRESENTATION IN PROFESSIONAL AND RESEARCH MEETINGS**

### **2008**

"Stochastic Modeling and Statistical Inference of Time Course Metabolic Data". Integrative BioSystems Institute Conference, October 18 -21, 2008, Atlanta, Georgia.

### **2007**

"Finding Cancer Genes Through Meta-analysis of Microarray Experiments: Rank Aggregation via the Cross Entropy Algorithm". CAMDA 2007, December 2007, Valencia, Spain.

"cIValid: An R package for Cluster Validation". July 2007, ISMB 2007, Vienna, Austria.

## **2005**

"Selecting a Clustering Algorithm for Statistical Consistency and Biological Relevance for Gene Expression Data". June 27, 2005, ISMB 2005, Detroit, Michigan.

## **MEETINGS ATTENDED**

### **2011**

Steering committee meeting for the "Celebration of Women in Statistics" in 2013, ASA, Alexandria, VA, Nov.4, 2011.

7th Annual Conference of US HUPO. Proteomics: New Developments and Grand Challenges, March 20 - 23, 2011 - Raleigh, NC

### **2007**

Joint Statistical Meetings, Washington DC, August 2007.

### **2006**

Joint Statistical Meetings, Seattle, Washington, August 2006.

International Biometric Society Conference IBC 2006, Montreal, Canada, July, 2006.

### **2003**

Joint Statistical Meetings, San Francisco, California, August 2003.

Joint Statistical Meetings, Atlanta, Georgia, August 2003.

### **2001**

Workshops for Junior Biostatisticians, 2001 ENAR, Charlotte, March 2001.

### **2000**

Genomics and Medicine Symposium, Emory University, Atlanta, September 2000.

Bioinformatics Conference, Georgia Research Alliance, Atlanta, June 2000 (invited).

Beyond Genome 2000, San Francisco, June 2000.

### **1998**

Joint Statistical Meeting, Dallas, Texas, August 1998.

## **TEACHING**

### **@ University of Louisville**

#### **Graduate Level Only**

1. Mathematics Statistics
2. Biostatistical Consulting

3. (Developed and taught) Statistics for Bioinformatics
4. (Developed and taught) High-throughput Data Analysis
5. (Developed and taught) Statistics for Proteomics
6. (Developed and taught) Multivariate Statistics for Genomics and Proteomics
7. (Developed and taught) Stochastic Modeling and Statistical Inference for Time-Course Data
8. (Developed and taught) Data mining for Genomics and Proteomics

**@ Georgia State University**

**Graduate**

1. Introductory Biostatistics
2. (Developed and Taught) Statistics for Bioinformatics
3. (Developed and taught) Survival Analysis
4. (Developed and taught) Statistical Computing with S+

**Undergraduate**

1. Introductory Biostatistics
2. (Developed and taught) Developed an interdisciplinary course on “Scientific Perspectives on World Hunger” Two other faculty from Nutrition and Geology – This course is used as one of the model courses in Models of Accessible Course Design under the GRADE project conducted by Georgia Institute of Technology (GT), IDET Communication Inc. (IDET), and the Southeast Disability and Business Technical Assistance Center (Southeast DBTAC).

**@ University of Georgia**

**Undergraduate**

2000 Summer long course for “Analysis of microarray data with R” for NSF/REU program, Summer, 2000.

**TEACHING OUTSIDE OF CURRENT INSTITUTION**

**@ Center for Disease Control and Prevention (CDC), Atlanta, GA**

**Professional Courses**

- 2001 (Developed and taught) Professional training course on Poisson regression  
 2000 (Developed and taught) A two day professional training course on Smoothing using S+

**STUDENT DIRECTION**

**Major or Co-Major Professor**

**PhD (@ University of Louisville)**

- 2009 Mourad Atlas, completed, Summer 2009 (Joined FDA).  
 2009 Vasyi Pihur (jointly with Somnath Datta), completed Summer 2009. (Joined John's Hopkins as a post-doc and Google Inc. thereafter).  
 2010 Farida Mostazabi (jointly with Somnath Datta), Graduated, June 2011. Currently at Medpace Inc., Cincinnati, OH 45227.

- 2012 Juliet Ndukum, Expected Completion, May 2012.
- 2013 Sutirtha Chakraborty (jointly with Somnath Datta), Expected Completion, May 2013.

**MS (@ University of Louisville)**

- 2007 Christopher N. Barnes, Feature Selection and Classification in High-throughput Data Analysis, graduated summer 2007 (in industry after PhD).
- 2008 Becky Patterson (jointly with P. Yoo), Testing the Effects of Predictors Using Data Generated by Non-identity Link Functions of the Single-Index Model: A Monte Carlo Approach, graduated spring 2008 (Employed in Graduate School UofL).
- 2010 Xiaohong Li, Completed, Spring, 2010 (Enrolled in the PhD program in Biostatistics at UofL).
- 2011 Jasmit Shah, completed, Summer , 2011, (Enrolled in the PhD program in Biostatistics at UofL).

**MS (@ Georgia State University)**

- 2001 Sylvie Bougi, A Comparative Power Study of Statistical Tests of Neutrality of DNA Markers Using Multigeneration Cytonuclear Data; graduated Spring 2001.
- 2002 Israel Hora, Time Series Analysis of Georgia Employment Data and Future Prediction of Employment Status, graduated, Fall, 2002.
- 2002 Ying, Yang (Biology), non-thesis report, graduated Summer, 2002.
- 2004 Jennifer Le-Rademacher, Partial Least Squares in Censored Survival Regression, graduated Summer 2004.
- 2004 Xiao Hong Zhu, graduated Summer 2004.
- 2004 Usha Ramakrishnan, graduated Fall 2004.
- 2005 Mourad Atlas, graduated Spring 2005.
- 2005 Lace Depadilla, graduated Spring 2005.
- 2005 Baofu Ma, graduated Summer 2005.
- 2005 Saurav Karmaker, graduated Summer 2005.

**Committee Member at University of Louisville**

- 2011 Stephanie Tabb, PHD, Graduated Fall 2011
- 2011 Hyejeong Jang, MS, Graduated May 2011
- 2008 Ling Lang PHD, Graduated August 2008
- 2007 Vasyl Pihur MS. Graduated August 2007