Curriculum Vitae

Paola Sebastiani

General Information

Birth:	January 16th, 1964. Brescia, Italy.
Citizenship:	Italian, USA permanent resident.
Current Position	Associate Professor.
Affiliation	Department of Biostatistics, Boston University, Boston MA

Research Interests

Bayesian Modeling: Bayesian model-based clustering. Bayesian models in genetics and genomics. Graphical models and Bayesian networks. Modeling incomplete data.

Design of Experiments: Optimal design of experiments. Experimental design in a decision theoretic framework. Information based design criteria.

Decision Theory: Decision theoretic foundations in statistical modeling and experimental design.

Machine Learning: Automation of statistical methods.

Applications: Artificial intelligence, bioinformatics, genetics, genomics, medical informatics, robotics, disease surveillance.

Summary

Paola Sebastiani is Associate Professor in the Department of Biostatistics at Boston University School of Public Health and Adjunct Associate Professor in Bioinformatics, Boston University. She is also founder and Chief Scientific Officer of Bayesware Ltd, a startup company developing and commercializing knowledge discovery software. She received a degree in Mathematics from the University of Perugia, Italy, 1987, a Master degree in Applied Stochastic Systems from the University College London, 1990, and a PhD in Statistics from the University of Rome, Italy, 1992. From 1990 to 1995 she was researcher at the Department of Statistics of the University of Perugia, Italy. Between 1995 and 2000 she held faculty positions at City University, London, the Open University, and Imperial College London, where she was Governor lecturer in the Mathematics Department. Between 2000 and 2003 she was Assistant Professor in the Department of Mathematics and Statistics, University of Massachusetts, Amherst. Her research interests include development and applications of Bayesian statistical methods in biomedical informatics, genomics and genetics. She has developed novel methodologies in machine learning and artificial intelligence; decision theory; graphical modeling and statistical experimental design. Her recent work focuses on the discovery of genetic modifiers of sickle cell anemia and exceptional longevity.

Contents

1. Education

DOCTORAL DEGREE 1989-1992	Department of Applied Probability and Statistics, University of Rome La Sapienza, Italy.
MASTER OF SCIENCE 1989-1990	Department of Statistical Science, University College London, London. *Distinction.*
DEGREE 1983-1987	Department of Mathematics, University of Perugia, Italy. Degree in Mathematics. *Summa cum Laude.*
MATURITA' SCIENTIFICA 1978-1983	Liceo Scientifico A. Volta, Italy, 60/60.

2. Professional Experience

ASSOCIATE PROFESSOR 2003-	Department of Biostatistics, Boston University, Boston MA
ASSOCIATE PROFESSOR 2003-	Bioinformatics Program, Boston University, Boston MA
ASSISTANT PROFESSOR 2000-2003	Department of Mathematics and Statistics, University of Massachusetts at Amherst, MA
GOVERNOR'S LECTURER 2000	Mathematics Department, Imperial College of Science, Technology and Medicine, London, UK.
CHIEF SCIENTIFIC OFFICER 1999 to date	Bayesware LLC, Boston, MA (www.bayesware.com).
LECTURER 1998 to 2000	Statistics Department, The Open University, Milton Keynes.
SENIOR RESEARCH FELLOW June–September 1998	Department of Computer Science, University of Massachusetts at Amherst, MA.
LECTURER 1995 to 1998	Department of Actuarial Science and Statistics, City University, London. Probation passed in May 1998.
VISITING LECTURER November–December 1994	Department of Statistics, London School of Economics and Political Science.
VISITING RESEARCHER July–October 1994	Department of Statistical Science, University College London.
VISITING RESEARCHER September–October 1992	Department of Statistical Science, University College London.
RESEARCHER 1990-1995	Faculty of Economics, University of Perugia, Italy. Tenured received in September 1993.
DOCTORAL STUDENT 1988-1990	Department of Applied Probability and Statistics, University of Rome La Sapienza, Italy.

3. Grants Awarded

NIH: R01 HG003354-01A2 2006-2009	Decoding Gene Expression Control Using Conditional Clustering by Dynamics. Subcontract to Childrens' Hospital [450,000 USD]. Funding agency: National Institutes for Health. USA

NIH: R21 HL080463 2005-2007	Genetic Dissection of Sickle Cell Anemia Phenotypes. Principal Investigator [360,000 USD]. Funding agency: National Institutes for Health. USA
A. P. SLOANE FOUNDATION 2003-2004	"Bioterrorism Surveillance: An Alarming Problem". Subcontractor to Children's Hospital [60,000 USD]. Funding agency: Alfred P. Sloane Foundation
NSF 2001-2004	"Systematic Phenotyping of SNPs using Bayesian Networks". Subcontract to Children's Hospital. [180,000 USD]. Funding agency: National Science Foundation. USA
NSF 2001-2004	"Musical Signal Recognition". Co-principal investigator [375,000 USD]. Funding agency: National Science Foundation. USA
DARPA 2000-2004	"Planning in the real world". Subcontract to EKSL, Department of Computer Science, UMass, Amherst. [15,000 USD]. Funding agency: Defense Department. USA
EUROPEAN COMMUNITY 1999-2000	"Bayesian Knowledge Extractor". Principal Investigator. [200,210 ECU]. Funding Agency: Eurostat
ROYAL SOCIETY 1998	Travel grant. [800 GBP]. Funding agency: Royal Society. UK
ROYAL SOCIETY 1996	Travel grant. [600 GBP]. Funding agency: Royal Society. UK
NUFFIELD FOUNDATION 1996	"Maximum-Entropy-Sampling". Principal Investigator. [3,000 GBP]. Funding agency: Nuffield Foundation. UK
MURST 1994	"Optimal experimental designs for non linear models". Principal Investigator. [6,000,000 Lit]. Funding agency: Italian Ministry of University
EUROPEAN SCIENCE FOUNDATION 1994	Visiting grant. Principal Investigator. [10,000 FF]. Funding agency: European Science Foundation.

4. Awards and Fellowships

TEACHING AWARD 2009	Award for teaching the course "Bayesian modeling in biomedical research and public health" at the Boston University School of Public Health
TEACHING AWARD 2005	Award for teaching the course "Statistical Methods in Functional Genomics" that was ranked best course taught in the fall 2004 at the Boston University School of Public Health
HONORARY MENTION 1999	Honorary mention at the Knowledge Discovery Cup, San Diego, CA, 1999 for the work "Modeling customers behavior using Bayesian Networks".
FELLOWSHIP	Research fellowship. Consiglio Nazionale delle Ricerche.

5. Software Developed

DISEASE SEVERITY CALCULATOR. 2006 to date	A system than computes the risk for early death in subjects with sickle cell anemia using a network model that integrates several laboratory and clinical variables.
BLINK. 2005 to date	Collection of Bayesian tools for the genomewide analysis of linkage disequilibrium.
BITMAPBUILDER. 2005 to date	Software to display linkage disequilibrium maps over wide ranges of the genome.
rTDT. 2003	A program that generalize the transmission disequilibrium test when there is missing genotype information.
BADGE. 2002–2005	(Bayesian differential analysis of gene expression data) is a program that implements a novel Bayesian method for the differential analysis of gene expression data measured with microarrays.
BEST. 2002	(Best enumeration of haplotype taggings) is a program for the determination of haplotype tagging SNPs from a set of specified haplotypes.
CAGED. 2001	(Cluster analysis of gene expression dynamics) is a program for the analysis of temporal expression profiles of gene expression data. Reviews of the program appeared in Science, July 2002, BioIT World, September 2002, The ISBA Bulletin (2003), and The Scientist (2003). The program has more than 1000 users including investigators from hospitals, academic institutions and research organization all over the world (http://genomethods.org/caged/users.htm)
RoC. 1999-2000	(A computer program for robust classification using incomplete information.
BAYESWARE DISCOVERER 1997-2000	Originally BKD (Bayesian knowledge extractor) is the first computer program for the automatic induction of Bayesian Networks from databases. Since 1997 the program has been distributed to thousands of investigators and used to model complex systems in several domains including genetics, genomics, artificial intelligence, biomedical informatics, and census data.

6. Editorials on articles I wrote.

BU TODAY July 31, 2008	Looking to Centenarians for Sickle-Cell Answers.
NATURE GENETICS Vol. 37, April 2005.	Defining stroke risks in sickle cell anemia
NATURE BIOTECHNOLOGY Vol. 37, April 2005.	Improving the ability to predict stroke
LANCET NEUROLOGY May 2005.	Genetic Model Predicts Stroke in Sickle-Cell Disease

SCIENCE Vol. 307, July 2002.	Crunching Gene Chips
BioIT WORLD September 2002	Group Thinks Outside the Box
THE SCIENTIST Vol 16, October 2002	Gene Expression Data Mining

7. Professional Services

EDITOR 2002-2003	"Machine Learning Journal". Special issue: Methods in Functional Genomics.
EDITORIAL BOARD 2005 to date	"Eurasip Journal of Signal Processing and Bioinformatics".
EDITORIAL BOARD 2004 to date	"Bayesian Analysis Journal".
EDITORIAL BOARD 2001 to 2005	"Machine Learning Journal".
EDITORIAL BOARD 1998 to 2003	"Evaluation of Intelligent Systems" website.
COMMITTEE MEMBERSHIP 2000–2001	Research Section of the Royal Statistical Society; Statistical Computing Section of the Royal Statistical Society
PROGRAM CHAIR CONFERENCE	Genomic Signal Processing 2008 (Cochair).
PROGRAM COMMITTEE CONFERENCE	Genomic Signal Processing 2005/2007; International Conference on Machine Learning (2003 to date); Artificial Intelligence and Statistics, 2005; MedInfo (Medical Informatics) 2004; American Association for Artificial Intelligence 2004; ISIPTA (2003 to date); Virtual Conference in Genomic Research (2003); Probabilistic and Graphical Models for Classification (2004 European conference on Machine Learning Workshop); Machine Learning in Bioinformatics (2003 International conference on machine learning workshop); Intelligent data analysis in medicine and pharmacogenomics (2001 to date); Intelligent Data Analysis symposium (2000 to date); "Bayesian Models in Medicine" (2001 to date); Royal Statistical Society meeting on "Statistical Methods in Genetics" (2001).
COUNCIL 1999 to 2005	Member of the Intelligent Data Analysis council.
REVIEWER FOR FUNDING AGENCY	British Economic and Social Research Council; NWO Netherlands Organization for Scientific Research; Pennsylvania Department of Health (DOH); National Institutes for Health (USA).
CONFERENCES ORGANIZER:	"Dynamic Models", the 10th Annual Statistics Conference of the Open University, December 1999. "Bayesian Knowledge Discovery in Official Statistics", session at the ISBA2000 conference.

REFEREE

Journals: Am J. Hum Genet, Art. Intel. in Med., Biometrika, Biometrics, Circulation, Genetics, Genome Res., IEEE Biomed. Eng., Intel. Data Anal. J, IEEE System, Man and Cyb, J. Americ Statist Assoc; IIE, Int. J. Approx Reasoni., J. Roy. Statist. Soc. B, J. Statist. Plann. Infer., J. Italian Statist. Soc, J. London Mathem. Soc., J. Mach. Learn. Resear. Mach. Learn., Physiol. Genomics, Rel. Eng. and Sys. Saf., Soft Computing, Statistica Sinica, Statist. in Med. *Conference:* Uncertainty in artificial intelligence, Knowledge discovery and data mining, Intelligent data analysis, American association for artificial intelligence

SOCIETY MEMBERSHIP

American Statistical Association (Member since 1998); Institute of Mathematical Statistics (Member since 1995); International Biometric Society (Member since 2002); International Association for Statistics and Computing (Fellow since 1999); International Society for Bayesian Analysis (Fellow since 1997); Royal Statistical Society (Fellow since 1992).

8. Publications

1 Journal Articles

1. Z. Zhao, N. Timofeev, S. W Hartley, D. K Chui, S. Fucharoen, M. H Steinberg, T. T. Perls, C. T Baldwin and <u>P. Sebastiani</u> (2008), Imputation of Missing Genotypes: An Empirical Evaluation of IMPUTE. *BMC Genetics.* Dec 12;**9**:85

2. V. G. Nolan, Y. Zhang, T. Lash, <u>P. Sebastiani</u> and M. H. Steinberg (2008), Association of Wind Speed and the Occurrence of Sickle Cell Acute Painful Episodes: Results of a Case-Crossover Study. *British Journal of Hematology.* **143**(3), 433–8.

3. A. E. Sedgewick, N. Timofeev, <u>P. Sebastiani</u>, J. C.C. So, E. S.K. Ma, L. Chong Chan, G. Fucharoen, S. Fucharoen, C. G. Barbosa, L. A. Farrer, C. T. Baldwin, M. H. Steinberg and D. H.K. Chui (2008), BCL11A is a major hbf quantitative trait locus in three different populations with β-hemoglobinopathies. *Blood Cells, Molecules, and Diseases.* **41**(3), 255–8.

4. J. Beane, <u>P. Sebastiani</u>, T. H. Whitfield, K. Steiling, Y. M. Dumas, M. E. Lenburg and A. Spira (2008), A Prediction Model for Lung Cancer Diagnosis that Integrates Genomic and Clinical Features. *Cancer Prevention Research* **1**, 56-64.

5. M. M. Abad-Grau, J. Ierache, C. Cervino and <u>P. Sebastiani</u> (2008), Evolution and challenges in the design of computational systems for triage assistance in the Emergency Department. *Journal of Biomedical Informatics* **41**(3):432-41.

6. L. Wang, M. Montano, M. Rarick and <u>P. Sebastiani</u> (2008), Conditional Clustering of Temporal Expression Profiles. *BMC Bioinformatics* **9**(1): 147.

7. <u>P. Sebastiani</u>, Z. Zhao Z, M. M. Abad-Grau, A. Riva A, S. W. Hartley, A. E. Sedgewick, A. Doria, M. Montano, E. Melista, D. Terry, T. T. Perls, M. H. Steinberg, C. T. Baldwin. (2008), A hierarchical and modular approach to the discovery of robust associations in genome-wide association studies from pooled DNA samples. *BMC Genetics* Jan 14;**9**(1):6

8. <u>P. Sebastiani</u>, L. Wang, V. G. Nolan, E. Melista, Q. Ma, C. T. Baldwin and M. H. Steinberg (2008), Fetal hemoglobin in sickle cell anemia: Bayesian modeling of genetic associations. *American Journal of Hematology.* **83**(3):189–95.

9. D. Terry, <u>P. Sebastiani</u>, S. Anderson and T. T. Perls (2008), Disentangling the roles of disability and morbidity in survival to exceptional old age. *Archives of Internal Medicine.* **168**(3):277–83.

10. J. Bean, <u>P. Sebastiani</u>, G. Liu, J. Brody, M. Lenburg, A. Spira (2007), Reversible and permanent effects of tobacco smoke exposure on airway epithelial gene expression. *Genome Biology* **8**(9):R201

11. <u>P. Sebastiani</u>, V. G. Nolan, C.T. Baldwin, M. M. Abad-Grau, L. Wang, A. H. Adewoye, L. C. McMahon, M.D., L. A. Farrer, J. G. Taylor, IV., G. J. Kato, M. T. Gladwin, M. H. Steinberg (2007), A network model to predict the risk of death in sickle cell disease. *Blood.* **110**(7):2727–2735.

12. <u>P. Sebastiani</u> and M. Abad-Grau (2007), Bayesian estimates of linkage disequilibrium. *BMC Genetics.* June 25 8:36.

13. M. Montano, J. Flanagan, L. Jiang, P Sebastiani, M. Rarick, N. LeBrasseur, C Morris, R. Jasuja and S. Bhasin (2007), Transcriptional profiling of testosterone-regulated genes in the skeletal muscle of hiv-infected men experiencing weight loss. *The Journal of Clinical and Endocrinology Metabolism* 92(7):2793–2802.

14. A. Spira, J. E. Beane, V. Shah, K. Steiling, G. Liu, F. Schembri, S. Gilman, Y. Dumas, P. Calner, P Sebastiani, S. Sridhar, J. Beamis, C. Lamb, T. Anderson, N. Gerry, J. Keane, M. E. Lenburg and J. S. Brody (2007), Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nature Medicine* **13** 361–366.

15. F. Ferrazzi, P. Sebastiani, M. F. Ramoni, and R. Bellazzi (2007), Bayesian approaches to reverse engineer cellular systems: a simulation study on nonlinear Gaussian networks. *BMC Bioinformatics.* **8**(Suppl 5):S2 doi:10.1186/1471-2105-8-S5-S2

16. M. Abad, R. Montes and P. Sebastiani (2006), Building chromosome-wide LD maps. *Bioinformatics.* **22**, 1933–1944

17. P. Sebastiani, H. Xe and M. F. Ramoni (2006), Bayesian analysis of comparative microarray experiments by model averaging. *Bayesian Analysis Journal.* **1**, 707–732.

18. M. Montano, M. Rarick, P. Sebastiani, P. Brinkmann, M. Russell, A. Navis, M. Essex, C. Wester and I. Thior (2006), Gene-expression profiling of hiv-1 infection and perinatal transmission in Southern Africa. *Genes and Immunity.* **7**, 298–309.

19. V. Nolan, A. Adewoye, C. Baldwin, L. Wang, Q. Ma, D. Wyszynski, J. Farrell, P. Sebastiani, L. Farrer and M. H. Steinberg (2006), Sickle cell leg ulcers: associations with haemolysis and SNPs in Klotho, TEK and genes of the TGF-β/BMP pathway. *British Journal of Hematology.* **133**, 570–578.

20. M. Rarick, C. McPheeters C, S. Bright S, A. Navis A, J. Skefos J, P. Sebastiani, M. Montano (2006), Evidence for cross-regulated cytokine response in human peripheral blood mononuclear cells exposed to whole gonococcal bacteria in vitro. *Microb Pathog.* **40**, 261–270.

21. P. Sebastiani, K. D. Mandl, P. Szolovits, I. S. Kohane and M. F. Ramoni (2006), Rejoinder: A Bayesian dynamic model for influenza surveillance. *Statistics in Medicine. Statistics in Medicine.* **25**, 1823–1825.

22. P. Sebastiani, K. D. Mandl, P. Szolovits, I. S. Kohane and M. F. Ramoni (2006), A Bayesian dynamic model for influenza surveillance. *Statistics in Medicine.* **25**, 1803–1816.

23. M. Montano, M. Rarick, P. Sebastiani, P. Brinkmann, J. Skefos and R. Ericksen (2006), HIV-1 burden influences host response to co-infection with Neisseria gonorrhoeae in vitro. *Int Immunol.*, **18**, 125–137.

24. P. Sebastiani, M. F. Ramoni, V. Nolan, C. T Baldwin, M. H. Steinberg (2005), From the Cover: Genetic dissection and prognostic modeling of overt stroke in sickle cell anemia. *Nature Genetics*, **37**, 435–440.

25. C. T. Baldwin, V. G. Nolan, D. F. Wyszynski, Q. Ma, P. Sebastiani, S. H. Embury, A. Bisbee, J. Farrell, L. Farrer and M. H. Steinberg (2005), Association of Klotho, Bone Morphogenic Protein 6 and Annexin A2 polymorphisms with sickle cell osteonecrosis. *Blood*, **106**, 372–375.

26. E. S. Klings, S. Safaya, A. H. Adewoye, A. Odhiambo, G. Frampton, M. Lenburg, N. Gerry, P. Sebastiani, M. H. Steinberg and H.W. Farber (2005), Differential gene expression in pulmonary artery endothelial cells exposed to sickle cell plasma. *Physiological Genomics.* **21**, 293–298.

27. L. Wang, M. F. Ramoni, K. D. Mandl and P. Sebastiani (2005), Factors affecting the performance of syndromic surveillance systems. *Artificial Intelligence in Medicine.* **34**, 269–278.

28. P. Sebastiani, and M. F. Ramoni (2005), Normative selection of Bayesian networks. *Journal of Multivariate Analysis.* **93**(2) 340–357.

29. P. Sebastiani, M. Abad, G. Alpargu and M. F. Ramoni (2004), Robust Transmissio/Disequilibrium Test for incomplete family genotypes. *Genetics.* **168**, 2329–2337.

30. K. Mandl, J. M. Overhage, M. M. Wagner, W. L. Lober, P. Sebastiani, F. Mostashari, J. A. Pavlin, P. H. Gesteland, T. Treadwell, E. Koski, L. Hutwagner, D. L Buckeridge, R. D. Aller, S. Grannis, (2004), Implementing syndromic surveillance: a practical guide informed by the early experience. *Journal of the American Medical Informatics Association.* **11**:141-150.

31. P. Sebastiani, Y. H. Yu, and M. F. Ramoni (2003), Bayesian machine learning and its potential applications to the genomic study of oral oncology. *Journal of Dental Research.* **17**:104-108.

32. P. Sebastiani, R. Lazarus, S. T. Weiss, L. M. Kunkel, I. S. Kohane and M. F. Ramoni (2003), Minimal haplotype tagging. *Proceedings of the National Academy of Sciences.* **100**(17):9900–9905,

33. P. Sebastiani, E. Gussoni, I. S. Kohane and M. F. Ramoni, (2003), Rejoinder. *Statistical Science.* **18**:69–70.

34. P. Sebastiani, E. Gussoni, I. S. Kohane and M. F. Ramoni (2003), Statistical challenges in functional genomics. *Statistical Science.* 33–60.

35. P. Sebastiani, M. F. Ramoni and I. Kohane (2003), Machine learning in the genomics era. *Machine Learning Journal.* **52**, 5–9.

36. M. F. Ramoni, P. Sebastiani and I. S. Kohane (2002), From the Cover: Cluster analysis of gene expression dynamics. *Proceedings of the National Academy of Sciences.* **99**(14):9121-9126.

37. M. F. Ramoni, P. Sebastiani and P. Cohen (2002), Bayesian clustering by dynamics. *Machine Learning.* **47**(1), 91–121.

38. P. Sebastiani and M. F. Ramoni (2001), Common trends in European school population. *Research in Official Statistics.* **4** (1), 169–183.

39. P. Sebastiani and M. F. Ramoni (2001), On the use of Bayesian networks to analyze survey data. *Research in Official Statistics.* **4** (1), 53–64.

40. P. Sebastiani and M. F. Ramoni (2001), Bayesian selection of decomposable models with incomplete data. *Journal of the American Statistical Association.* **96** (456), 1375–1386.

41. M. F. Ramoni and P. Sebastiani (2001), Robust learning with missing data. *Machine Learning.* **45**(2), 147-170

42. M. F. Ramoni, <u>P. Sebastiani</u> and R. Dybowsky (2001), Robust prediction for intensive-care patients. *Methods of Information in Medicine.* **40**, (1), 39–45.

43. M. F. Ramoni and <u>P. Sebastiani</u> (2001), Robust Bayes classifier. *Artificial Intelligence.* **125**, 207–224.

44. <u>P. Sebastiani</u>, M. F. Ramoni and A. Crea (2000), Profiling customers from in-house data. *ACM SIGKDD Explorations*, **1**, 91–96.

45. <u>P. Sebastiani</u> and M. F. Ramoni (2000), Bayesian inference with missing data using bound and collapse. *Journal of Computational and Graphical Statistics.* **9**, 779–800.

46. <u>P. Sebastiani</u> and H.P. Wynn (2000), Maximum entropy sampling and optimal Bayesian experimental design. *Journal of the Royal Statistical Society. Series B.* **62**, 145–157.

47. A. P. Dawid and <u>P. Sebastiani</u> (1999), Coherent dispersion criteria for optimal experimental design. *Annals of Statistics*, **27**, 65–81.

48. M. F. Ramoni and <u>P. Sebastiani</u> (1998), Parameter estimation in Bayesian networks from incomplete databases. *Intelligent Data Analysis. An International Journal*, **2**, 139–160.

49. <u>P. Sebastiani</u> and R. Settimi (1998), First-order optimal design for nonlinear problems. *Journal Statistical Planning and Inference*, **74**, 177–192.

50. <u>P. Sebastiani</u> and R. Settimi (1997), A note on D-optimal designs for a logistic regression model. *Journal Statistical Planning and Inference*, **59**, 359–368.

51. S. Colazza, M.G. Rosi, <u>P. Sebastiani</u> and M. Ursini (1996), Host discrimination in the egg Parasitoids Trissolcus Basalis. *Acta Aecologica*, **17**, 109–125.

52. <u>P. Sebastiani</u> (1996), On the derivatives of matrix powers. *SIAM Journal of Matrix Analysis* **17**, 640–648.

53. M.L. Eaton, A. Giovagnoli and <u>P. Sebastiani</u> (1996), A predictive approach to the Bayesian design problem with application to normal regression models. *Biometrika*, **83**, 111–125.

54. J. Burridge and <u>P. Sebastiani</u> (1994), D-optimal Designs for generalized linear models with variance proportional to the square of the mean. *Biometrika*, **81**, 295–304.

55. J. Burridge and <u>P. Sebastiani</u> (1992), Optimal designs for generalized linear models. *Journal of the Italian Statistical Society*, **2**, 183–202.

56. M. I. Taticchi, P. Pandolfi, M. T. Biondi and <u>P. Sebastiani</u> (1992), Population dynamics of Daphnia Galeata Sars and Bosmina Longirostris Muell in Trasimeno lake. *Bollettino di Zoologia*, **59**, 427–435.

57. A. Giovagnoli and <u>P. Sebastiani</u> (1989), Experimental designs for mean and variance estimation in variance component models. *Computational Statistics and Data Analysis*, **8**, 21–28, Correction ibidem 1990, **10**, 93.

58. A. Giovagnoli and <u>P. Sebastiani</u> (1988), Introduzione ai modelli lineari con più componenti di varianza. *Quaderni di Statistica e Matematica Applicata* **4** 79–112, Università di Perugia.

2 Refereed Conference Proceedings

59. M. M. Abad-Grau and P. Sebastiani (2007), Multivariate imputation of genotype data using short and long range disequilibrium. *Lecture Notes in Computer Science: Proceedings of Computation and Simulation in Modelling Biological Systems.* 187–194.

60. F. Ferrazzi, P. Sebastiani, I. S. Kohane, M. F. Ramoni, R. Bellazzi (2006), Dynamic Bayesian networks in modelling cellular systems: a critical appraisal on simulated data. *19th IEEE Symposium on Computer-Based Medical Systems (CBMS'06)*, 544–549.

61. L. Wang, M. F. Ramoni and P. Sebastiani (2006), Clustering short gene expression profiles. *Lecture Notes in Computer Science: Research in Computational Molecular Biology: 10th Annual International Conference, (RECOMB 2006)*, 60–68.

62. M. M. Abad and P. Sebastiani (2006), Bayesian correction for SNP ascertainment bias. *Lecture Notes in Computer Science: Modeling Decisions for Artificial Intelligence: 3rd International Conference, (MDAI 2006)*, 262–273.

63. F. Ferrazzi, R. Amici, R., P. Sebastiani, I. S. Kohane, M. F. Ramoni and R. Bellazzi (2006), Can we use linear Gaussian networks to model dynamic interactions among genes? Results from a simulation study. In *Genomic Signal Processing and Statistics, (GENSIPS 2006)*, 13–14

64. F. Ferrazzi, P. Sebastiani, R. Bellazzi, I.S. Kohane and M. F. Ramoni (2006), Identification of feedback structures from gene expression time series. *Proceedings of Intelligent Data Analysis in Medicine and Pharmacology (IDAMAP 2006)*, 57–60.

65. P. Sebastiani, L. Wang, K. D. Mandl and M. F. Ramoni (2004), Improving outbreak detection by signal integration. *MMWR: Mortality and Morbidity Weekly Reports (Supplement).* **257**, 57.

66. S. Lodi, P. Sebastiani, D. Cocchi and M. F. Ramoni (2004), Bayesian clustering of gene expression dynamics: an application. *New Developments in Classification and Data Analysis.* Springer-Verlag. 259–266.

67. S. Lodi, D. Cocchi and P. Sebastiani (2003), Applications of Bayesian clustering by dynamics. *Book of short papers. CLADAG 2003: Meeting of the Classification and Data Analysis Group of the Italian Statistical Society* CLUEB, Bologna, 249–252.

68. T. Rath, M. Carreras and P. Sebastiani (2003), Automated detection of influenza epidemics with Hidden Markov Models. *Lecture Notes in Computer Science: Proceedings of the Fifth International Symposium on Intelligent Data Analysis.* 521–531.

69. C.T. Morrison, P. Cohen, and P. Sebastiani (2002), On the development of visual object memory: the stay/go decision problem. *Proceedings of the 2nd International Conference on Development and Learning.* IEEE Computer Society Press, 239–244

70. P. Sebastiani and M. F. Ramoni (2001), Clustering continuous time series. *Proceeding of the 18th International Conference on Machine Learning (ICML-2001).* Morgan Kaufmann, San Francisco (CA), 497–504.

71. P. Sebastiani and M. F. Ramoni (2001), Bayesian clustering by dynamics of European socio-economic indicators. *Bayesian Methods: Selected Papers from Sixth World Meeting of the International Society for Bayesian Analysis (ISBA)*, European Communities, Luxembourg, 479–488.

72. M. Ramoni and P. Sebastiani (2001), Automated Bayesian modeling. *Bayesian Methods: Selected Papers from Sixth World Meeting of the International Society for Bayesian Analysis (ISBA)*, European Communities, Luxembourg, 429–438

73. H. P. Wynn and P. Sebastiani (2001), The quantization of the attention function under a Bayes information theoretic model. *MaxEnt 2000: Twentieth International Workshop on Bayesian Inference and Maximum Entropy in Science and Engineering.* AIP Conference Proceedings, 159–168.

74. P. Sebastiani and H. P. Wynn, (2001), Design to maximize information. *MaxEnt 2000: Twentieth International Workshop on Bayesian Inference and Maximum Entropy in Science and Engineering.* AIP Conference Proceedings, 192–203.

75. P. Sebastiani, M. F. Ramoni and P. Cohen (2001), Bayesian analysis of sensory inputs of a mobile robot. *Proceedings of the 5th Workshop on Case Studies in Bayesian Statistics.* Lecture Notes in Statistics, Springer, New York, 363–379.

76. M. F. Ramoni, P. Sebastiani and P. Cohen (2000), Multivariate Bayesian clustering by dynamics. *Proceeding of the National Conference of the American Association for Artificial Intelligence (AAAI-2000).* AAAI Press/The MIT Press: Menlo Park/Cambridge, 633–638.

77. P. Sebastiani and M. F. Ramoni (2000), Robust Bayesian classification. *Proceedings of the 14th Conference for Statistical Computing (Compstat 2000). Full papers*, Physica Verlag, Heidelberg, 445–450.

78. P. Sebastiani and M. F. Ramoni (2000), Official statistics via Bayesian networks. *Proceedings of the 14th Conference for Statistical Computing (Compstat 2000). Short Papers*, Statistica Netherlandica.

79. P. Cohen, M. F. Ramoni, P. Sebastiani and J. Warwick (2000), Unsupervised clustering of robot activities: a Bayesian approach. *Proceedings of the 4th International Conference on Autonomous Agents (Agents 2000)*, ACM Press, New York, 134–136.

80. M. F. Ramoni, P. Sebastiani and R. Dybowsky (1999), Robust outcome prediction for intensive-care unit patients. *Proceedings of the AIMDM-99 Workshop on Prognostic Models in Medicine: Artificial Intelligence and Decision Analytic Approaches.* Amsterdam, 45–49.

81. P. Sebastiani, M. F. Ramoni, P. Cohen, J. Warwick and J. Davis (1999), Discovering dynamics using Bayesian clustering. In *Proceedings of the 3rd International Symposium on Intelligent Data Analysis.* Lecture Notes in Computer Science, Springer, New York, 199–210.

82. P. Sebastiani, M. F. Ramoni and P. Cohen (1999), Unsupervised classification of sensory input in a mobile robot. In *Proceedings of International Joint Conference in Artificial Intelligence Workshop on Neural, Symbolic and Reinforcement Methods for Sequence Learning.* AAAI Press, Cambridge, MA, 23–28.

83. P. Sebastiani and M. F. Ramoni (1999), Model folding for data subject to nonresponse. *Proceedings of Uncertainty 99: The 7th International Workshop on Artificial Intelligence and Statistics.* Morgan Kaufman, San Mateo, CA, 287–292.

84. M. F. Ramoni and P. Sebastiani (1999), Learning conditional probabilities from incomplete data: an experimental comparison. *Proceedings of Uncertainty 99: The 7th International Workshop on Artificial Intelligence and Statistics.* Morgan Kaufmann, San Mateo, CA, 260 – 265.

85. P. Sebastiani and M. F. Ramoni (1998), Decision theoretic foundations of graphical model selection. *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence.* Morgan Kaufmann, San Mateo, CA, 464–471.

86. P. Sebastiani and H.P. Wynn (1998), Risk based optimal designs. *Proceedings of the 1998 Conference on Model Oriented Data Analysis and Experimental Design*, Physica Verlag, Heidelberg, 247–257.

87. P. Sebastiani and M. F. Ramoni (1998), Induction of graphical models from incomplete samples. *Proceedings of the 13th Conference for Statistical Computing (Compstat98)*, Physica Verlag, Heidelberg, 449–454.

88. M. F. Ramoni and P. Sebastiani (1997), The use of exogenous knowledge to learn Bayesian networks from incomplete databases. *Proceedings of the 2nd International Symposium on Intelligent Data Analysis.*, Lecture Notes in Computer Science, Springer New York, 537–548.

89. M. F. Ramoni and P. Sebastiani (1997), Learning Bayesian networks from incomplete databases. *Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufman, San Mateo, CA, 401 – 408.

90. M. F. Ramoni and P. Sebastiani (1997), Robust learning with missing data. *Proceedings of the 6th Workshop on AI & Statistics. Preliminary Papers*, Ft. Lauderdale, Florida, 1997, 339–406.

91. R.G. Cowell, A.P. Dawid and P. Sebastiani (1996), A comparison of sequential learning methods for incomplete data. *Bayesian Statistics 5*, Oxford University Press, 533–542.

92. P. Sebastiani and R. Settimi (1994), Planning experiments in non linear situations. *Proceedings of the 9th International Workshop on Statistical Modeling*, Exeter, July 1994.

93. P. Sebastiani and M. Ursini (1994), Modeling the behaviour of a parasitoid insect. *Atti della XXXVII Riunione Scientifica della Società Italiana di Statistica*, SIS, 171–177.

94. A. Forcina and P. Sebastiani (1993), Graphical analysis of conditional independence. *Proceedings of the 8th International Workshop on Statistical Modeling*, Leuven, 85–87.

95. P. Sebastiani (1991), Optimal designs for generalized linear models. *Atti del Convegno SIS "Sviluppi metodologici nei diversi approcci all'inferenza statistica"*, Cagliari, Pitagora Bologna, **1**, 167–170.

3 Book Chapters, Invited Discussions and Editorials

96. M. Montano and <u>P. Sebastiani</u> (2008), Efforts to characterize host response to HIV-1 infection. In *RNA Viruses: Host Gene Responses to Infections*, D. Yang Ed.

97. <u>P. Sebastiani</u>, A. Puca and T. T. Perls (2008), Deciphering the familial and genetic influences upon exceptional longevity. In *Exceptional Longevity*. John Hopkins University Press. In press

98. T. T. Perls and <u>P. Sebastiani</u> (2008), Genetics of exceptional longevity. In *Molecular Biology of Aging*. L. Guarente et al. (Eds) Cold Spring Harbor Laboratory Press.

99. <u>P. Sebastiani</u> and T. T. Perls (2007), Complex genetic models. In *Bayesian Belief Networks: A Practical Guide to Applications*. O. Pourret, P. Nam and B. G. Marcot (Eds). Wiley, NY. 53–72.

100. <u>P. Sebastiani</u> and M. Abad (2007), Bayesian networks for genetics analysis. In *Systems Bioinformatics: An Engineering Case-Based Approach*. G Alterowitz and M. F. Ramoni (Eds). Artech House, Boston. 205–228.

101. <u>P. Sebastiani</u>, J. Jeneralcuk and M. F. Ramoni (2006), Design and analysis of screening experiments with microarrays. In *Screening*, A. Dean and S Lewis (editors). Springer Verlag, NY. 115–138.

102. P. Sebastiani, M. F. Ramoni (2005), Learning Bayesian networks. In *Encyclopedia of Data Warehousing and Mining*. Kluwer, Amsterdam. **2**, 674–678.

103. P. Sebastiani, M. Abad and M. F. Ramoni (2005), Bayesian networks. In *The Data Mining and Knowledge Discovery Handbook*. 193–230.

104. P. Sebastiani, M. Abad and M. F. Ramoni (2005), Bayesian networks for genomic analysis. In *Genomic Signal Processing and Statistics*, Dougherty, E. R., Shmulevich, I., Chen, J., and Wang, Z. J. (Eds.). 281–320.

105. <u>P. Sebastiani</u> and K. Mandl (2004), Biosurveillance and outbreak detection. In *Data Mining: Next Generation Challenges and Future Directions*, H. Kargupta A. Joshi K. Sivakumar and Y. Yesha (Eds). MIT Press. 185–198.

106. <u>P. Sebastiani</u>, M. F. Ramoni and I. Kohane (2003), *Methods in Functional Genomics*. Special issue of the Machine Learning Journal. **52**.

107. M. F. Ramoni and <u>P. Sebastiani</u> (2003), Bayesian methods in intelligent data analysis. In *Intelligent Data Analysis. An Introduction.*, M. R. Berthold and D.J. Hand (Eds). Springer Verlag, New York. 128–166. 2nd Ed. 131–168.

108. <u>P. Sebastiani</u>, M. F. Ramoni and I. Kohane (2003), Bayesian clustering of gene expression dynamics. In *The Analysis of Gene Expression Data: Methods and Software*, Parmigiani, G., Garrett, E. S., Irizarry, R., and Zeger, S. (Eds). Springer Verlag, NY. 400–427.

109. <u>P. Sebastiani</u>, M. F. Ramoni and P. Cohen (2001), Sequence learning via Bayesian clustering by dynamics. In *Sequence Learning: Paradigms, Algorithms, and Applications*, L. Giles and R. Sun, editors Springer Verlag, New York. 11–34.

110. M. F. Ramoni and P. Sebastiani (1999), Bayesian methods in intelligent data analysis. In *Intelligent Data Analysis. An Introduction.*, Eds M. Berthold and D.J. Hand. Springer Verlag, New York. 128–166.

111. P. Sebastiani (1996), Invited discussion to the papers by Atkinson and by Bates et al. Ordinary Meeting of the Royal Statistical Society, April 1995. *Journal of the Royal Statistical Society. Series B*, **58**, 96–97.

4 Other Conference Proceedings and Abstracts

1. N. Timofeev, P. Sebastiani, S. H. Hartley, C. T. Baldwin, and M. H. Steinberg (2008), Fetal Hemoglobin in Sickle Cell Anemia: A Genome-Wide Association Study of the Response to Hydroxyurea Blood (ASH Annual Meeting Abstracts), Nov 2008; 112: 2471.

2. P. Sebastiani, N. Timofeev, S. H. Hartley, D. Dworkis, L. Farrer, C. T. Baldwin, T. T. Perls, and M. H. Steinberg (2008), Genome-Wide Association Studies Suggest Shared Polymorphisms Are Associated with Severity of Sickle Cell Anemia and Exceptional Longevity. Blood (ASH Annual Meeting Abstracts), Nov 2008; 112: 1446.

3. D. Dworkis, P. Sebastiani, E. Melista, J. Parente, G. Lester, J. F Quinn, D. H.K. Chui, M. H. Steinberg, and C. T. Baldwin Fetal Hemoglobin in Sickle Cell Anemia: A Novel Method for High-Resolution Discovery of Associated Genomic Copy Number Variations Blood (ASH Annual Meeting Abstracts), Nov 2008; 112: 2491.

4. P. Sebastiani, L. Wang, T. T. Perls, D. Terry, V. G. Nolan, M. Montano, C. T. Baldwin and M. H. Steinberg (2007), A repertoire of genes that modify the risk of death in sickle cell anemia. *Blood (ASH Annual Meeting Abstracts), Nov 2007.* In press.

5. V. G. Nolan, Y. Zhang, T. Lash, P. Sebastiani and M. H. Steinberg (2007), Association of wind speed and the occurrence of sickle cell acute painful episodes: results of a case-crossover study. *Blood (American Society of hematology, Annual Meeting Abstracts), Nov 2007.* In press.

6. L. Wang, V. Nolan, C. Baldwin, M. H. Steinberg and P. Sebastiani (2007), Discovery of genetic modifiers of a finite domain phenotype using Bayesian beta regression. *Proceedings of the Joint Statistical Meetings, (JSM 2007).* In press.

7. Y. Yang and P. Sebastiani (2007), An evaluation of different experimental designs of temporal microarray experiments. *Proceedings of the Joint Statistical Meetings, (JSM 2007).* In press.

8. A. Ozonoff, J. Ashba and P. Sebastiani (2006), Calculation of excess influenza mortality for small geographic regions. *Proceedings of the Joint Statistical Meetings, (JSM 2006).*

9. A. Ozonoff, S. Sukpraprut and P. Sebastiani (2006), Modeling seasonality of influenza with Hidden Markov Models. *Proceedings of the Joint Statistical Meetings, (JSM 2006).*

10. P. Sebastiani, V. G. Nolan, C. T. Baldwin, M. M. Abad-Grau, L. Wang, A. H. Adewoye, L. C. McMahon, L. A. Farrer, J. G. Taylor, VI, G. J. Kato, M. T. Gladwin, M. H. Steinberg (2006), Severity of sickle cell disease: modeling interrelationships among hemolysis, pulmonary hypertension and risk of death. In *Blood (American society of hematology, Annual Meeting Abstracts), Nov 2006*, 108: 1221.

11. P. Sebastiani, M. M. Abad-Grau, A. Riva, V. G. Nolan, E. Melista, M. H. Steinberg and C. T. Baldwin (2006), Fetal hemoglobin (HbF) in sickle cell anemia: genome-wide association studies using pooled dna samples can reveal genetic associations with HbF concentration. In *Blood (American Society of Hematology, Annual Meeting Abstracts), Nov 2006*, 108: 786.

12. P. Sebastiani, E. Hadley, M. Province, A. Yashin, K. Christensen, J. Vaupel, C. Kammerer, W. Rossi, T. Perls, A. American Society of Hematology (2006), A family longevity selection score. *Proceedings of the 2006 Meeting of the American Society of Human Genetics.* 411

13. P. Sebastiani, C. T. Baldwin, V. Nolan, D. F. Wyszynski, Q. Ma, J. Farrell, A. Bisbee, K. Waraska, L. A. Farrer, M. H. Steinberg (2005), Polymorphisms (SNPs) in multiple genes of the TGF-BMP pathway are associated with a global measure of sickle cell disease severity. *Blood. (American Society of Hematology, Annual meeting abstract), Nov 2005*, 106: 74.

14. V. G. Nolan, M. A. Wilcox, P. Sebastiani, C. T. Baldwin, D. F. Wyszynski, Q. Ma, A. Bisbee, J. J. Farrell, L. A. Farrer, and M. H. Steinberg (2005), Gene-gene interactions and the pathophysiology of sickle cell disease: modeling the effects of snps on sickle cell-associated vasoocclusive events using classification and regression trees and stochastic gradient boosting. *Blood (American Society of Hematology, Annual Meeting Abstracts), Nov 2005*, 106: 3183.

15. A. H. Adewoye, V. G. Nolan, C. T. Baldwin, D. F. Wyszynski, Q. Ma, J. J. Farrell, A. Bisbee, E. Homan, P. Sebastiani, L. A. Farrer and M. H. Steinberg (2005), Association of polymorphisms of the transforming growth factor-/bone morphogenetic protein (tgf-/bmp) pathway with sickle cell bacteremia. *Blood (American Society of Hematology, Annual Meeting Abstracts), Nov 2005*, 106: 3170.

16. V. G. Nolan, P. Sebastiani, C. T. Baldwin, D. F. Wyszynski, Q. Ma, A. Bisbee, J. Farrell, L. A. Farrer, M. H. Steinberg (2005), Modeling genetic polymorhphisms and sickle cell associated vasooclusive events using classification and regression trees (CART). *28th Annual Meeting of the National Sickle Cell Disease.*

17. P. Sebastiani, V. Nolan, C. T. Baldwin and M. H. Steinberg (2005), Modeling phenotype interactions in sickle cell disease. *28th Annual Meeting of the National Sickle Cell Disease.*

18. L. Wang, P. Sebastiani and M. F. Ramoni (2005), Gene expression pattern identification using bayesian clustering by dynamics. *Proceedings of Genomics Signal Processing and Statistics, 2005.*

19. F. Ferrazzi, P. Sebastiani, L. Wang, R. Bellazzi, I. S. Kohane and M. F. Ramoni (2005), Discovering feedback control mechanisms in gene expression. *Proceedings of Genomics Signal Processing and Statistics, 2005.*

20. P. Sebastiani, V. Nolan, C. T. Baldwin and M. H. Steinberg (2004), Modeling phenotype interactions in sickle cell anemia. *Blood (American Society of Hematology Annual Meeting Abstracts), Nov 2004*, **104:**11, 460a.

21. P. Sebastiani, M. F Ramoni, V. Nolan, C. T. Baldwin and M. H. Steinberg (2004), Multigenic dissection and prognostic modeling of overt stroke in sickle cell anemia. *Blood. (American Society of Hematology, Annual Meeting Abstracts), Nov 2004.* **104:**11, 462a.

22. P. Sebastiani and M.F. Ramoni (2003), Bayesian differential analysis of gene expression data. *Proceedings of the Joint Statistical Meeting. (JSM 2002).* Electronic publication.

23. M. F. Ramoni and P. Sebastiani (2001), Bayesian bioinformatics. *Proceedings of the Workshop "Bayesian models in medicine"*. Cascais, Pt, 2001. Page 1-4.

24. P. Sebastiani and M. F. Ramoni (1999), Iterative approximation of the marginal likelihood from incomplete data. *Book of Abstracts of the 2nd European Conference on Highly Structured Stochastic Systems*, University of Pavia, 262–264.

25. M. F. Ramoni, P. Sebastiani and P. Cohen (1999), Bayesian analysis of sensory inputs of a mobile robot. *Book of Abstracts of the 2nd European Conference on Highly Structured Stochastic Systems*, University of Pavia, 235–236.

26. R. G. Cowell, M. Ramoni and P. Sebastiani (1999), An experimental evaluation of the predictive accuracy of Bayesian networks. *Book of Abstracts of the 2nd European Conference on Highly Structured Stochastic Systems*, University of Pavia, 75–76.

27. P. Sebastiani and H.P. Wynn (1997), Bayesian experimental design and Shannon information. *Proceedings of the Joint Statistical Meeting (JSM 1997)*, 176 –181.

28. M.L. Eaton, A. Giovagnoli and P. Sebastiani (1993), A predictive approach to the Bayesian design problem with application to normal regression models. *Bulletin of the International Statistical Institute, Proceedings 49th Session*, Contributed papers, **1**, 405–406.

5 Book Review

29. A. Giovagnoli and P. Sebastiani (1992), Review of the book by A.C.Atkinson and Donev, "Optimum Experimental Designs". *Journal of the Italian Statistical Society*, **1**, 425–426.

6 Bulletin and Newsletters

30. P. Sebastiani (2002), Caged: cluster analysis of gene expression dynamics. *ISBA Bulletin*, **9**(2), 11–12.

31. P. Sebastiani (2002), Computational challenges in the post-genome era. *Department Newsletter.* Department of Mathematics and Statistics, University of Massachusetts, Amherst MA.

7 Theses and Dissertations

32. P. Sebastiani (1992), *Piani di Esperimento Ottimi per Modelli Lineari Generalizzati.* Dissertazione Dottorale, Dottorato di Ricerca in Statistica, V Ciclo.

33. P. Sebastiani (1987), *Modelli Lineari con Piu' Componenti di Varianza.* Tesi di Laurea. Facolta di Scienze, Unversità di Perugia, Italy.

9. Presentations

1 Invited Conference Presentations

[1] Bayesian Modeling of Complex Traits: Diagnostic Versus Prognostic Models. 2009 Bayesian Biostatistics Conference, Houston, January 2009.

[2] After the first wave of genome wide association studies: whats next? IDAMAP 2008, Washington DC, November 2008.

[3] Candidate and genome wide association studies. 20th Anniversary Celebration of the Georgia Centenarian Study, Atlanta, August 2008.

[4] Leveraging pleiotropy to increase the power of genome wide association studies. Computational Models for Analyzing Genotype-Phenotype Associations in Rare Diseases. Bethesda, July 2008.

[5] Gene Networks. 68th meeting of the ADA, San Francisco June 2008.

[6] Bayesian modeling of complex traits. ENAR 2008. Crystal City, March 2008.

[7] The new genetics of sickle cell anemia. 2007 Red Blood Cell Meeting. Harvard Medical School. Boston October 2007.

[8] Graphical models for temporal data. Annual conference of the International Biometric Society of the Eastern North American Region, Atlanta 2007.

[9] Bayesian modeling of complex traits. Joint Statistical Meeting, August 2006. Seattle WA

[10] Modeling gene environment interaction. Workshop on Genetic Regulatory Network. College Station, November 2005.

[11] Genes and networks. Genomics Signal Processing and Statistics 2004, Baltimore, May 2004.

[12] Decoding gene expression control using generalized gamma networks. Virtual Seminars on Genomics and Bioinformatics, University of South Dakota, May 2003.

[13] Decoding gene expression control using generalized gamma networks. New England Statistics Symposium. Storrs, April 2003.

[14] Decoding gene expression control using generalized gamma networks. Samsi Workshop: Graphical Models. Research Triangle Park, NC. February 2003.

[15] The statistical insignificance of fold ratio in the analysis of gene expression. UMass Intercampus Bioinformatics Conference. Boston, MA. May 2001.

[16] Experimental design to maximize information. MaxEnt 2000: 20th International Workshop on Bayesian Inference and Maximum Entropy in Science and Engineering, Gif sur Yvette, France. July 2000.

[17] Bayesian clustering by dynamics of european socio-economic indicators. International Society for Bayesian Analysis, Creete, Greece. May 2000.

[18] Knowledge discovery via Bayesian networks. Statistical Computing Section of the Royal Statistical Society. London, UK. February 2000.

[19] Bayesian experimental design and Shannon information. International Society for Bayesian Analysis, Istambul, Turkey. August 1997.

[20] Maximum entropy sampling and optimal Bayesian experimental design. Royal Statistical Society, General Applications Section. London, UK. March 1996.

[21] Invited to second the vote of thanks on papers by Atkinson and Bates et al. Ordinary Meeting of the Royal Statistical Society. London, UK. April 1995.

2 Conference Presentations

[22] Novel Genetic Modifiers of the Severity of Sickle Cell Anemia Identified Through A Genome-Wide Association Study. 3rd Annual Sickle Cell Disease Scientific Meeting. Ft Lauderdale, February 2009.

[23] Genome-Wide Association Studies Suggest Shared Polymorphisms Are Associated with Severity of Sickle Cell Anemia and Exceptional Longevity. ASH 2008. San Francisco, Decmber 2008.

[24] Phenotypic complexity of sickle cell anemia and exceptional longevity share common genetic variants. Genomics of Complex Diseases. Boston September 2008.

[25] Results from GWAS in Centenarians. Meeting of the "Longevity Consortium". Boulder, June 2008.

[26] A repertoire of genes that modify the risk of death in sickle cell anemia. Annual meeting of the American Society of Hematology 2007. Atlanta, December 2007.

[27] Genetics of Exceptional Longevity. Meeting of the "Longevity Consortium". Bethesda, December, November 2007.

[28] Genetics of Exceptional Longevity: The New England Centenarian Study. Gerontological Society of America: Annual Meeting. San Francisco, November 2007.

[29] Bayesian modeling of complex traits. Emerging Design and Analysis Issues in Genomic Studies in Population Science. Harvard School of Public Health, Boston, October 2007.

[30] Severity of sickle cell disease: modeling interrelationships among hemolysis, pulmonary hypertension and risk of death. Annual meeting of the American Society of Hematology 2006. Orlando, December 2006.

[31] A family longevity selection score. Annual meeting of the American Society of Human genetics. New Orleans, October 2006.

[32] Chromosome 6q22-25 is associated with hbf concentration in sickle cell anemia (hbss): a dense snp analysis using pooled dna. Annual Sickle Cell meeting 2006. Memphis, April 2006.

[33] Clustering short gene expression profiles. 11th Annual conference on "Research in computational molecular biology". Venice, April 2006.

[34] Polymorphisms (Snps) in multiple genes of the TGF-BETA/BMP pathway are associated with a global measure of sickle cell disease severity. Annual meeting of the American Society of Hematology 2005. Atlanta, December 2005.

[35] Modeling phenotype interactions in sickle cell disease. Annual Sickle Cell meeting 2005. Cincinnati, April 2005.

[36] Multigenic dissection and prognostic modeling of overt stroke in sickle cell anemia. Annual meeting of the American Society of Hematology 2004. San Diego, December 2004.

[37] Modeling phenotype interactions in sickle cell disease. Annual meeting of the American Society of Hematology 2004. San Diego, December 2004.

[38] Bayesian network modeling of complex traits in association studies. Harvard Genomics Conference, Boston May 2004.

[39] Discovery of complex traits by Bayesian network modeling. 27th Sickle Cell Disease Symposium, Los Angeles, April 2004.

[40] A Dynamic graphical model for influenza surveillance. Annual conference of the International Biometric Society of the Eastern North American Region 2004, Pittsburgh, March 2004.

[41] Screening experiments with microarrays. New Directions in Design and Analysis of Experiments, Chicago, May 2003.

[42] Bayesian differential analysis of gene expression data, Joint Statistical Meeting 2002, New York, August 2002

[43] The statistical insignificance of fold ratio in the analysis of gene expression. Fifteenth New England Statistics Symposium. Storrs. April 2001.

[44] Profiling customers from in-house data. 5th Workshop on Case Studies in Bayesian Statistics. Pittsburgh, PA, September 1999.

[45] Iterative approximation of the marginal likelihood from incomplete data. 2nd European Conference on Highly Structured Stochastic Systems, Pavia, Italy, September 1999.

[46] Robust outcome prediction for intensive-care unit patients. Prognostic Models in Medicine: Artificial Intelligence and Decision Analytic Approaches, Aalborg, DK, June 1999.

[47] Model folding for data subject to nonresponse. Uncertainty99, Ft Lauderdale, Fl. January 1999.

[48] Bayesian knowledge discoverer. Competive Research Prototype Presentation of the 1998 Conference on Knowledge Discovery and Data mining. New York, August 1998.

[49] Decision Theoretic Foundations of Graphical Model Selection. Fourteen Conference on Uncertainty in Artificial Intelligence. Madison, WI, July 1998.

[50] Learning bayesian networks from incomplete databases. Thirteenth Conference on Uncertainty in Artificial Intelligence Providence, RI, July 1997.

[51] Learning graphical models from incomplete data. Practical Bayesian Statistics 4, Nottingham, July 1997.

[52] Robust parameter learning in bayesian networks with missing data. Sixth Workshop on AI and Statistics, Ft Lauderdale, FL. January 1997.

[53] Maximum entropy sampling and optimal bayesian experimental design. 4th World Congress of the Bernoulli Society, Vienna, August 1996.

[54] Computational methods for optimal designs for non linear problems. European Meeting of Statisticians, Aarhus, August 1995.

[55] Computational complexity in bayesian optimal design. Workshop on Model Uncertainty, Bath, July 1995.

[56] A comparison of learning methods for incomplete data. International Meeting of Bayesian Statistics, Alicante, June 1994.

[57] A Comparison of Learning Methods for Incomplete Data. HSSS Workshop, Cortona, April 1994.

[58] Modelling the behaviour of a parasitoid insect. XXXVII Riunione Scientifica SIS, Sanremo, April 1994.

[59] Graphical analysis of conditional independence. 8th International Workshop on Statistical Modeling, Leuven, July 1993.

[60] D-optimal designs for generalized linear models with constant coefficient of variation. Royal Statistical Society Conference, Sheffield, September 1992.

[61] Piani di esperimento D-ottimi per modelli di regressione logistica. XXXVI Riunione Scientifica SIS, Pescara, April 1992.

[62] Optimal design for generalized linear models. European Meeting of Statisticians, Barcellona, August 1991.

3 Invited Seminars

[63] Prognostic modeling of complex traits. University of Michigan, Ann Arbor, April 2007.

[64] Bayesian networks for the genetic dissection of complex traits. Northern Illinois University. Graduate Colloquium. Dekalb, October 2006

[65] The challenge of complex traits. Northern Illinois University. Graduate Colloquium. Dekalb, October 2006

[66] Genetic Dissection of Complex Traits. NHLBI Advisory Council. Bethesda, February 2006.

[67] Genetic dissection of complex traits by Bayesian networks. Department of Biostatistics, Yale University. October 2005.

[68] Genetic dissection of stroke in sickle cell anemia, Advisory Board, NHLBI. Bethesda, June 2005.

[69] Genes and networks. Department of Biostatistics, Harvard School of Public Health. February 2005.

[70] Design and analysis of microarray experiments. Rheumatology Grand Round, BMC, February 2005.

[71] Genes and networks. Boston Chapter of the American Statistical Association. October 2004.

[72] Challenges in the analysis of microarray experiments. Beyond Genomics. July 2004.

[73] Genes and networks. Department of Biostatistics. Johns Hopkins University, February 2004.

[74] An integrated model for influenza surveillance. Department of Biostatistics, Harvard School of Public Health. January 2004.

[75] Challenges in the analysis of microarray experiments. Center for Imaging Science. Johns Hopkins University. November 2003.

[76] Children are sentinels of influenza illness and mortality. Department of Biostatistics. Boston University. October 2003.

[77] Statistical challenges in the design and analysis of microarray data. Department of Statistics. GlaxoSmithKline. August 2003.

[78] Computational challenges in the post-genome era. Department of Statistics and Decision Sciences, Duke University, Durham NC. February 2003.

[79] Computational challenges in the post-genome era. Department of Genetics and Genomics, Boston University, Boston MAS. January 2003.

[80] Bayesian analysis of gene expression data. Department of Biostatistics, Boston University, Boston MA. January 2003.

[81] BADGE: Bayesian analysis of differential gene expression. Cancer Genomics Group: Whitehead Institute, Cambridge MA. November 2002.

[82] Statistical challenges in the post-genome era. Department of Mathematics and Statistics, Boston University, Boston MA. November 2002.

[83] Bayesian differential analysis of gene expression data. Department of Statistics, University of Connecticut, Storrs CN. April 2002.

[84] Bayesian differential analysis of gene expression data. Grand Round Series: Harvard Medical School, Boston MA. March 2002.

[85] Stochastic analysis of gene expression data. Children's Hospital Informatics Program., Boston MA January 2002.

[86] Bayesian Clustering By Dynamics. Department of Statistics and Mathematics. University of Chicago, Chicago IL. April 2001.

[87] Knowledge discovery via Bayesian networks. Department of Mathematics and Statistics. University of Massachusetts at Amherst, Amherst MA. December 2000.

[88] Learning Bayesian Networks from Data. Department of Computer Science. University of Massachusetts at Amherst, Amherst MA. October 2000.

[89] Knowledge discovery via Bayesian networks. Office of National Statistics, London UK. February 2000.

[90] Bayesian clustering by dynamics. Department of Mathematics and Statistics. University of Massachusetts at Amherst, Amherst MA. February 2000.

[91] Robust Bayesian classification. Faculty of Mathematical Sciences, University of Southampton, Southampton UK. December 1999.

[92] Robust Bayesian classification. Department of Statistical Sciences, University College London, London UK. November 1999.

[93] Finding episodes boundaries. Department of Computer Science, University of Massachussets in Amherst, Amherst MA. October 1999.

[94] Bayesian knowledge discovery. Department of Statistics, University of Nottingham, Nottingham, UK. December 1998.

[95] Introduction to Bayesian inference. Cycle of Seminars at the Department of Computer Science, University of Massachussets in Amherst, Amherst MA. August 1998.

[96] Bayesian experimental design. Department of Mathematical Sciences, University of Durham. May 1997.

[97] Robust learning with missing data. Department of Actuarial Science and Statistics, City University, London UK. February 1997.

[98] BUGS: bayesian analysis using gibbs sampling. Department of Actuarial Science and Statistics, City University, London UK. November 1996.

[99] Maximum entropy sampling and optimal Bayesian experimental design. Department of Statistics, University of Glasgow, UK. June 1996.

[100] A coherent approach to optimal design of experiments. Department of Statistical Science University College of London, UK. December 1994.

[101] A coherent approach to optimal design of experiments. Department of Mathematics and Statistics. University of Plymouth, Plymouth UK. November 1994.

[102] Modeling the behaviour of insects via markov processes. Department of Entomology, University of Perugia, Perugia IT. May 1993.

[103] Optimal designs for generalized linear models. Department of Statistical Science, University College London, London UK. October 1992.

[104] An introduction to Bayesian networks, Department of Mathematics, University of Perugia, Perugia It. 1990.

4 Tutorials

[105] BioSecurity 2003–Workshop: Syndromic Surveillance: Outbreak Detection and Disease Monitoring. Washington DC, June 2003.

[106] BioSecurity 2003–Workshop: Syndromic Surveillance: Outbreak Detection and Disease Monitoring. Washington DC, October 2003.

10. Teaching Experience

1 Graduate Courses

[1] BAYESIAN MODELING IN BIOMEDICAL RESEARCH (New graduate course introduced in Fall 2006) An upper level graduate course that introduces advanced Bayesian statistical methods to analyse biomedical data using stochastic computations.

[2] STATISTICAL METHODS IN EPIDEMIOLOGY (Fall 2003, Spring 2004, Spring 2005) An upper level graduate course that introduces advanced statistical methods to analyse epidemiology data in the presence of confounding.

[3] DESIGN AND ANALYSIS OF MICROARRAY EXPERIMENTS (New graduate course introduced in Fall 2004, fall 2005, fall 2007) An upper level graduate course that introduces advanced statistical methods for the design and analysis of microarray experiments.

[4] STATISTICAL METHODS IN BIOINFORMATICS (New graduate course introduced in Spring 2003) An upper level graduate course that introduces basic biology, the technology of microarrays and statistical and machine learning techniques for the analysis of microarray data. The course also describes databases and sequencing.

[5] STATISTICAL MODELING (1996-1997, 1998-1999, 1999-2000). MSc in Social Research Methods and Statistics, City University. The course provides an overview of statistical methods for the study of association between variables. Emphasis is on graphical models. Software: Glim4.

[6] STATISTICAL MODELING (1996-1997); EXPERIMENTAL DESIGN AND DATA ANALYSIS (1995-1997); ROBUST ENGINEERING DESIGN AND TAGUCHI METHODS (1995-1998). MSc in Quality Management, Statistical Methods and Reliability, City University. These highly intensive courses provide an introduction to statistical methods for data analysis and design of experiments. Software: Minitab.

[7] INTRODUCTION TO SPLUS (1994). MSC in Statistics, London School of Economics and Political Science, London.

[8] GRAPHICAL MODELING (1993). Graduate course for the PhD in Statistical Methods for social and economic sciences, University of Perugia, Italy. The course gives an introduction to graphical models for the analysis of association between variables. Software: Glim4.

2 Undergraduate Courses

[9] DESIGN OF EXPERIMENTS STAT506 (Spring 2002, 2003). University of Massachusetts at Amherst. A senior/graduate course, which provides an overview of experimental design and analysis of experimental data. Topics include one and two way anova, factorial experiments, block designs.

[10] BAYESIAN DATA ANALYSIS STAT597E (Fall 2001). University of Massachusetts at Amherst. A senior/graduate course, which provides an overview of Bayesian methods and Bayesian graphical models in modern data analysis. Topics include overview of Bayesian methods in estimation and hypothesis testing, multiple regression, graphical models and Bayesian computations.

[11] BAYESIAN INFERENCE AND DECISION THEORY (1995-1997). BSc in Actuarial Science and Statistical Science with Management Studies, City University, London. The course provides

an overview of Bayesian methods and Bayesian graphical models in statistical inference and decision theory, with emphasis on actuarial and financial applications.

[12] EXPERIMENTAL DESIGN (1995-1997). BSc in Actuarial Science, Statistical Science with Management Studies, Mathematical Sciences with Statistics, City University, London. The course provides an overview of the design and analysis of factorial experiments. Software: Minitab.

[13] DATA ANALYSIS II (1997). BSc in Actuarial Science, Statistical Science with Management Studies, Mathematical Sciences with Statistics, City University, London. The course aims to introduce the students to Generalized Linear Models, with particular focus on the study of associations between variables, and the analysis of prospective studies. Software: Splus.

[14] COMPUTATIONAL STATISTICS (1990 – 1994). BSc in Economics, University of Perugia, Italy. This is a second course in Statistics for students of Economics, and it aims to improve students skills in regression analysis. Software: Splus/SAS.

3 Service Teaching

[15] STAT140 (Fall 2002). Introduction to statistics. University of Massachusetts at Amherst. An introductory course to statistics for students with no mathematical background.

[16] S501 (Fall 2000) Methods of Applied Statistics. University of Massachusetts at Amherst. An applied statistics course for graduate students and upper level undergraduates with no background in statistics who will need statistics in their further studies and their work. The focus is on understanding and using statistical methods in research and applications.

[17] S141 (Fall 2000) Quantitative Concepts for Management. School of Management Major. University of Massachusetts at Amherst. The course provides an introduction to statistical methods for business studies. Software: Minitab.

[18] FOMGT250 (Spring 2000) School of Management Major. University of Massachusetts at Amherst. The course provides an introduction to advanced statistical methods for business studies. Software: Minitab.

[19] STATISTICS (1997-1998) BSc in Mechanical Engineering, Air Transport Engineering, City University, London. The course provides an introduction to probability and statistical methods, with particular focus on the analysis of experimental data. Software: Minitab.

[20] STATISTICS AND COMPUTING (1995-1997) BSc in Insurance and Investment, City University, London. This is a second course in statistics, with emphasis on regression techniques. Software: Minitab.

[21] STATISTICAL INFERENCE (1993-1994). BSc in Mathematics, University of Perugia, Italy. The course gives an overview of both classical and Bayesian inference.

[22] OPERATIONAL RESEARCH (1990-1991). Diploma in Informatics, Faculty of Economics, University of Perugia, Italy. The course aims to improve students skills in optimization techniques. Software: Derive.

4 Distance Teaching

[23] Statistics and Society (1998-1999). First course in Statistics provided by the Open University.

[24] Statistical Modeling (1998-1999). Second course in Statistics provided by the Open University. The course gives an introduction to Multiple linear regression and Generalised Linear Models. Software: Genstat.

5 Graduate Students Supervised

1993–1996. Member PhD committee, candidate M. Baioletti, University of Perugia, Italy.

2003–2006. Member PhD committee, candidate Toni Rath, Department of Computer Science, University of Massachusetts, Amherst.

2003–2007. Primary advisor, PhD candidate Ling Wang, Biostatistics, Boston University.

2004–2006. Member PhD committee, candidate Gang Xu, Bioinformatics Program, Boston University.

2004–2007. Member PhD committee, candidate Jennifer Bean, Bioinformatics Program, Boston University.

2004–2007. Member PhD committee, candidate Vishal Sha, Bioinformatics Program, Boston University.

2006–2008. Member PhD committee, candidate Yi Yu, Bioinformatics Program, Boston University.

2006 to date. Primary advisor, PhD candidate Vikki Nolan, Epidemiology, Boston University.

2006 to date. Primary advisor, PhD candidate Amanda Sedgewick, Bioinformatics Program, Boston University.

2006 to date. Primary advisor, PhD candidate Zhenming Zhao, Biostatistics, Boston University.

2006 to date. Primary advisor, PhD candidate Yijun Yang, Biostatistics, Boston University.

2007 to date. Primary advisor, PhD candidate Nadia Timofeev, Biostatistics, Boston University.

2008 to date. Member PhD committee, candidate Kelly Graham, Md/PhD program in Molecular Medicine, Boston University.

2008 to date. Member PhD committee, candidate Lyndsey A. Emery, Md/PhD program in Molecular Medicine, Boston University.

1996 to date. Advisor of Master students: A. Sinclair, A. John, MSc, City University, London, UK, M Qin, S. Bath, L. Mookanamtarambil, L. Wang, University of Massachusetts, Claire Reynolds, Boston University.

11. Administration

SEMINAR ORGANIZER	City University, London (1995–1996). Imperial College, London (2000), University of Massachusetts, Amherst (2001).
LIBRARY OFFICER 1996–1998	City University, London.
PUBLICATION OFFICER 1996–1998	Administrator of the departmental research web site. City University, London.
FACULTY SEARCH COMMITTEE	University of Massachusetts, Amherst (2002–2003); Boston University School of Public Health (2004–2005, 2007–2008)
GRADUATE SCHOOL PROGRAM 2006 to date	Member. Boston University School of Public Health
EXECUTIVE COMMITTEE 2006 to date	Member. Boston University School of Public Health

February 21, 2009