

# Alexander Statnikov

**E-mail:** [Alexander.Statnikov@med.nyu.edu](mailto:Alexander.Statnikov@med.nyu.edu)

**Website:** <http://www.nyuinformatics.org>

**Office:** 1 (212) 263-3641

**Fax:** 1 (212) 263-5995

**Cell:** 1 (615) 545-3685

**Work address:**

227 East 30<sup>th</sup> Street, 7<sup>th</sup> Floor, Office 736,  
Center for Health Informatics and Bioinformatics,  
New York University Langone Medical Center,  
New York, NY 10016, USA

---

## Research Interests

- Developing, testing, and applying computational causal discovery methods for biomedical and other data of high-dimensionality;
- Building, optimizing, and validating predictive models;
- Conducting comprehensive benchmarking studies of methods for analysis of high-dimensional data.

## Current Academic Affiliation and Positions

- **Assistant Professor**, *Center for Health Informatics and Bioinformatics*, July 2009 - Present  
*Department of Medicine, Division of Translational Medicine*,
- **Director**, *Computational Causal Discovery Laboratory*,
- **Benchmarking Director**, *Best Practices Informatics Consultation Service*,
- **Affiliate**, *Computational Biology Graduate Training Program, Sackler Institute of Graduate Biomedical Sciences*,
- **Affiliate**, *Pathobiology and Translational Medicine Training Program, Department of Pathology*  
New York University Langone Medical Center, New York, New York.

## Education

- Ph.D.**, Biomedical Informatics (GPA 3.9), *Vanderbilt University*, Nashville, Tennessee, December 2008
- Master of Science**, Biomedical Informatics (GPA 3.9), *Vanderbilt University*, Nashville, Tennessee, August 2005
- Master of Science**, Applied Mathematics (GPA 3.9), *Case Western Reserve University*, Cleveland, Ohio, August 2002
- Bachelor of Science**, Mathematics (GPA 3.6), *Case Western Reserve University*, Cleveland, Ohio, August 2001
- Applied Mathematics (GPA 3.9), *Moscow State Institute of Electronics and Mathematics*, Moscow, Russia, September 1996 - July 1999

## Professional Work Experience

- Senior Scientific Programmer**, *Discovery Systems Laboratory, Department of Biomedical Informatics, Vanderbilt University Medical Center*, Nashville, Tennessee, May 2002 - May 2009
- Scientific Programmer**, *Zyxbio, LLC*, Cleveland, Ohio, February 2002 - May 2002
- Research Assistant and Programmer**, *Department of Mathematics, Case*, May 2000 -

Western Reserve University, Cleveland, Ohio.

**Research Assistant and Programmer**, *Mechanical Engineering Research Institute, Russian Academy of Sciences*, Moscow, Russia.

**Programmer and System Administrator**, *NPO Stroysservice Ltd.*, Moscow, Russia.

May 2002

February 1997 -  
August 1999

December 1998 -  
August 1999

## Consulting

- **Prediction Sciences, LLC** (La Jolla, California): Proteomic-based stroke diagnosis biomarker discovery and diagnostic signatures development.
- **Sanctis, LLC** (Leesburg, Virginia): Molecular signatures development.
- **Discovery Holdings, LLC** (Nashville, Tennessee): Machine learning algorithms development.

## Publications

### Peer-Reviewed Articles\*:

1. Lytkin NI, McVoy L, Weitkamp JH, Aliferis CF, **Statnikov A**. Expanding the Understanding of Biases in Development of Clinical-Grade Molecular Signatures: A Case Study in Acute Respiratory Viral Infections. *PLoS ONE*, 2011; 6(6): e20662.
2. Alekseyenko AV, Lytkin NI, Ai J, Ding B, Padyukov L, Aliferis CF, **Statnikov A**. Causal Graph-Based Analysis of Genome-Wide Association Data in Rheumatoid Arthritis. *Biology Direct*, 2011 May; 6(1): 25.
3. Shmelkov E, Tang Z, Aifantis I, **Statnikov A**. Assessing Quality and Completeness of Human Transcriptional Regulation Pathways on a Genome-Wide Scale. *Biology Direct*, 2011 Feb; 6(1): 15.
4. Narendra V, Lytkin NI, Aliferis CF, **Statnikov A**. A Comprehensive Assessment of Methods for De-Novo Reverse-Engineering of Genome-Scale Regulatory Networks. *Genomics*, 2011 Jan; 97(1): 7-18.
5. Aliferis CF, Alekseyenko AV, Aphinyanaphongs Y, Brown S, Fenyo D, Fu L, Shen S, **Statnikov A**, Wang J. Trends and Developments in Bioinformatics in 2010: Prospects and Perspectives. (in press) *IMIA Yearbook of Medical Informatics*, 2011.
6. Lemeire J, Meganck S, Cartella F, Liu T, **Statnikov A**. Inferring the Causal Decomposition under the Presence of Deterministic Relations. *Proceedings of the 19<sup>th</sup> European Symposium on Artificial Neural Networks (ESANN)*, 2011.
7. Guyon I, **Statnikov A**, Aliferis CF. Time Series Analysis with the Causality Workbench. *Journal of Machine Learning Research Workshop and Conference Proceedings, Volume 12: Causality in Time Series (NIPS 2009)*, 12: 115-139, 2011.
8. **Statnikov A**, Lytkin NI, McVoy L, Weitkamp JH, Aliferis CF. Using Gene Expression Profiles from Peripheral Blood to Identify Asymptomatic Responses to Acute Respiratory Viral Infections. *BMC Research Notes*, 2010 Oct; 3(1): 264.
9. Espinosa L, Cathelin S, D'Altri T, Trimarchi T, **Statnikov A**, Guiu J, Rodilla V, Inglés-Esteve J, Nomdedeu J, Bellosillo B, Besses C, Abdel-Wahab O, Kucine N, Sun SC, Song G, Mullighan CC, Levine RL, Rajewsky K, Aifantis I, Bigas A. The Notch/Hes1 Pathway Sustains NF- $\kappa$ B Activation through CYLD Repression in T Cell Leukemia. *Cancer Cell*, 2010 Sep; 18(3): 268-281.
10. **Statnikov A**, Aliferis CF. Analysis and Computational Dissection of Molecular Signature Multiplicity. (**Cover Article**) *PLoS Computational Biology*, 2010 May; 6(5): e1000790.

---

\* Following the practice in biomedical informatics and computer science, this section includes peer-reviewed articles published in journals and in major, rigorously peer-reviewed biomedical informatics and computer science conference proceedings (the acceptance rates of most computer science conferences in the present vitae range from less than 25% to less than 5%).

11. **Statnikov A**, McVoy L, Lytkin N, Aliferis CF. Improving Development of the Molecular Signature for Diagnosis of Acute Respiratory Viral Infections. *Cell Host & Microbe*, 2010 Feb; 7(2): 100-1.
12. Guyon I, Pellet JP, **Statnikov A**. Development of Projects for the Causality Workbench. *Proceedings of the AAAI Artificial Intelligence and Development (AI-D) Spring Symposium*, 2010.
13. Aliferis CF, **Statnikov A**, Tsamardinos I, Mani S, Koutsoukos X. Local Causal and Markov Blanket Induction Algorithms for Causal Discovery and Feature Selection for Classification. Part I: Algorithms and Empirical Evaluation. *Journal of Machine Learning Research*, 11(Jan): 171-234, 2010.
14. Aliferis CF, **Statnikov A**, Tsamardinos I, Mani S, Koutsoukos X. Local Causal and Markov Blanket Induction Algorithms for Causal Discovery and Feature Selection for Classification. Part II: Analysis and Extensions. *Journal of Machine Learning Research*, 11(Jan): 235-284, 2010.
15. **Statnikov A**, Aliferis CF. TIED: An Artificially Simulated Dataset with Multiple Markov Boundaries. *Journal of Machine Learning Research Workshop and Conference Proceedings, Volume 6: Causality: Objectives and Assessment (NIPS 2008)*, 6: 249-256, 2010.
16. Mani S, Aliferis CF, **Statnikov A**. Bayesian Algorithms for Causal Data Mining. *Journal of Machine Learning Research Workshop and Conference Proceedings, Volume 6: Causality: Objectives and Assessment (NIPS 2008)*, 6: 121-136, 2010.
17. Aliferis CF, **Statnikov A**, Tsamardinos I, Schildcrout JS, Shepherd BE, Harrell FE. Factors Influencing the Statistical Power of Complex Data Analysis Protocols for Molecular Signature Development from Microarray Data. *PLoS ONE*, 2009; 4(3): e4922.
18. Fananapazir N, **Statnikov A**, Aliferis CF. The FAST-AIMS Clinical Mass Spectrometry Analysis System. *Advances in Bioinformatics*, vol. 2009, Article ID 598241, 2009.
19. **Statnikov A**, Wang L, Aliferis CF. A Comprehensive Comparison of Random Forests and Support Vector Machines for Microarray-Based Cancer Classification. *BMC Bioinformatics*, 2008; 9: 319.
20. **Statnikov A**, Li C, Aliferis CF. A Statistical Reappraisal of the Findings of an Esophageal Cancer Genome-Wide Association Study. *Cancer Research*, 2008; 68: 3074-3075.
21. Guyon I, Aliferis CF, Cooper GF, Elisseeff A, Pellet JP, Spirtes P, **Statnikov A**. Design and Analysis of the Causation and Prediction Challenge. *Journal of Machine Learning Research Workshop and Conference Proceeding, Volume 3: Causation and Prediction Challenge (WCCI 2008)*, 3: 1-33, 2008.
22. **Statnikov A**, Li C, Aliferis CF. Effects of Environment, Genetics and Data Analysis Pitfalls in an Esophageal Cancer Genome-Wide Association Study. *PLoS ONE*, 2007; 2(9): e958.
23. **Statnikov A**, Aliferis CF. Are Random Forests Better than Support Vector Machines for Microarray-Based Cancer Classification? *Proceedings of the AMIA Annual Symposium*, 2007.
24. Aliferis CF, **Statnikov A**, Tsamardinos I. Challenges in the Analysis of Mass-Throughput Data: A Technical Commentary from the Perspective of Statistical Machine Learning. *Cancer Informatics*, 2007 Feb 16; 2:133-62.
25. Aphinyanaphongs Y, **Statnikov A**, Aliferis CF. A Comparison of Citation Metrics to Machine Learning Filters for the Identification of High Quality MEDLINE Documents. *Journal of the American Medical Informatics Association*. 2006 Jul-Aug; 13: 446-455.
26. **Statnikov A**, Hardin D, Aliferis CF. Using SVM Weight-Based Methods to Identify Causally Relevant and Non-Causally Relevant Variables. *Proceedings of the Neural Information Processing Systems (NIPS) 2006 Workshop on Causality and Feature Selection*, 2006.
27. Tsamardinos I, **Statnikov A**, Brown LE, Aliferis CF. Generating Realistic Large Bayesian Networks by Tiling. *Proceedings of the 19<sup>th</sup> International Florida Artificial Intelligence Research Society (FLAIRS) Conference*, 2006.
28. Levy S, **Statnikov A**, Aliferis CF. Biomarker Selection from High-Dimensionality Data. *Pharmaceutical Discovery*. 2005 Microarray Supplement, 2005 Sep; 37-44.
29. **Statnikov A**, Tsamardinos I, Dosbayev Y, Aliferis CF. GEMS: A System for Automated Cancer Diagnosis and Biomarker Discovery from Microarray Gene Expression Data. *International Journal of Medical Informatics*. 2005 Aug; 74(7-8): 493-501.

30. **Statnikov A**, Aliferis CF, Tsamardinos I, Hardin D, Levy S. A Comprehensive Evaluation of Multicategory Classification Methods for Microarray Gene Expression Cancer Diagnosis. *Bioinformatics*. 2005 Mar; 21(5): 631-43.
31. Aphinyanaphongs Y, Tsamardinos I, **Statnikov A**, Hardin D, Aliferis CF. Text Categorization Models for High-Quality Article Retrieval in Internal Medicine. *Journal of the American Medical Informatics Association*. 2005 Mar-Apr; 12(2): 207-16.
32. Duda S, Aliferis CF, Miller R, **Statnikov A**, Johnson K. Extracting Drug-Drug Interaction Articles from MEDLINE to Improve the Content of Drug Databases. *Proceedings of the AMIA Annual Symposium*, 2005.
33. **Statnikov A**, Aliferis CF, Tsamardinos I. Methods for Multi-Category Cancer Diagnosis from Gene Expression Data: A Comprehensive Evaluation to Inform Decision Support System Development. *Studies in Health Technology and Informatics*. 2004; 107(Pt 2): 813-817.
34. Aliferis CF, Tsamardinos I, **Statnikov A**. HITON: A Novel Markov Blanket Algorithm for Optimal Variable Selection. *Proceedings of the AMIA Annual Symposium*, 2003.
35. Tsamardinos I, Aliferis CF, **Statnikov A**. Time and Sample Efficient Discovery of Markov Blankets and Direct Causal Relations. *Proceedings of the 9<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2003.
36. Aliferis CF, Tsamardinos I, Massion P, **Statnikov A**, Fananapazir N, Hardin D. Machine Learning Models for Classification of Lung Cancer and Selection of Genomic Markers Using Array Gene Expression Data. *Proceedings of the 16<sup>th</sup> International Florida Artificial Intelligence Research Society (FLAIRS) Conference*, 2003.
37. Tsamardinos I, Aliferis CF, **Statnikov A**. Algorithms for Large Scale Markov Blanket Discovery. *Proceedings of the 16<sup>th</sup> International Florida Artificial Intelligence Research Society (FLAIRS) Conference*, 2003.
38. Frey L, Fisher D, Tsamardinos I, Aliferis CF, **Statnikov A**. Identifying Markov Blankets with Decision Tree Induction. *Proceedings of the third IEEE International Conference on Data Mining (ICDM)*, 2003.
39. Aliferis CF, Tsamardinos I, **Statnikov A**, Brown LE. Causal Explorer: A Probabilistic Network Learning Toolkit for Biomedical Discovery. *Proceedings of the International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences (METMBS)*, 2003.
40. Aliferis CF, Tsamardinos I, Massion P, **Statnikov A**, Hardin D. Why Classification Models Using Array Gene Expression Data Perform So Well: A Preliminary Investigation of Explanatory Factors. *Proceedings of the International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences (METMBS)*, 2003.

#### **Books:**

41. **Statnikov A**, Aliferis CF, Hardin DP, Guyon I. A Gentle Introduction to Support Vector Machines in Biomedicine, Volume 1: Theory and Methods. *World Scientific Publishing*, Singapore: 2011.
42. **Statnikov A**, Aliferis CF, Hardin DP, Guyon I. A Gentle Introduction to Support Vector Machines in Biomedicine, Volume 2: Applications. (in preparation) *World Scientific Publishing*, Singapore: 2011.
43. Guyon I, Aliferis CF, Cooper GF, Elisseff A, Pellet JP, Spirtes P, **Statnikov A** (editors). Challenges in Machine Learning, Volume 2: Causation and Prediction Challenge. *Microtome Publishing*, Brookline, Massachusetts, 2010.
44. Guyon I, Dror G, **Statnikov A**, Vincent L (editors). Challenges in Machine Learning, Volume 6: Active Learning and Experimental Design. (in preparation) *Microtome Publishing*, Brookline, Massachusetts, 2011.

**Book Chapters:**

45. Guyon I, Aliferis CF, Cooper GF, Elisseeff A, Pellet JP, Spirtes P, **Statnikov A**. Causality Workbench. In *Causality in the Sciences*. Edited by Illari PM, Russo F and Williamson J. (in press) *Oxford University Press*, 2011.
46. **Statnikov A**, Tsamardinos I, Brown LE, Aliferis CF. Causal Explorer: A Matlab Library of Algorithms for Causal Discovery and Variable Selection for Classification. In *Challenges in Machine Learning. Volume 2: Causation and Prediction Challenge*. Edited by Guyon I, Aliferis CF, Cooper GF, Elisseeff A, Pellet JP, Spirtes P and Statnikov A. *Microtome Publishing*, Brookline, Massachusetts, 2010.

**Peer-Reviewed Software Presentations Published in Conference Proceedings:**

47. **Statnikov A**, Tsamardinos I, Aliferis CF. Using the GEMS System for Supervised Analysis of Cancer Microarray Gene Expression Data. *AMIA Annual Symposium*, 2005.
48. **Statnikov A**, Tsamardinos I, Aliferis CF. Using the GEMS System for Cancer Diagnosis and Biomarker Discovery from Microarray Gene Expression Data. *12<sup>th</sup> National Conference on Artificial Intelligence (AAAI)*, 2005.
49. **Statnikov A**, Tsamardinos I, Aliferis CF. Using GEMS for Cancer Diagnosis and Biomarker Discovery from Microarray Gene Expression Data. *13<sup>th</sup> Annual International Conference on Intelligent Systems for Molecular Biology (ISMB)*, 2005.

**Posters and Abstracts:**

50. Attur M, **Statnikov A**, Aliferis CF, Li Z, Krasnokutsky S, Samuels J, Greenberg JD, Patel J, Oh C, Abramson SB. Elevated Expression of Inflammatory Mediators Cyclooxygenase-2, Its Product Prostaglandin E2 and Interleukin-1 Beta by Peripheral Blood Leukocytes in Symptomatic Knee Osteoarthritis. *American College of Rheumatology and Association of Rheumatology Health Professionals (ACR/ARHP) Annual Scientific Meeting*, 2011.
51. Apanasovich T, **Statnikov A**, Alekseyenko AV, Lobach I. Role of Age When First Got Drunk in Gene-Environment Interplay in Alcoholism. *The Genetics and Epigenetics of Substance Abuse: NIDA/NIAAA Symposium at the World Congress of Psychiatric Genetics*, 2011.
52. Shmelkov E, Grigoryan A, **Statnikov A**, Cardozo T. In Silico Prediction of the Neutralization Range of Human Anti-HIV Monoclonal Antibodies. *ACM International Conference on Bioinformatics and Computational Biology (ACM-BCB), Second Immunoinformatics and Computational Immunology Workshop (ICIW 2011)*, 2011.
53. Zimmerman LJ, Slebos R, Coleman JA, Aliferis CF, Hardin DP, **Statnikov A**, Liebler DC. Assessment of Plasma Proteome Stability. *Clinical Proteomic Technologies for Cancer 2009 Annual Meeting*, 2009.
54. Zimmerman LJ, Coleman JA, Hardin DP, **Statnikov A**, Aliferis CF, Liebler DC. Development of Metrics for Assessment of Plasma Quality. *57th ASMS Conference on Mass Spectrometry and Allied Topics*, 2009.
55. Kokkotou E, Lois A, Triggs C, Conboy L, McDougall L, **Statnikov A**, Aliferis CF, Pothoulakis C, Kaptchuk T, Lembo A. Serum Biomarker Analysis of Placebo Responses in Patients with Irritable Bowel Syndrome. *Neurogastroenterology and Motility Joint International Meeting*, 2008.
56. Aliferis CF, **Statnikov A**, Tsamardinos I, Kokkotou E, Massion PP. Application and Comparative Evaluation of Causal and Non-Causal Feature Selection Algorithms for Biomarker Discovery in High-Throughput Biomedical Datasets. *Neural Information Processing Systems (NIPS) 2006 Workshop on Causality and Feature Selection*, 2006.

57. Aliferis CF, **Statnikov A**, Massion PP. Pathway Induction and High-Fidelity Simulation for Molecular Signature and Biomarker Discovery in Lung Cancer Using Microarray Gene Expression Data. *APS Conference: Physiological Genomics and Proteomics of Lung Disease*, 2006.
58. **Statnikov A**, Tsamardinos I, Aliferis CF. Using GEMS for Cancer Diagnosis and Biomarker Discovery from Microarray Gene Expression Data. *13<sup>th</sup> Annual International Conference on Intelligent Systems for Molecular Biology (ISMB)*, 2005.

#### **Selected Technical Reports:**

59. **Statnikov A**, Kasparova E, Aliferis CF. Applying Decision Support Models in the Presence of Incomplete Evidence. Technical Report DSL TR-06-02, *Department of Biomedical Informatics, Vanderbilt University*, 2006.
60. **Statnikov A**, Tsamardinos I, Aliferis CF. New Efficient and Correct Algorithms for Identification of Direct Causal Relationships and Markov Blankets from Data. Technical Report DSL TR-06-01, *Department of Biomedical Informatics, Vanderbilt University*, 2006.
61. Tsamardinos I, Aliferis CF, **Statnikov A**, Brown LE. Scaling-Up Bayesian Network Learning to Thousands of Variables Using Local Learning Technique. Technical Report DSL TR-03-02, *Department of Biomedical Informatics, Vanderbilt University*, 2003.
62. **Statnikov A**, Tsamardinos I, Aliferis CF. An Algorithm for Generation of Large Bayesian Networks. Technical Report DSL TR-03-01, *Department of Biomedical Informatics, Vanderbilt University*, 2003.
63. Aliferis CF, Tsamardinos I, **Statnikov A**. Large-Scale Feature Selection Using Markov Blanket Induction for the Prediction of Protein-Drug Binding. Technical Report DSL TR-02-06, *Department of Biomedical Informatics, Vanderbilt University*, 2002.

#### **Theses:**

64. **Statnikov A**. Algorithms for Discovery of Multiple Markov Boundaries: Application to the Molecular Signature Multiplicity Problem. Ph.D. Thesis. *Department of Biomedical Informatics, Vanderbilt University*, Advisor: Dr. Constantin F. Aliferis, Committee Members: Dr. Gregory F. Cooper, Dr. Douglas P. Hardin, Dr. Daniel R. Masys, Dr. Ioannis Tsamardinos, December 2008.
65. **Statnikov A**. Automatic Cancer Diagnostic Decision Support System for Gene Expression Domain. Master's Thesis. *Department of Biomedical Informatics, Vanderbilt University*, Advisors: Dr. Constantin F. Aliferis and Dr. Ioannis Tsamardinos, August 2005.
66. **Statnikov A**. Numerical Methods for Image Reconstruction for the Calibration of the NASA-Glenn Icing Research Wind Tunnel: A Computer-Based Approach. Master's Thesis. *Department of Mathematics, Case Western Reserve University*, Advisor: Dr. Steven H. Izen, August 2002.

#### **Patents:**

67. **Statnikov A**, Aliferis CF. Computer Implemented Method for Discovery of Markov Boundaries from Datasets with Hidden Variables. *United States Patent Application*, #12/689,944, Filed: 2010.
68. Aliferis CF, **Statnikov A**. Local Causal and Markov Blanket Induction Method for Causal Discovery and Feature Selection from Data. *United States Patent Application*, #12/700,689, Filed: 2010.
69. **Statnikov A**, Aliferis CF. A Computer Implemented Method for Determining All Markov Boundaries and Its Application for Discovering Multiple Maximally Accurate and Non-Redundant Predictive Models. *United States Patent Application*, #12/610,046, Filed: Oct 30, 2009.
70. **Statnikov A**, Aliferis CF, Tsamardinos I, Fananapazir N. Method and System for Automated Supervised Data Analysis. *United States Patent*, #7912698, Filed: Aug 28, 2006, Issued: Mar 22, 2011.

**Papers in Preparation or under Review:**

- **Statnikov A**, Lytkin NI, Lemeire J, Aliferis CF. Algorithms for Discovery of Multiple Markov Boundaries. (submitted), 2011.
- Feig JE, Vengrenyuk Y, Wu C, Reiser V, Puig O, **Statnikov A**, Aliferis CF, Garabedian MJ, Fisher EA. Regression of Atherosclerosis is Characterized by the Molecular Re-Programming of the Macrophage Transcriptome. (in preparation), 2011.
- Mani S, Cooper GF, **Statnikov A**. Causal Discovery Algorithms Based on Y Structures. (in preparation), 2011.
- Bai J, Alekseyenko AV, **Statnikov A**, Wang IM, Wong P. Strategic Application of Gene Expression in Drug Development and Clinical Applications. (in preparation), 2011.
- Marbach D, Costello JC, Küffner R, Vega N, Prill RJ, Camacho DM, Allison KR, the DREAM5 Consortium, Kellis M, Collins JJ, Stolovitzky G. Wisdom of Crowds for Robust Gene Network Inference. (submitted), 2011.

(The list of publications in the field of *Multicriteria Analysis* is available upon request)

**Invited Talks**

1. Assessing quality and completeness of human transcriptional regulatory pathways on a genome-wide scale. *New York University Langone Medical Center*, April 2011.
2. A comprehensive assessment of methods for de-novo reverse-engineering of genome-scale regulatory networks. *New York University Langone Medical Center*, November 2010.
3. Computational advances in causal discovery: Inferring regulatory networks and biological pathways. *U.S. Food and Drug Administration, Office of Clinical Pharmacology*, May 2010.
4. Computational advances in causal discovery: Inferring regulatory networks and biological pathways. *Glycomics Workshop, New York University Langone Medical Center*, March 2010.
5. Analysis and computational dissection of molecular signature multiplicity. *New York University Langone Medical Center*, December 2008.
6. Analysis and computational dissection of molecular signature multiplicity. *Department of Biomedical Informatics, University of Utah*, November 2008.
7. Analysis and computational dissection of molecular signature multiplicity. *Department of Biomedical Informatics, Columbia University*, November 2008.
8. Effects of environment, genetics and data analysis in an esophageal cancer genome-wide association study. *Department of Biostatistics, Vanderbilt University*, October 2007.
9. Are random forests better than support vector machines for microarray-based cancer classification? *AMIA Annual Symposium*, November 2007.
10. Using SVM weight-based methods to identify causally relevant and non-causally relevant variables. *Neural Information Processing Systems (NIPS) Workshop on Causality and Feature Selection*, December 2006.
11. Using the GEMS system for supervised analysis of cancer microarray gene expression data. *AMIA Annual Symposium*, October 2005.
12. Using the GEMS system for cancer diagnosis and biomarker discovery from microarray gene expression data. *12<sup>th</sup> National Conference on Artificial Intelligence (AAAI)*, July 2005.
13. Using GEMS for cancer diagnosis and biomarker discovery from microarray gene expression data. *13<sup>th</sup> Annual International Conference on Intelligent Systems for Molecular Biology (ISMB)*, June 2005.
14. Methods for multi-category cancer diagnosis from gene expression data: A comprehensive evaluation to inform decision support system development. *MEDINFO*, September 2004.
15. Toward a system to support automated development and evaluation of cancer diagnostic models from gene expression data. *Department of Biomedical Informatics, Vanderbilt University*, November 2003.

## Educational Activities

### Students:

- Varun Narendra (M.D./Ph.D. student): Rotation Advisor
- Evgeny Shmelkov (Ph.D. student): Rotation Advisor, Chair of Ph.D. Committee
- Aaron Elkin (Undergraduate student): Summer Internship Advisor
- Thang Le (Ph.D. student in Computer Science, Rutgers University): Ph.D. Committee Member
- Kassatihun Gebre-Amlak (M.D. student, Vanderbilt University): Emphasis Research Program Mentor.

### Course Instruction:

- Instructor of the “*Introduction of Translational and Clinical Bioinformatics*” course for training pathology residents, Department of Pathology, *New York University School of Medicine*: March - April 2011.
- Co-instructor of the “*Introduction of Translational and Clinical Bioinformatics*” course for training pathology residents, Department of Pathology, *New York University School of Medicine*: September - November 2009.
- Co-instructor of the “*Biomedical Artificial Intelligence*” course, Department of Biomedical Informatics, *Vanderbilt University*: January - February 2009.
- Instructor of undergraduate Mathematics courses “*Elementary Functions and Analytic Geometry*” and “*Calculus I*”, Department of Mathematics, *Case Western Reserve University*: August 2001 - May 2002.

### Course Lectures:

- Bioinformatics: Molecular Signature Discovery in the “*Introduction to Biomedical Informatics*” course, CTSI Translational Research Education and Careers (TREC) Program, *New York University School of Medicine*: November 2011.
- Development of Molecular Signatures from High-Throughput Assay Data in the “*Molecular Signaling and Drug Development*” course, Computational Biology Graduate Training Program, *New York University School of Medicine*: May 2011.
- Bioinformatics: Molecular Signature Discovery in the “*Introduction to Biomedical Informatics*” course, CTSI Translational Research Education and Careers (TREC) Program, *New York University School of Medicine*: December 2010.
- Development of Molecular Signatures from High-Throughput Assay Data in the “*Molecular Signaling and Drug Development*” course, Computational Biology Graduate Training Program, *New York University School of Medicine*: April 2010.
- Applications of Support Vector Machines in Clinical Bioinformatics in the “*Biomedical Artificial Intelligence*” course, Department of Biomedical Informatics, *Vanderbilt University*: March 2008

### Tutorials:

- “*A Gentle Introduction to Support Vector Machines in Biomedicine*”, *AMIA Annual Symposium*, 2010
- “*A Gentle Introduction to Support Vector Machines in Biomedicine*”, *AMIA Annual Symposium*, 2009
- “*Support Vector Machines without Tears*”, *AMIA Annual Symposium*, 2008

## Grants

### Active:

1 I01 CX000242-01	Department of Veterans Affairs	10/01/09-09/30/13	\$188,000
PI: Massion	Role: Investigator	Effort: 2.4 calendar	
<i>Molecular Biomarkers of Small Cell Lung Cancer Behaviour</i>			

1 U54 RR024386-01A2	NIH/NCRR	07/14/09-03/31/14	\$3,879,473
PI: Cronstein	Role: Bioinformatics Service Manager	Effort: 2.4 calendar	
<i>NYU-HHC Clinical and Translation Science Award</i>			

2 U01 CA111295-05	NIH/NCI	09/30/04-06/20/15	\$608,087
PI: Pass	Role: Investigator	Effort: 0.3 calendar	
<i>North American Mesothelioma Early Detection Research Network Grant</i>			

### Pending:

1 R01 LM 011179-01A1	NIH/NLM	07/01/12-06/30/16	\$150,000
PI: Statnikov	Role: Principal Investigator	Effort: 3.0 calendar	
<i>Methods for Accurate and Efficient Discovery of Local Pathways</i>			

1 R21 AA020356-01A1	NIH/NIAAA	12/01/11-11/30/13	\$150,000
PI: Lobach	Role: Investigator	Effort: 0.6 calendar	
<i>Novel Analysis of Gene-Gene and Gene-Environment Interactions in Alcoholism</i>			

1 R01 AG042427-01	NIH/NIA	07/01/12-06/30/16	\$250,000
PI: Lobach	Role: Investigator	Effort: 1.2 calendar	
<i>Analysis of Gene-Environment Interplay in Alzheimer's Disease</i>			

R01 AR052873	NIH/NIAMS	07/01/12-06/30/17	\$495,000
PI: Abramson	Role: Investigator	Effort: 1.2 calendar	
<i>Leukocyte Gene Expression and Genetic Biomarkers of OA Incidence and Progression</i>			

1 P50 AG042425-01	NIH/NIA	07/01/12-06/30/17	\$750,000
PI: Naftolin	Role: Investigator	Effort: 1.2 calendar	
<i>Specialized Centers of Research (SCOR) of Sex Differences</i>			

1 R01 MH098097-01	NIH/NIMH	07/01/12-06/30/15	\$175,000
PI: Malaspina	Role: Investigator	Effort: 0.6 calendar in year 3	
<i>Willingness to Work for Goals and Avolition</i>			

### Completed:

3 R01 AR0566672-01S1	NIH/NIAMS	09/23/09-09/22/11	\$236,187
PI: Cronstein	Role: Investigator	Effort: 1.2 calendar	
<i>The Pharmacology of Dermal Fibrosis</i>			

## Scientific Leadership and Community Service

### Forums:

- AMIA (American Medical Informatics Association) 2012 Summit on Translational Bioinformatics: Scientific Program Committee Member
- ICIBM 2012 (International Conference on Intelligent Biology and Medicine): Program Committee Member
- AMIA (American Medical Informatics Association) 2011 Fall Symposium: Poster Awards Committee Member
- AMIA (American Medical Informatics Association) 2010 Fall Symposium: Scientific Program Committee Member, Poster Awards Committee Member
- AISTATS 2010 (Artificial Intelligence in Statistics) Active Learning and Experimental Design Workshop: Organizing Committee Member
- NIPS 2008 (Neural Information Processing Systems) Causality Workshop: Organizing Committee Member

### Challenges:

- Scientific Advisor of *ChaLearn.org* (Challenges in Machine Learning): 2011 - Present
- Co-Organizer of the “*Active Learning Challenge*”: 2010
- Co-Organizer of the “*Causality Challenge #2: Potluck*”: 2008
- Co-Organizer of the “*Causality Challenge #1: Causation and Prediction*”: 2007 - 2008

### Other:

- European Commission: Future and Emerging Technologies, Ad-hoc Grant Reviewer: 2010
- AMIA (American Medical Informatics Association) Genomics Workgroup Student Representative: 2005 - 2007

## Editorial Responsibilities for Scientific Publications

- Editorial Board Member, *Journal of Machine Learning Research*: 2011 - Present
- Editorial Board Member, *Journal of Microbial and Biochemical Technology*: 2011 - Present
- Associated Editor, *Journal of Integrated OMICS*: 2011 - Present

## Scientific Reviewer Activities

### Ad-hoc reviewer for the following journals:

- Bioinformatics
- BMC Bioinformatics
- BMC Medical Genomics
- British Journal of Cancer
- Computational Statistics and Data Analysis
- Decision Support Systems
- FEBS Letters
- IEEE Transactions on Knowledge and Data Engineering
- Inflammation
- International Journal of Approximate Reasoning
- Journal of the American Medical Informatics Association
- Journal of Artificial Intelligence Research
- Journal of Biomedical Informatics
- Journal of Clinical and Laboratory Medicine
- Journal of Machine Learning Research
- Methods of Information in Medicine
- Nucleic Acids Research
- Pattern Recognition Letters
- PLoS Genetics

### Reviewer for the following conferences:

- 11<sup>th</sup> IEEE International Conference on Data Mining series (ICDM 2011)
- American Medical Informatics Association (AMIA) 2011 Fall Symposium
- 22<sup>nd</sup> International Joint Conference on Artificial Intelligence (IJCAI 2011)
- 19<sup>th</sup> European Symposium on Artificial Neural Networks (ESANN 2011)
- American Medical Informatics Association (AMIA) 2011 Summit on Translational Bioinformatics
- 2010 International Joint Conference on Neural Networks (IJCNN)
- American Medical Informatics Association (AMIA) 2010 Fall Symposium
- American Medical Informatics Association (AMIA) 2009 Fall Symposium
- 15<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2009)
- Pacific Symposium on Biocomputing (PSB) 2009
- 22<sup>nd</sup> Annual Conference on Neural Information Processing Systems (NIPS 2008)
- American Medical Informatics Association (AMIA) 2008 Fall Symposium
- European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2008)
- IEEE World Congress on Computational Intelligence (WCCI 2008)
- American Medical Informatics Association (AMIA) 2007 Fall Symposium
- 3<sup>rd</sup> Indian International Conference on Artificial Intelligence (IICAI 2007)
- American Medical Informatics Association (AMIA) 2006 Fall Symposium
- 10<sup>th</sup> Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2006)
- 11<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2005)

### Reviewer for the following books:

- “*Computational Intelligence in Bioinformatics*” (Editor in Chief: Arpad Kelemen)

## Honors and Awards

- ISMB 2005 Best Poster Award (June 2005)
- Medinfo 2004 Student Paper Competition, Gold Medal (September 2004)
- Vanderbilt University Medical School Dean's Scholarship (September 2004 - December 2008)
- American Mathematical Society Waldemar J. Trjitzinsky National Award (May 2000)
- Case Western Reserve University Dean's Honors List (2000)

## Professional and Scientific Societies

- AAAI (American Association for Artificial Intelligence): 2005 - Present
- AAAS (American Association for the Advancement of Science): 2007 - Present
- ACM (Association for Computing Machinery): 2008 - Present
- AMIA (American Medical Informatics Association): 2003 - Present
- AMS (American Mathematical Society): 2001 - Present
- ASA (American Statistical Association): 2008 - Present
- ISCB (International Society for Computational Biology): 2004 - Present

## Other Experience

- **Teaching assistant**, Department of Mathematics, *Case Western Reserve University*: January 2000 - May 2001
- **Webmaster**, Department of Mathematics, *Case Western Reserve University*: January 2000 - May 2002
- **Webmaster**, Department of Mechanical Engineering, *Cleveland State University*: December 2000 - May 2001

## Personal

- Born in Moscow, Russia, 1979
- United States citizen
- Fluent in English and Russian