

Curriculum Vitae
Constantin F. Aliferis
(Web page: www.dsl-lab.org)

Education and Training

Undergraduate/Graduate	Athens University	M.D., 1990	Medicine
Graduate:	University of Pittsburgh	M.S., 1994	Intelligent Systems
	University of Pittsburgh	Ph.D., 1998	Intelligent Systems/ Medical Informatics

Postgraduate Fellowship Training (Biomedical Informatics):

1991-1994: University of Pittsburgh, Section of Medical Informatics, School of Medicine

Academic appointments:

June 1998-May 2000	Athens University, Greece	Research Associate in Epidemiology
June 2000-present	Vanderbilt University	Assistant Professor of Biomedical Informatics
June 2000-present	Vanderbilt-Ingram Cancer Center	Investigator
June 2000-Sept. 2001	Vanderbilt University	Director, MS/PhD Program in Biomedical Informatics
Sept 2001-present	Vanderbilt University	Director, Discovery Systems Laboratory
May 2005-present	Vanderbilt University	Assistant Professor of Cancer Biology (secondary appointment)
May 2007-present	Vanderbilt University	Assistant Professor of Computer Science (secondary appointment)
May 2007-present	Vanderbilt University	Assistant Professor of Biostatistics (secondary appointment)

Honors

2007-8	Co-organizer of Causal Discovery Challenge PASCAL (Pattern Analysis, Statistical Modeling and Computational Learning) network of the European Commission's IST-funded Network of Excellence for Multimodal Interfaces.
2007	Guest Editor Journal of Machine Learning Research, Special Topic on Causality
2007	Editorial Board member: Open Artificial Intelligence Journal
2007	Scientific Program Committee: American Medical Informatics Fall Symposium
2006	Co-organizer: NIPS 2006 workshop on causality and feature selection
2006	Invited talk in American Physiological Society Conference on Physiological Genomics and Proteomics of Lung Disease
2006	Scientific Program Committee & Session Chair: American Medical Informatics Fall Symposium
2005	Research spotlight: Cancer Informatics
2005	Editorial Board member: Cancer Informatics
2005	Editorial Board member: Biomedical Knowledge
2005	Best Poster Award in Intelligent Systems in Molecular Biology
2005	Contributing author to "Tietz Textbook of Clinical Chemistry and Molecular Diagnostics"

2004, 5, 6	Scientific Program Committee: International Workshop on Feature Selection for Data Mining
2002- 2006	American Medical Informatics Association Education Committee
2002- 2003	American Medical Informatics Association Professional Relations Committee
2002	Finalist, Best paper award competition, AMIA Fall Symposium
2000, 7	Session Chair, American Medical Informatics Association Fall Meeting
1997	Invited editorial: Yearbook of Medical Informatics
1996	Selection of publication in the Yearbook of Medical Informatics
1995	Student Paper Award from American College of Medical Informatics (ACMI) in World Congress of Medical Informatics MEDINFO 1995.
1994	Student Paper Award in the 18 th Annual AMIA Fall Symposium
1993	Finalist, Best paper award competition, AMIA Fall Symposium

Membership on Study Sections/Review Panels

2005	NIH/NLM RO1/R21/R13 Review Panel
2005	DOD Epidemiology, Breast Cancer review panel
2004	NIH/NLM Grants Review Panel for Exploratory/Developmental Research grants in Biomedical Informatics and Informatics for Disaster Management grants
2001	NIH Cardiovascular Sciences Initial Review Group ZRG1-CVA-01, Bioengineering Research Partnership Review

Publications

A. Peer reviewed articles¹

1. “The Problem of Statistical Gene Instability in Microarray Studies: External Reproducibility and Biological Importance of Unstable Genes and their Molecular Signatures” **C.F. Aliferis**, A. Statnikov, S. Prata, E. Kokkotou (in preparation).
2. “Extending Text Categorization Filters in Medicine” Y. Aphinyanaphongs, N. Wilczynski, **C. Aliferis**, B. Haines (in preparation).
3. “Dissecting the role of environment, genetics and data analysis bias in genome-wide association studies: a case study in esophageal cancer”. A. Statnikov, C. Li, **C.F. Aliferis**. (submitted).
4. “Gene Expression Microarrays Do Predict Clinical Outcomes”. **C.F. Aliferis**, A. Statnikov, I. Tsamardinos., J. Schildcrout, B. Shepherd, F. Harrell Jr (submitted).
5. “Local Regulatory-Network Inducing Algorithms For Biomarker Discovery From Mass-Throughput Datasets”. **C.F. Aliferis**, A. Statnikov, E. Kokkotou, P.P. Massion, I. Tsamardinos (submitted).
6. “Are Random Forests Better than Support Vector Machines for Microarray-Based Cancer Classification?” A. Statnikov, **C.F. Aliferis** (submitted).
7. “Genomics and Proteomics of Lung Disease: Conference Summary” J.U. Raj, **C. Aliferis**, R.M. Caprioli, A.W. Cowley, Jr., P.F. Davies, M.W. Duncan, D.J. Erle, S.C. Erzurum, P.W. Finn, H. Ischiropoulos, N. Kaminski, S.R. Kleeberger, G.D. Leikauf, J.E. Loyd, T.R. Martin, S. Matalon, J. H. Moore, J. Quackenbush, T. Sabo-Attwood, S.D. Shapiro, J.E. Schnitzer, D.A. Schwartz, L.M.

¹ Following the practice in computer science and biomedical informatics, this section includes *full-length* papers published in peer-reviewed journals and *full-length* papers in major, rigorously peer-reviewed biomedical informatics and computer science conference proceedings (acceptance rates of computer science conferences in present vitae range from 25% to less than 5%).

- Schwiebert, D. Sheppard, L.B. Ware, S.T. Weiss, J.A. Whitsett, M.M. Wurfel, M.A. Matthay. (to appear: Lung Cellular and Molecular Physiology).
8. "A Causal Modeling Framework for Generating Clinical Practice Guidelines from Data". – S. Mani and **C. Aliferis**. "AI in medicine Europe (AIME 2007) conference Amsterdam, July 2007.
 9. "Text Categorization Models for Identifying Unproven Cancer Treatments on the Web" Y. Aphinyanaphongs, **C.F. Aliferis**. To appear in International Medical Informatics Congress, MEDINFO 2007.
 10. "Learning Causal and Predictive Clinical Practice Guidelines from Data". S. Mani, **C. F. Aliferis**, S. Krishnaswami, T. Kotchen. To appear in International Medical Informatics Congress, MEDINFO 2007.
 11. "A comparison of Impact Factor, Clinical Query Filters, and Pattern Recognition Query Filters in Terms of Sensitivity to Topic". L. Fu, L. Wang, Y. Aphinyanaphongs, **C. F. Aliferis**. To appear in International Medical Informatics Congress, MEDINFO 2007.
 12. "Using SVM Weight-Based Methods to Identify Causally Relevant and Non-Causally Relevant Variables". Statnikov A., Hardin D., Aliferis CF. NIPS 2006 Workshop on Feature Selection and Causality.
 13. "Challenges in the Analysis of Mass-Throughput Data: A Technical Commentary from the Statistical Machine Learning Perspective". **Aliferis CF**, Statnikov A, Tsamardinos I. *Cancer Informatics*. 2006; 2: 133–162.
 14. "Prospective validation of text categorization models for indentifying high-quality content-specific articles in PubMed". Y. Aphinyanaphongs, **C.F. Aliferis**. Proceeding of 2006 Annual Fall Conference of the American Medical Informatics Association.
 15. "A Comparison of Citation Metrics to Machine Learning Filters for the Identification of High Quality MEDLINE Documents". Y. Aphinyanaphongs, A. Statnikov, **C.F. Aliferis**. *J Am Med Inform Assoc*. 2006 Jul-Aug;13(4):446-55.
 16. "The Max-Min Hill Climbing Bayesian Network Structure Learning Algorithm". I. Tsamardinos, L.E. Brown, **C.F. Aliferis**. *Machine Learning*. 2006 Mar 29; [Epub ahead of print].
 17. "An Algorithm for Generation of Large Bayesian Networks". I. Tsamardinos, L. E. Brown, A. Statnikov, **C. F. Aliferis**. In Proceedings of the 19th International Florida Artificial Intelligence Research Society (FLAIRS) Conference 2006
 18. "Modeling Clinical Judgment and Implicit Guideline Compliance in the Diagnosis of Melanomas Using Machine Learning". A. Sboner, **C.F. Aliferis**. AMIA Symposium, 2005
 19. "Using citation data to improve retrieval from MEDLINE". E.V.Bernstam, J.R.Herskovic, Y. Aphinyanaphongs, **C.F. Aliferis**, M.G.Sriram, W.R. Hersh,. *J Am Med Inform Assoc*. 2006 Jan-Feb;13(1):96-105
 20. "Extracting Drug-Drug Interaction Articles from MEDLINE to Improve the Content of Drug Databases". S. Duda, **C.F. Aliferis**, R.A. Miller, A. Statnikov, K.B.Johnson, , AMIA Symposium, 2005
 21. "Formative Evaluation of a Prototype System for Automated Analysis of Mass Spectrometry Data". N. Fananapazir, M. Li, D. Spentzos, **C.F. Aliferis**. AMIA Symposium, 2005
 22. "A Comparison of Novel and State-of-the-Art Polynomial Bayesian Network Learning Algorithms" Laura E. Brown, Ioannis Tsamardinos, **C. F. Aliferis** (in: *American Association for Artificial Intelligence (AAAI) conference 2005*).

23. "Predicting Dire Outcomes of Patients with Community Acquired Pneumonia" G.F. Cooper, V. Abraham , **C.F. Aliferis**, J. Aronis, B.G. Buchanan , R.Caruana, M. J. Fine, J.E. Janosky, G. Livingston, S. Monti, T. Mitchell, P. Spirtes. *J Biomed Inform.* 2005 Oct;38(5):347-66
24. "Gene Expression Model Selector (GEMS): a system for decision support and discovery from array gene expression data". A. Statnikov, I. Tsamardinos, Y. Dosbayev, **C.F. Aliferis**. *Int J Med Inform.* 2005 Aug;74(7-8):491-503.
25. "Text Categorization Models for Retrieval of High Quality Articles in Internal Medicine". Y. Aphinyanaphongs, I. Tsamardinos, A. Statnikov, D. Hardin, **C.F. Aliferis**. *J Am Med Inform Assoc.* 2005 Mar-Apr;12(2):207-16.
26. "A Comprehensive Evaluation of Multicategory Classification Methods for Microarray Gene Expression Cancer Diagnosis". A. Statnikov, **C.F. Aliferis**, I. Tsamardinos, D. Hardin, S. Levy. *Bioinformatics.* 2005 Mar 1;21(5):631-43.
27. "Methods for Multi-Category Cancer Diagnosis from Gene Expression Data: A Comprehensive Evaluation to Inform Decision Support System Development". A. Statnikov, **C.F. Aliferis**, I. Tsamardinos. *In Proceedings of the 11th World Congress on Medical Informatics (MEDINFO)*, September 7-11, 2004, San Francisco, California, USA
28. "A Theoretical Characterization of Linear SVM-Based Feature Selection". D. Hardin, I. Tsamardinos, **C.F. Aliferis** , *In Twenty-First International Conference on Machine Learning (ICML)*, 2004
29. "Learning Boolean Queries for Article Quality Filtering". Y. Aphinyanaphongs, **C.F. Aliferis** . *In Proceedings of the 11th World Congress on Medical Informatics (MEDINFO)*, September 7-11, 2004, San Francisco, California, USA
30. "A Novel Algorithm for Scalable and Accurate Bayesian Network Learning". L.E. Brown, I. Tsamardinos, **C.F. Aliferis**. *In Proceedings of the 11th World Congress on Medical Informatics (MEDINFO)*, September 7-11, 2004, San Francisco, California, USA
31. "HITON, A Novel Markov Blanket Algorithm for Optimal Variable Selection". **C. F. Aliferis**, I. Tsamardinos, A. Statnikov. *In Proceedings of the 2003 American Medical Informatics Association (AMIA) Annual Symposium*, 2003, pages 21-25.
32. "Why Classification Models Using Array Gene Expression Data Perform So Well: A Preliminary Investigation Of Explanatory Factors". **C.F. Aliferis**, I. Tsamardinos, P. Massion, A. Statnikov, D. Hardin. *In Proceedings of the 2003 International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences (METMBS)*, June 23-26, 2003, Las Vegas, Nevada, USA, CSREA Press
33. "Machine Learning Models For Classification Of Lung Cancer and Selection of Genomic Markers Using Array Gene Expression Data". **C.F. Aliferis**, I. Tsamardinos, P. Massion, A. Statnikov, N. Fananapazir, D. Hardin. *In Proceedings of the 16th International Florida Artificial Intelligence Research Society (FLAIRS) Conference*, May 12-14, 2003, St. Augustine, Florida, USA, AAAI Press, pages 67-71
34. "Time and Sample Efficient Discovery of Markov Blankets and Direct Causal Relations". I. Tsamardinos, **C.F. Aliferis**, A. Statnikov. *In Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August 24-27, 2003, Washington, DC, USA, ACM Press, pages 673-678
35. "Algorithms for Large Scale Markov Blanket Discovery". I. Tsamardinos, **C.F. Aliferis**, A. Statnikov. *In Proceedings of the 16th International Florida Artificial Intelligence Research Society (FLAIRS) Conference*, May 12-14, 2003, St. Augustine, Florida, USA, AAAI Press, pages 376-380

36. "Towards Principled Feature Selection: Relevance, Filters, and Wrappers". . I. Tsamardinos and **C.F. Aliferis**. In *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics*, January 3-6, 2003, Key West, Florida, USA
37. "Causal Explorer: A Probabilistic Network Learning Toolkit for Biomedical Discovery". **C. Aliferis**, I. Tsamardinos, A. Statnikov, L.E. Brown. In *Proceedings of the 2003 International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences (METMBS)*, June 23-26, 2003, Las Vegas, Nevada, USA, CSREA Press
38. "Text Categorization Models for Retrieval of High Quality Articles in Internal Medicine." Y. Aphinyanaphongs, **C.F. Aliferis**. In *Proceedings of the 2003 American Medical Informatics Association (AMIA) Annual Symposium*, 2003, Washington, DC, USA, pages 31-35.
39. "Identifying Markov Blankets with Decision Tree Induction." L. Frey, D. Fisher, I. Tsamardinos, **C.F. Aliferis**, A. Statnikov. In *Proceedings of the Third IEEE International Conference on Data Mining (ICDM)*, November 19-22, 2003, Melbourne, Florida, USA, IEEE Computer Society Press, pages 59-66.
40. "Machine Learning Models For Lung Cancer Classification Using Array Comparative Genomic Hybridization". **C. F. Aliferis**, D. Hardin, P.Massion. *Proc AMIA Symp.* 2002; 7-11.
41. "Issues and Opportunities in Public Health Informatics: A Panel Discussion". R. Kukafka, P.W. O'Carrol, J.L. Gerberding, E.H. Shortliffe, **C. Aliferis**, J.R. Lumpkin, W.A. Yasnoff. *J Public Health Manag Pract.* 2001 Nov;7(6):31-42.
42. "Temporal Representation Design Principles: An Assessment in the Domain of Liver Transplantation". **C.F. Aliferis**, and G. F. Cooper. *Proc AMIA Symp.* 1998;170-4.
43. "Representing and Developing Temporally Abstracted Knowledge as a means Towards Facilitating Time Modeling in Medical Decision-Support Systems". **C.F. Aliferis**, G.F. Cooper, M.E. Pollack, B.G. Buchanan, M.M. Wagner. *Comput. Biol. Med.*, 27:411-434, 1997.
44. "An Evaluation of Machine Learning Methods for Predicting Pneumonia Mortality". G.F. Cooper, **C.F. Aliferis**, R. Ambrosino, J. Aronis, B.G. Buchanan, R. Caruana, M.J. Fine, C. Glymour, G. Gordon, B.H. Hanusa, J.E. Janosky, C. Meek, T. Mitchell, T. Richardson, and P. Spirtes. *Artificial Intelligence in Medicine*, 9, 107-138, 1997.
45. "A Temporal Analysis of QMR." Aliferis, **C. F., Cooper**, G. F., Miller, R. A., Buchanan, B. G., Bankowitz, R., and Giuse, N. *Journal of the American Medical Informatics Association*, 3, 79-91,1996.
46. "A Structurally and Temporally Extended Bayesian Belief Network Model: Definitions, Properties, and Modeling Techniques". **C.F. Aliferis**, G. Cooper. *Proc. Uncertainty in Artificial Intelligence*, 1996; 28-39.
47. "On the Heuristic Nature of Medical Decision-Support Systems." **C.F. Aliferis**, and R. Miller, *Methods of Information in Medicine*, 34: 5-14, 1995.
48. "Temporal Reasoning Abstractions in QMR. **C.F. Aliferis**, G.F. Cooper, B.G. Buchanan, R.A. Miller, R.Bankowitz, and N. Giuse. *Proc. MEDINFO.* 1995;8 Pt 1:847-51.
49. "A New Formalism for Temporal Modeling in Medical Decision-Support Systems". **C.F Aliferis**, G.F. Cooper. *Proc Annu Symp Comput Appl Med Care.* 1995; 213-7.
50. "An Evaluation of an Algorithm for Inductive Learning of Bayesian Belief Networks Using Simulated Data Sets". **C. F. Aliferis**, G. Cooper. *Proc. Uncertainty in Artificial Intelligence*, 1994; 8-14.

51. "A Temporal Analysis of QMR: Abstracted Temporal Representation and Reasoning and Initial Assessment of Diagnostic Performance Trade-Offs". **C.F. Aliferis**, G.F. Cooper, R. Bankowitz. *Proc Annu Symp Comput Appl Med Care*. 1994; 709-15.
52. "Data Explorer : A Prototype Expert System for Statistical Analysis". **C. F. Aliferis**, E. Chao, G. F. Cooper. *Proc Annu Symp Comput Appl Med Care*. 1993;:389-93.
53. "Incidence of Gynecomastia in 954 Young Males and its Relationships to Somatometric Parameters". E. Georgiadis, L. Papandreou, C. Evangelopoulou, **C.F. Aliferis**, C. Lymberis, C. Panitsa, M.Batrinou. *Ann Hum Biol*, 21; 579-87, 1994.
54. "Body Height, Weight, and Obesity in Young Greeks". E. Georgiadis, **C.F. Aliferis**, L. Papandreou, C. Mantzoros, D. Trichopoulos. *Iatriki* 63(3) 272-276, 1993.
55. "Effects of Socioeconomic Factors on the Weight and the Obesity of Young Greeks ". E. Georgiadis, L. Papandreou, C. Mantzoros, **C.F. Aliferis**, K. Konstantinou. *Iatriki* 63(6) 595-601,1993.
56. "Evidence for an Increase in the Prevalence of Known Diabetes in a Sample of an Urban population in Greece". N. Katsilambros, **K. Aliferis**, C. Darviri, P.Tsapogas, Z. Alexiou, N. Tritos, M. Arvanitis. *Diabet Med*: 10, 87-90 1993.
57. "Serum Lipids and Arterial Blood Pressure in Relation to Waist-To-Hip Ratio in Young Males" (letter). N. Katsilambros, E. Georgiadis, **C.F. Aliferis**, L. Papandreou, D. Triantaphyllou, S. Kouroutis, N. Grigoriadis, A. Tzavaras. *Am J Clin Nutr* 57; 697-8, 1993.
58. "Pattern of Treatment Among Diabetic Patients in a Sample of Urban Greek Population". N.Katsilambros, **K. Aliferis** et al. *Diabete & Metabolisme* 19, 130-132, 1993.
59. "Are Adrenal and Testicular Androgen Levels Correlated?" E. Georgiadis, C. Mantzoros, **C. F.Aliferis**, M. Batrinou. *Hormone and Metabolic Research* 10 (24) 488-91Oct. 1992.
60. "Effects of Socioeconomic Factors on the Height of Young Greeks". E. Georgiadis, L. Papandreou, C. Mantzoros, **C. F.Aliferis**. *Hippokrates* (4) 237-243, 1991.
61. "Computer-Assisted Growth Assessment": A Program for Monitoring Growth and Prognosing Final Height of Children'. **C. F.Aliferis**, E. Georgiou, K. Dalles, C. Proukakis. *Materia Medica Greca* 18(1) 62-68, 1990.

B. Peer-reviewed Abstracts & Peer-reviewed Software Demonstrations (Selected):

- 62 "Formative Comparative Evaluation of Traditional and Recent Quality-Content Filters for Answering Clinical Questions with MEDLINE." Aphinyanaphongs, Y, Jerome, R, **Aliferis C**. Medical Library Association Annual Meeting and Exhibition 2007, Philadelphia, PA.
- 63 "Application and Comparative Evaluation of Causal and Non-Causal Feature Selection Algorithms for Biomarker Discovery in High-Throughput Biomedical Datasets". **Aliferis CF**, Statnikov A., Tsamardinos I, Kokkotou E, Massion PP. NIPS 2006 Workshop on Feature Selection and Causality.
- 64 "Pathway induction and high-fidelity simulation for molecular signature and biomarker discovery in lung cancer using microarray gene expression data". **Aliferis CF**, Statnikov A, Massion P. In Proceedings of 2006 American Physiological Society Conference "Physiological Genomics and Proteomics of Lung Disease". November 2-5, 2006.

- 65 "Modelling liver transplant survival: Comparing techniques of deriving predictor sets". Hoot, N., Feurer, I., Pinson, C.W., **Aliferis, C.F.**, *Journal of Gastrointestinal Surgery*, Apr 2005
- 66 "GEMS: A System for Cancer Diagnosis and Biomarker Discovery from Microarray Gene Expression Data". Statnikov A, Tsamardinos I, **Aliferis CF**. *AMIA Annual Symposium*, 2005.
- 67 "Using the GEMS System for Cancer Diagnosis and Biomarker Discovery from Microarray Gene Expression Data". Statnikov A, Tsamardinos I, **Aliferis CF**. 12th National Conference on Artificial Intelligence (AAAI), 2005.
- 68 "Using GEMS for Cancer Diagnosis and Biomarker Discovery from Microarray Gene Expression Data". Statnikov A, Tsamardinos I, **Aliferis CF**. 13th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB), 2005.
- 69 "Health Education Software System for Accidents Prevention". **C. Aliferis**, S. Palamas, E. Maragkaki, M. Chioni, E. Petridou. Panhellenic Pediatric Conference, 1999.
- 70 "A Multimedia Health Promotion Software System for Teaching Accidents and First Aids to Children and Young Adults". E. Petridou, **C. Aliferis**, S. Palamas, E. Maragkaki, M. Hioni. 5th World Conference on Injury Prevention and Control, 1999.
- 71 "Improving the Cost-Effectiveness of Health Care Through Machine Learning Applied to Large Clinical Databases". G. Cooper, **C. Aliferis**, B.G. Buchanan, M.J. Fine, C. Glymour, G. Gordon, C. Meek, T. Mitchell, F. provost, T. Richardson, R. Scheines, and P. Spirtes. 1994 Annual Fall Meeting of the Biomedical Engineering Society.
- 72 "A Method for the Assessment and Monitoring of the Growth and Final Height of Greek Children With the Help of a Computer". **C. Aliferis**, E. Georgiou, K. Dalles, G. Raptis, C. proukakis. 15th Panhellenic Medical Congress, 1989.
- 73 "Incidence and Relationships of Gynecomastia in 954 Healthy Young Males". E.Georgiadis, L.Papandreou, C.Evangelopoulou, **C.Aliferis**, C.Panitsa, M.Batrinou. 7th International Congress on Senology, 3-7th May 1992, Rhodes, Greece. Abstracts book p. 198.
- 74 "W/H Ratio and Blood Pressure in Young Female Persons". N.Katsilambros, E.Georgiadis, **C.Aliferis**, C. Evangelopoulou, L.Papandreou, K.Kostas". 4th European Conference on Obesity, 7-9 May 1992 Noordwijkerhout, Netherlands. Int. J. Obesity (16) suppl. 1, p.69, May 1992
- 75 "W/H Ratio and BMI in a Group of Young Healthy Males, and Relationships to Blood Glucose and Other Factors". E. Georgiadis, L. Papandreou, **K. Aliferis**, N. Grigoriadis, A. Tzavaras, N. Katsilambros. 14th IDF Congress, Washington D.C. May 1991. Diabetes (40) 491a (supl. 1).
- 76 "Body Fat Distribution and Blood Pressure in Young Men". N. Katsilambros, E. Georgiadis, L. Papandreou, **C. Aliferis**, N. Gregoriadis, D. Triantafyllou. 3d Mediterranean Congress of Angiology, Rhodes, Greece May 1991. Abstracts book p. 18.
- 77 "W/H Ratio and BMI in Young Healthy Females and Relationships to Blood Pressure and Biochemical Factors". E. Georgiadis, **C. Aliferis**, L. Papandreou, K. Evangelopoulou, J. Katsarakis, N. Katsilambros. 6th European Nutrition Conference, Athens May 1991. Abstracts book p.73.

- 78 "Serum Lipids, Blood Glucose and Arterial Blood Pressure in a Group of Young Men - Relationships to Body Fat Distribution". N. Katsilambros, E. Georgiadis, **C. Aliferis**, D. Triantafyllou, S. Kouroutis, N. Gregoriadis, A. Tzavaras, L. Papandreou. Presented at the 26th annual meeting of the European Diabetes Epidemiology Study Group, Lund, Sweden May 1991.

C. Book Chapters & peer-reviewed tutorials (Selected):

79. "Causal Feature Selection". I. Guyon, **C.F. Aliferis**, A. Elisseeff. In: Computational Methods of Feature Selection, H. Liu and H. Motoda (Eds). Chapman and Hall (to appear).
80. "Selection And Interpretation Of Laboratory Procedures". E. Shultz, **C. F. Aliferis**, D. Aronsky. In: "Tietz Textbook of Clinical Chemistry and Molecular Diagnostics", Carl A. Burtis, Edward R. Ashwood, David E. Bruns (eds.). Elsevier Saunders, St Louis, MO, 2006.
81. "Machine Learning Methods for Biomedical Discovery and Decision Support". **C. F. Aliferis**, I. Tsamardinos. Full-day tutorial in World Congress in Medical Informatics, 2004.
82. "Machine Learning Methods for Biomedical Discovery and Decision Support". **C. F. Aliferis**, I. Tsamardinos. Full-day tutorial in American Medical Informatics Fall Symposium, 2003.
83. "Biomedical Informatics Training Program at Vanderbilt University". D. Aronsky, **C.F. Aliferis**, K. B. Johnson, N. Lorenzi, R.A. Miller. IMIA Yearbook of Medical Informatics 2004.
84. "On the Heuristic Nature of Medical Decision-Support Systems." **C.F. Aliferis**, and R. Miller, in: Yearbook of Medical Informatics Schattauer Publishers, JH van Bommel, A. T. McCray eds. 1996; 371-380.
85. "Medical Informatics in Preventive Medicine". **C. F. Aliferis**. In the Textbook of Preventive Medicine (E. Petridou and D. Trichopoulos eds.), Paschalidis Publications, Athens 2000

D. Invited Reviews/Editorials (Selected):

86. "Synopsis of Decision Support Systems". **C. Aliferis**, in "Yearbook of Medical Informatics 1997", J.H.van Bommel, A.T. McCray eds., Schattauer, 1997.
87. "Biomarker Selection from High-Dimensionality Data", S. Levy, A. Statnikov, **C.F. Aliferis**. Microarray Technology, September 2005.

E. Selected Technical Reports, and Education-Related Documents Not Published in Peer-Reviewed Forums:

88. "Proposal For A Program Of Study Leading To The Master Of Science And Doctor Of Philosophy Degrees In Biomedical Informatics", October 2000 (157 pages + financial projections).
89. "Scaling-Up Bayesian Network Learning to Thousands of Variables Using Local Learning Technique" I. Tsamardinos, **C. F. Aliferis**, A. Statnikov, L. E. Brown. DSL TR-03-02, 2003
90. "Algorithms For Large-Scale Local Causal Discovery And Feature Selection In The Presence Of Limited Sample Or Large Causal Neighborhoods" **C. F. Aliferis**, I. Tsamardinos Technical Report DSL-02-08

91. “Using Local Causal Induction To Improve Global Causal Discovery: Enhancing The Sparse Candidate Set” **C. F. Aliferis**, I. Tsamardinos. Technical Report DSL-02-04
92. “Large-Scale Feature Selection Using Markov Blanket Induction For The Prediction Of Protein-Drug Binding”. **C. F. Aliferis**, I. Tsamardinos Technical Report DSL-02-06

F. Software Packages, Patents (Selected)

93. **Causal Explorer:** A comprehensive software package for Bayesian network Induction, Feature/Biomarker Selection and Causal Structure Discovery. Available from: http://discover.mc.vanderbilt.edu/discover/public/causal_explorer/index.html
94. **GEMS (Gene Expression Model Selector)** available from: <http://www.gems-system.org/>
95. **EBM Search** available from: www.ebmsearch.org .
96. **Content and Quality Assessment Method and Apparatus for Biomedical Information Retrieval.** Yindalon Aphinanaphongs and **C.F. Aliferis** (US patent application number: 60/570,879).
97. **Method, System, And Apparatus for Causal Discovery and Variable Selection For Classification,** **C.F. Aliferis** and Ioannis Tsamardinos (US patent 10/439,374).
98. **Method And System For Automated Supervised Data Analysis.** **C.F. Aliferis**, A. Statnikov, N. Fanapazir, I. Tsamardinos (US patent application number: 11510847).

Professional Activities

RESEARCH :

Ongoing

1. **Co-Investigator:** Biomarker Profiles in the Diagnosis/Prognosis of ARDS
1U01 HL081332-01 (Ware)
2. **Co-Investigator:** Clinical Proteomic Technology Assessment for Cancer
1 U24 CA126479-01 (Liebler)
3. **Co-Principal Investigator:** “Causal Discovery Challenge” from PASCAL (Pattern Analysis, Statistical Modeling and Computational Learning). PASCAL is the European Commission's IST-funded Network of Excellence for Multimodal Interfaces.

Pending

4. **Principal Investigator:** Next-Generation Software for Discovery of Cancer Molecular Signatures and Markers. 1R01CA128602-01.
5. **Principal Investigator:** Causal Discovery Algorithms for Translational Research with High-Throughput Data. (2R01 LM007948-04).

6. **Co-Investigator:** Genomics and Genetics of Acute Graft-Versus-Host Disease
NCI, (Schuening)

7. **Co-Principal Investigator:** Causal Discovery Workbench and Challenge Program. NSF proposal
0725746.

8. **Co-investigator:** Probe-Centric Meta-Analysis of Microarray Data in Cancer. PAR-06-410 (Ni)

Completed

9. **Principal Investigator:** Principled methods for very-large-scale causal discovery. 1R01 LM007948-01

10. **Principal Investigator:** Pilot Project Computational Models of Lung Cancer:
Connecting Classification, Gene Selection, and Molecular Sub-typing NIH/NLM BISTI Planning Grant 1
P20 LM 007613-01 (Stead)

11. **Co-Principal Investigator** in Training Grant T15 LM07450-01 NIH/National Library of Medicine.

12. **Principal Investigator:** Vanderbilt Academic Venture Capital Fund support for the Discovery Systems
Laboratory (2001-2004).

13. **Principal Investigator:** Vanderbilt University Discovery Grant to study Complex Modeling of Clinical
Trial Data With Gene Expression Covariates & Development of Optimal Re-analysis Policies (2001)

14. **Principal Investigator:** WWW system to deliver pediatric care guidelines. Funded by the Ministry of
Health, Greek Government (2000-2).

15. **Co-Investigator:** SPOR in GI Cancer 1 P50 CA 95103-01 (Coffey)

16. **Co-Investigator:** Pharmacogenics of Arrhythmia Therapy 1U01 HL65962-01A1 (Roden)

17. **Co-Investigator:** SPOR in Breast Cancer 1 P50 CA (Arteaga)

Consulting:

- Prediction Sciences LLC (La Jolla, CA) (Proteomic-based stroke diagnosis models development).

- Ontar Corporation (Nort Andover, MA) (Tool for Pharmacovigilance for Infectious Disease, Combat
Casualty Care and Biological Warfare and Chemical Defense and Vaccines).

Reviewer Service

- Journal of Machine Learning Research
- Machine Learning Journal,
- Journal of the American Medical Informatics Association,
- BMC Bioinformatics,
- Cancer Informatics,
- Medical Decision Making,
- Journal of Biomedical Informatics.
- International Journal of Medical Informatics,

- Artificial Intelligence in Medicine,
- Transactions of the IEEE,
- Journal of Public Health Management and Practice,
- Future Generation Of Computer Systems,
- International Joint Conference in Artificial Intelligence (IJCAI),
- Pacific Symposium in Biocomputing (PSB),
- IEEE Transactions in Information technology in Biomedicine
- MEDINFO,,
- FLAIRS,
- AMIA Fall Conference: papers reviewer & tutorials reviewer.
- Book proposal reviewer for Elsevier

UNIVERSITY TEACHING:

1. Primary instructor for BMIF 315 “Foundations of Biomedical Informatics Methods” (graduate level core courses in MS/PhD Program and open to all graduate students). Taught in Spring 2007 and planned to be taught every year. From 2008 and on this course will be offered to EECS students too.
2. Primary instructor for BMIF 330, and BMIF 330a: “Biomedical Artificial Intelligence: Decision Support Systems and Machine Learning” (graduate level core courses in MS/PhD Program and open to all graduate students). Note: taught in 2002, 2003, 2004, 2006, with C. Aliferis as primary instructor; it will be offered from 2008 onward with Subramani Mani as primary instructor and C. Aliferis as co-instructor. From 2008 and on this course will be offered to EECS students too. Together BMIF315 and BMIF 330 will serve as primary courses in Machine Learning for EECS graduate students.
3. Lecturer in CANB 342 (graduate-level course in Cancer Biology for PhD students in that program).
4. Lecturer in MCI Clinical Proteomics class (graduate level).
5. Has offered independent studies in: Artificial Intelligence/Machine Learning, and Information Retrieval.

STUDENT MENTORING:

(i) Primary research advisor to:

1. Yin Aphinyanaphongs (MD/PhD student), expected to graduate in summer of 2007.
2. Alexander Statnikov (PhD student) expected to graduate in spring or summer of 2008.
3. Nafeh Fananapazir (MD/MS student) expected to graduate in summer of 2007.
4. Lawrence Fu (PhD student) expected to graduate in spring or summer of 2008.
5. Firas Wehbe (PhD student) expected to graduate in spring or summer of 2008.

(ii) Distinctions awarded to primary advisees:

- **Donald Lindberg Fellowship** awarded in 2005 to Yin Aphinyanaphongs (one such fellowship awarded annually accompanied by a \$25,000 grant award)
- **Gold Medal in 2004 MEDINFO** (tri-annual world congress of Medical informatics) Student Paper Competition for Alexander Statnikov (selected among 100 international competitors, accompanied by travel award and monetary prize)
- **First Prize in AMIA (American Medical Informatics Association) 2003 meeting** Student Paper Competition for Yin Aphinyanaphongs in (selected among 40 international competitors, accompanied by travel award and monetary prize)
- **Finalist in AMIA (American Medical Informatics Association) 2005 meeting** Student Paper Competition for Andrea Sboner (selected among 100 international competitors, accompanied by

travel award) (Andrea Sboner was visiting student to Vanderbilt/Discovery Systems Laboratory from Trento University).

(iii) Current and past member of thesis committees for:

- Laura Brown,
- Tricia Thornton,
- Joel Parker,
- Stephany Duda,
- Firas Wehbe,
- Lawrence Fu,
- Andrea Sboner (visiting student)

(iv) Post-doctoral project advisor to Eva Kasparova Ph.D.

SERVICE:

- **Architect and Founding Director, M.S./Ph.D. Program in Biomedical Informatics**, Vanderbilt University (2000- 2001): designed and successfully defended in front of Vanderbilt and external reviewers the Vanderbilt M.S./Ph.D. program in Biomedical Informatics. Wrote the 200+ page Program Proposal and Curriculum. Headed recruitments and day-to-day operations of the program.
- **Director & Founder, Discovery Systems Laboratory** (September 2001-present)
- **Organized the Gene Expression and Proteomics Data Analysis Interest Group** (summer 2000-summer 2001).
- **Department representative** to Graduate Faculty Assembly (2005-2007).
- Regular participant in the Vanderbilt-Ingram Cancer Center/Department of Biostatistics “**Hi-Dimensionality Core**” (service core analyzing all hi-dimensionality data generated in the cancer center).
- Member of the **consulting faculty of the Vanderbilt “Omics Clinic”** (Research design and analysis consulting service for mass-throughput data jointly organized by the Departments of Biostatistics and Biomedical Informatics).
- **Participation in University and Medical School Committees (selected):**
 1. Research Informatics Advisory Committee, Vanderbilt University (December 2000-2005).
 2. Vanderbilt MS-PhD Program in Biomedical Informatics Recruitment and Admissions Committees (2000, 2001).
 3. Ad-hoc DBMI Post-Doctoral Fellows Committee (May 2000).
 4. Education Committee for BISTI Planning Grant (2002-2003).
 5. Recruiting Team of the Vanderbilt Trans-Institutional Bioinformatics Initiative (2001-2004).
 6. Committee for Medical School Curriculum (2002).
 7. Ad-hoc committee for assessment of NuTek Inc. (2002).
 8. Retreat for Personalized Medicine (2002, 2005).
 9. Reviewer for Pilot and Feasibility Committee of the Vanderbilt Diabetes Center (2003).
 10. Faculty candidate & student candidate interviewing (for Department of Biomedical Informatics and other departments on an ad-hoc basis)
 11. Member of Biomedical Informatics MS/PhD Academic Program Committee (2003-2004)
 12. Member of Biomedical Informatics MS/PhD Curriculum Committee and ad-hoc sub-committee (2004-2005)
 13. Retreat for McKesson retreat (2006, 2007).
 14. Member of Mentoring committee for Terri Ni Ph.D., Research Assistant Professor.

INVITED TALKS AND RESEARCH-RELATED PRESENTATIONS & SEMINARS (SELECTED):

1. "Methodological rigor of "omics" data analysis: efficiency, stability, guidelines, and sensationalism". **Penn Bioinformatics Forum, University of Pennsylvania**, 3-14-2007
2. "Presentation to Journal Club for Analysis of Complex Datasets" (jointly with L. Wang). **Department of Biostatistics, Vanderbilt University**, 2-6-2007.
3. ""Omics"-Based Next-Generation Decision Support Tools". **McKesson – Vanderbilt Retreat** November 21, 2006.
4. "Discovery Systems Laboratory Research in Machine Learning for Discovery and Decision Support". NLM Training Meeting, **Vanderbilt University**, 6-27-2006.
5. "Pathway Induction And High-fidelity Simulation For Molecular Signature And Biomarker Discovery In Lung Cancer Using Microarray Gene Expression Data". **American Physiological Society Conference on Physiological Genomics and Proteomics of Lung Disease**, November 2-5, 2006.
6. "Presentation to Journal Club for Analysis of Complex Datasets" (jointly with F. Harrell). **Department of Biostatistics, Vanderbilt University**, 18 Apr 2006.
7. "The Rashomon Effect in discovery of biomarkers from mass-throughput cancer data: what it is and how it affects biological interpretability". Invited talk in **5th Annual joint retreat of Host-Tumor Interaction Program of the Vanderbilt-Ingram Cancer Center and the Dept. of Cancer Biology**, November 18-19th 2005
8. "Bayesian Network-Based Data Analysis: Algorithms and Experiments" **Vanderbilt University, Biostatistics Seminar Series**, June 8, 2005.
9. "Presentation to Journal Club for Analysis of Complex Datasets" (with F. Harrell). **Department of Biostatistics, Vanderbilt University**, 18 Apr 2006.
10. "Clinical Bioinformatics: Experiments in Diagnosis, Survival Prediction and Biomarker Discovery with Array Gene Expression and Mass Spectrometry Data Using A Novel Markov Blanket Induction Algorithm". **Siemens Corporate Research**, 1-21-2005.
11. "Biomedical Informatics methods and Systems In Support of Basic and Clinical Research", Research Seminar, **Beth Israel Deacones Medical Center, Harvard University**, 11-22-2004.
12. "Experiments in Diagnosis, Survival Prediction and Biomarker Discovery with Array Gene Expression and Mass Spectrometry Data Using a Novel Markov Blanket Induction Algorithm", Informatics program Research Seminar, **Children's Hospital Boston, Harvard University**, 11-23-2004.
13. "Feature Selection and Causal Discovery Using Local Structure Induction ", Research Seminar, **Department of Biomedical Informatics, Columbia University**, 4-23-2004
14. "Feature Selection and Discovery Using Local Structure Induction: Algorithms, Applications, and Challenges", **Stanford Medical Informatics Colloquium**, 4-8-2004
15. "Large-Scale Variable Selection with Causal Interpretation", School of Health Information Sciences, **University of Texas at Houston**, 3-24-2004
16. "Public Health Informatics, Special Panel: Issues & Opportunities" (plenary panel Speaker) **AMIA 2001, Spring Congress**, 2001
17. "Machine Learning For Bioinformatics: Feature Selection Methods", Seminar to **Bristol-Myers Squibb Company**, **Vanderbilt University**, 9-17-2002
18. D. Hardin and C.F. Aliferis, "Machine Learning Models For Lung Cancer Classification Using Array Comparative Genomic Hybridization", **Vanderbilt Biomathematical Study Group**, 5-4-2002
19. "Statistical Genomics: Making Sense Of All the Data: Bayesian Networks", **Vanderbilt University Statistical Genomics Seminar**, 5-22-2001
20. "Complex Data Modeling Using Machine Learning", Seminar to **Bristol-Myers Squibb Company**, **Vanderbilt University**, 3-13-02 and 3-14-2002
21. "Using Machine Learning and Statistical Methods to Predict Dire Outcomes Of Patients With Community Acquired Pneumonia", **National Center for Biotechnology Information (NCBI)**, October 2000.