

# BIOMEDICAL INFORMATICS PROGRAM

*Committee: (Chair and Program Director) Russ B. Altman; (Associate Director) Lawrence M. Fagan; Douglas L. Brutlag, Betty Cheng, Parvati Dev, Alan M. Garber, Teri Klein, Henry Lowe, Mark A. Musen, Gio Wiederhold*

## **Participating Faculty and Staff by Department:**

Opportunities for research are not limited to the specific faculty and departments listed.

*Anesthesia:* David M. Gaba (Professor)

*Biochemistry:* Douglas L. Brutlag (Professor)

*Biostatistics:* Richard A. Olshen (Professor)

*Business:* Alain C. Enthoven (Professor, emeritus), Alan M. Garber (Professor, by courtesy)

*Chemistry:* Vijay Pande (Assistant Professor)

*Civil and Environmental Engineering:* Raymond E. Levitt (Professor)

*Computer Science:* Serafim Batzoglou (Assistant Professor), Richard E. Fikes (Professor, Research), Leo Guibas (Professor), Daphne Koller (Associate Professor), Jean-Claude Latombe (Professor), Gio Wiederhold (Professor, Research, emeritus), Terry Winograd (Professor)

*Electrical Engineering:* Albert Macovski (Professor, emeritus)

*Genetics:* Russ B. Altman (Associate Professor), David Botstein (Professor), Mike Cherry (Associate Professor, Research), Stanley N. Cohen (Professor), Stuart Kim (Associate Professor), Teri Klein (Senior Research Scientist), Richard M. Myers (Professor)

*Health Research and Policy:* Byron W. Brown, Jr. (Professor, emeritus), Mark A. Hlatky (Professor), Richard A. Olshen (Professor), Robert Tibshirani (Professor)

*Management Science and Engineering:* Samuel Holtzman (Consulting Associate Professor), Ronald A. Howard (Professor), Ross D. Shachter (Associate Professor)

*Mathematics:* Samuel Karlin (Professor, emeritus)

*Mechanical Engineering:* Scott L. Delp (Associate Professor)

*Medicine:* Russ B. Altman (Associate Professor), Terrance Blaschke (Professor), Robert W. Carlson (Professor), Parvati Dev (Senior Research Scientist), Lawrence M. Fagan (Senior Research Scientist), Alan M. Garber (Professor), Mary Goldstein (Associate Professor), Michael Higgins (Consulting Associate Professor), Peter D. Karp (Consulting Assistant Professor), Teri Klein (Senior Research Scientist), John Koza (Consulting Professor), Henry Lowe (Associate Professor, Research; Senior Associate Dean for Information Resources and Technology), Mark A. Musen (Associate Professor), Douglas K. Owens (Associate Professor), Glenn Rennels (Consulting Assistant Professor), Gillian Sanders (Assistant Professor, Research), Peter Small (Associate Professor), Michael Walker (Consulting Assistant Professor)

*Neurosurgery:* John R. Adler (Professor), Ramin Shahidi (Assistant Professor, Research)

*Obstetrics and Gynecology:* W. LeRoy Heinrichs (Professor, emeritus)

*Pathology:* Arend Sidow (Assistant Professor)

*Radiation Oncology:* Arthur L. Boyer (Professor), Lei Xing (Assistant Professor, Research)

*Radiology:* Gary H. Glover (Professor), Sandy A. Napel (Associate Professor), Norbert J. Pelc (Professor), Geoffry Rubin (Associate Professor)

*Statistics:* Trevor J. Hastie (Professor), Art Owen (Professor)

*Structural Biology:* Michael Levitt (Professor)

*Surgery:* Thomas Krummel (Professor), Charles Taylor (Assistant Professor, Research)

*Department Offices:* MSOB 215

*Mail Code:* 94305-5479

*Department Phone:* (650) 723-6979

*Web site:* <http://www.smi.stanford.edu/academics/>

Courses given in Biomedical Informatics Program have the subject code BIOMEDIN. For a complete list of subject codes, see Appendix B.

This interdisciplinary program was created in response to a recognized need for well-trained researchers and academic leaders in the expanding field of biomedical informatics.

The program in Biomedical Informatics emphasizes research to develop novel computational methods that can advance biomedicine. Students receive training in the investigation of new approaches to conceptual modeling and to development of new algorithms that address challenging problems in the biological sciences and clinical medicine. Students with a primary interest in developing new informatics methods and knowledge are best suited for this program. Students with a primary interest in the biological or medical application of existing informatics techniques may be better suited for training in the application areas themselves.

## GRADUATE PROGRAMS

The Biomedical Informatics Program is interdepartmental and offers instruction and research opportunities leading to M.S. and Ph.D. degrees in Biomedical Informatics. All students are required to complete the core curriculum requirements outlined below, and also to elect additional courses to complement both their technical interests and their goals in applying informatics methods to clinical settings, biology, or imaging. Students who fail to maintain a 3.0 grade point average (GPA) in a category of the core curriculum are expected to pass a comprehensive exam in that area before the graduate degree is granted. In addition, all degree candidates must pass an oral examination that tests the student's ability to integrate the various components of the curriculum and to relate them to the overall field of biomedical informatics.

The core curriculum is common to all degrees offered by the program but is adapted or augmented depending on the interests and prior experience of the student. Deviations from the core curriculum outlined below must be justified in writing and approved by the student's Biomedical Informatics academic adviser and the chair of the Biomedical Informatics Committee. It should be noted, however, that the program is intended to provide flexibility and to complement other opportunities in applied medical research that exist at Stanford. Although most students are expected to comply with the basic program of study outlined here, special arrangements can be made for those with unusual needs or those simultaneously enrolled in other degree programs within the University. Similarly, students with prior relevant training will have the curriculum adjusted to eliminate requirements that were met as part of their prior training.

## CORE CURRICULUM

All students are expected to participate regularly in the Biomedical Informatics Student Seminar (201) and Colloquia (200), regardless of whether they register for credit in those courses. In addition, all students are expected to fulfill requirements in the following five categories:

1. *Core Biomedical Informatics* (15 units): students are expected to understand current applications of computers in biology and medicine and to develop a broad appreciation for research in the management of biomedical information. Required courses are the two quarter sequence Introduction to BIOMEDIN 210 and 211 both of which should be taken during the first year in the program. Students must also take an additional 6 units of Biomedical Informatics course work (which may include cross-listed courses from other departments, but not including BIOMEDIN 200, 201, 299, 302, or 303), selected in consultation with the academic adviser. BIOMEDIN 212 and 214 are strongly encouraged.
2. *Computer Science* (9 units): the student is expected to acquire a knowledge of the use of computers, computer organization, programming, and symbolic systems. It is assumed that students will have had by matriculation prior computing experience at least equivalent to a course introducing the fundamentals of data structures and algorithms such as CS 103A,B, 103X, 106A,B, 106X, or other courses approved

by academic adviser or executive committee. All students are required to take a minimum of 9 units of courses in the Department of Computer Science. If similar courses have not been taken previously, these units must include CS 121, 161, and a course that requires significant programming and knowledge of machine architectures (for example, EE 182, CS 110, or the CS 193 series). For those who have taken such courses previously, replacement units may be taken from any other course in CS selected by the student and approved by the academic adviser. A course in databases is especially recommended. With the exception of CS 110, all other courses applied to the degree requirements must be numbered 137 or higher.

3. *Probability, Statistics, and Decision Science* (9 units): students are required to take at least three courses that span the following five topics: basic probability theory, Bayesian statistics, decision analysis, machine learning, and experimental-design techniques. Prior courses in statistics at least equivalent to STATS 60 and calculus equivalent to MATH 42 are prerequisites. A prior course in linear algebra equivalent to MATH 103 or 113 is recommended. For the probability requirements, students may, for example, take MS&E 120, STATS 116, or MS&E 221. For the statistics requirements, sequences after STATS 116 can include STATS 200 followed by a course in stochastic modeling, machine learning or data mining, such as STATS 202, 215, or 315A,B, or CS 228 or 229. Options for decision analysis include MSE 152 or 252, or cost effectiveness analysis (BIOMEDIN 432). Specific courses should be chosen in consultation with the student's academic adviser. Also recommended is a course in the psychology of human problem solving.
4. *Biomedical Domain Knowledge* (9 units): students are expected to acquire an understanding of pertinent life sciences and how to analyze a domain of application interest. Prior courses in biology at least equivalent to BIOSCI 31 and 32 are prerequisites. All students without formal health care training must take BIOMEDIN 207.
5. *Social and Ethical Issues* (6 units): candidates are expected to be familiar with key issues regarding ethics, public policy, financing, organizational behavior, management, and pertinent legal topics. Students may select at least 6 units from suitable courses that include, for example, BIOMEDIN 250, 256, and 432; CS 201; HRP 390, 391, and 392, or any other advanced course in policy and social issues proposed by the student and approved by the Biomedical Informatics academic adviser.

The core curriculum generally entails a minimum of 46 units of course work, but can require substantially more or less depending upon the courses selected and the previous training of the student. The varying backgrounds of students are well recognized and no one is required to take courses in an area in which he or she has already been adequately trained; under such circumstances, students are permitted to skip courses or substitute more advanced work. Students design appropriate programs for their interests with the assistance and approval of their Biomedical Informatics academic adviser. At least 27 units of formal course work are expected.

## MASTER OF SCIENCE (ACADEMIC)

The University requirements for the M.S. degree are described in the "Graduate Degrees" section of this bulletin.

This degree is designed for individuals who wish to undertake in-depth study of biomedical informatics. Normally, a student spends two years in the program and will implement and document a substantial project during the second year. The first year involves acquiring the fundamental concepts and tools through course work and research project involvement. All first- and second-year students are expected to devote 50 percent or more of their time participating in research projects. Research rotations are not required, but can be done with approval of the academic adviser or training program director. Graduates of this program are prepared to contribute creatively to basic or applied projects in biomedical informatics. Biomedical Informatics recently has been approved to offer a Master of Science (Professional) degree. This new program is primarily designed for working professionals who already have advanced

training in one discipline and wish to acquire interdisciplinary skills. Details of this new program may be found on our web site at <http://www.smi.stanford.edu/academics/>.

## PROGRAM REQUIREMENTS

Programs of at least 54 units that meet the following guidelines are normally approved:

1. Completion of the core curriculum.
2. A minimum of 6 additional units of courses in Computer Science numbered 135 or higher, courses in Management Science and Engineering or Statistics numbered 200 or higher, PSYCH 256 or 267, or relevant courses in other departments approved by the student's academic adviser.
3. Electives: additional courses to bring the total to 54 or more units.
4. Teaching: all students are expected to act as Teaching Assistants (TAs) for at least one course during their first two years of training. This will generally be one of the informatics short courses, although another course approved by the program faculty may occasionally be substituted.

## DOCTOR OF PHILOSOPHY

The University's basic requirements for the doctorate (residence, dissertation, examination, and so on) are discussed in the "Graduate Degrees" section of this bulletin.

Individuals wishing to prepare themselves for careers as independent researchers in biomedical informatics, with applications experience in bioinformatics, clinical informatics, or imaging informatics, should apply for admission to the doctoral program. The following are additional requirements imposed by the Biomedical Informatics Interdisciplinary Committee:

1. A student should plan and successfully complete a coherent program of study including the core curriculum, oral examination, and additional requirements for the master's program. In addition, doctoral candidates are expected to take at least two more advanced courses (see categories under item '2' of the master's program requirements). In the first year, two or three research rotations are strongly encouraged. The master's requirements, including the oral examination, should be completed by the end of the second year in the program (six quarters of study, excluding summers). Doctoral students are generally advanced to Ph.D. candidacy after completing the oral examination. A student's academic adviser has primary responsibility for the adequacy of the program, which is regularly reviewed by the Graduate Study Committee of the Biomedical Informatics program.
2. To remain in the Ph.D. program, each student must attain a grade point average (GPA) as outlined in the master's programs above, and must pass a comprehensive exam covering introductory level graduate material in any curriculum category in which he or she fails to attain a GPA of 3.0. The student must fulfill these requirements and apply for admission to candidacy for the Ph.D. by the end of six quarters of study (excluding summers). In addition, reasonable progress in the student's research activities is expected of all doctoral candidates.
3. During the third year of training, generally in the Winter Quarter, each doctoral student is required to give a preproposal seminar that describes evolving research plans and allows program faculty to assure that the student is making good progress toward the definition of a doctoral dissertation topic. By the end of nine quarters (excluding summers), each student must orally present a thesis proposal to a dissertation committee that generally includes at least one member of the Graduate Study Committee of the Biomedical Informatics program. The committee determines whether the student's general knowledge of the field, and the details of the planned thesis, are sufficient to justify proceeding with the dissertation.
4. As part of the training for the Ph.D., each student is required to be a teaching assistant for two courses approved by the BMI exec; one should be completed in the first two years of study.
5. The most important requirement for the Ph.D. degree is the dissertation. Prior to the oral dissertation proposal and defense, each student

must secure the agreement of a member of the program faculty to act as dissertation adviser. The principal adviser need not be an active member of the Biomedical Informatics program faculty, but all committees should include at least one departmental faculty member.

6. No oral examination is required upon completion of the dissertation. The oral defense of the dissertation proposal satisfies the University oral examination requirement.
7. The student is expected to demonstrate an ability to present scholarly material orally and present his or her research in a lecture at a formal seminar.
8. The student is expected to demonstrate an ability to present scholarly material in concise written form. Each student is required to write a paper suitable for publication, usually discussing his or her doctoral research project. This paper must be approved by the student's academic adviser as suitable for submission to a refereed journal before the doctoral degree is conferred.
9. The dissertation must be accepted by a reading committee composed of the principal dissertation adviser, a member of the program faculty, and a third member chosen from anywhere within the University.

## COURSES

**BIOMEDIN 200. Biomedical Informatics Colloquium**—Series of colloquia offered by program faculty, students, and occasional guest lecturers. Credit available only to students in a Biomedical Informatics degree program. May be taken no more than three times for credit.

*1 unit, Aut, Win, Spr (Staff)*

**BIOMEDIN 201. Biomedical Informatics Student Seminar**—For all students and faculty. Participants report on recent relevant articles from the Biomedical Informatics literature or their research projects. The ongoing experience, with feedback from faculty, is intended to teach presentation skills to Biomedical Informatics trainees. Credit available only to students in an Biomedical Informatics degree program. (May be taken no more than three times for credit.)

*1 unit, Aut, Win, Spr (Staff)*

**BIOMEDIN 207. Introduction to Medicine**—(Same as IMMUNOL 230.) For doctoral students. Information and approaches used by physicians to understand human disease, by focusing on two multisystem disorders: type I and type II diabetes mellitus. Sources include computer demonstrations, web resources, disease simulation software, guided use of medical school teaching materials, and taped lectures in small groups. Students carry out quarter-long, team projects related to the class work.

*3-4 units, Spr (Mellins, Parnes)*

**BIOMEDIN 210. Introduction to Biomedical Informatics: Fundamental Methods**—(Same as CS 270.) Issues in the modeling, design, and implementation of computational systems for use in biomedicine. Topics: basic knowledge representation, controlled terminologies in medicine and biological science, fundamental algorithms, information dissemination and retrieval, knowledge acquisition, and ontologies. Emphasis is on the principles of modeling data and knowledge in biomedicine and on translation of resulting models into useful automated systems. Recommended: understanding of the basic principles of object-oriented systems at the level of CS 107.

*3 units, Aut (Musen)*

**BIOMEDIN 211. Introduction to Clinical Systems**—(Same as CS 271.) Survey of the major applications in clinical informatics, including imaging systems, information systems, and decision-support technology. Emphasis is on the system requirements, relevant data, standards, algorithms, and implementation issues in each area.

*3 units, Win (Lowe)*

**BIOMEDIN 212. Biomedical Informatics Project Course**—(Same as CS 272.) A hands-on, software-building course following up on introductory courses in biomedical informatics. Designed to allow students to create novel software. Students conceive, design, specify, implement, evaluate, and report on a software project in the domain of biomedicine.

Focus is on pragmatics of creating written proposals, providing status reports, and preparing final reports. Also aspects of software engineering at an introductory level such as version control, UML, and testing. Prerequisites: 210 or 214, or consent of instructor; CS 106.

*3 units, Aut (Altman, Cheng, Klein)*

**BIOMEDIN 214. Representations and Algorithms for Computational Molecular Biology**—(Same as CS 274.) Introduction to the basic computational issues and methods used in bioinformatics, including access and use of biological data sources on the Internet. Topics: basic algorithms for alignment of biological sequences and structures, computing with strings, phylogenetic tree construction, hidden Markov models, computing with networks of genes, basic structural computations on proteins, protein structure prediction, protein threading techniques, homology modeling, molecular dynamics and energy minimization, statistical analysis of 3D biological data, integration of diverse data sources, knowledge representation and controlled terminologies for molecular biology, graphical display of biological data, genetic algorithms and genetic programming applied to biological problems. See instructor for unit options. Prerequisites: programming skills and understanding of matrix algebra.

*1-5 units, Spr (Altman)*

**BIOMEDIN 216. Lectures on Representations and Algorithms for Molecular Biology**—Lecture series for BMI 214. Recommended: familiarity with biology.

*1 unit, Spr (Altman)*

**BIOMEDIN 226. Genetic Algorithms and Genetic Programming**—(Enroll in CS 426.)

*3 units (Kozal) not given 2002-03*

**BIOMEDIN 228. Influence Diagrams and Probabilistic Networks**—(Enroll in MS&E 355.)

*3 units, Win (Shachter)*

**BIOMEDIN 230. Seminar on Knowledge Acquisition for Expert Systems**—(Same as CS 525.) For graduate students. Discussion of experimental approaches to the construction of expert-system knowledge bases. Topics: interviewing techniques, formal and informal approaches to modeling expert knowledge, and automated tools that facilitate knowledge acquisition. Enrollment limited to 20. Prerequisite: one course in artificial intelligence.

*1-2 units (Musen) alternate years, given 2003-04*

**BIOMEDIN 233. Intermediate Biostatistics: Analysis of Discrete Data**—(Same as HRP 261, STATS 261.) The 2x2 table. Chi-square test. Fisher's exact test. Odds ratios. Sampling plans; case control and cohort studies. Series of 2x2 tables. Mantel-Haenszel. Other tests.  $k \times m$  tables. Matched data logistic models. Conditional logistic analysis, application to case-control data. Log-linear models. Generalized estimating equations for longitudinal data. Cell phones and car crashes: the crossover design. Special topics: generalized additive models, classification trees, bootstrap inference.

*3 units, Win (Olshen)*

**BIOMEDIN 234. Biomedical Genomics**—Overview of how genomics is influencing medical research and health-care delivery, illuminating the genomic discoveries being translated into diagnostic and therapeutic medical applications. Themes: the relevance of human genome project and functional genomics to inherited and acquired diseases, and the role of public databases and computational methods for solving problems in biology. Human genetic variation, SNPs, comparative genomics, computer models of biological processes, microbial genomics, pharmacogenomics, structure-based drug design, gene therapy. Case studies demonstrate the use of information technologies for converting molecular biological data into knowledge that can improve patient care and accelerate the discovery of new therapeutics.

*3 units, Win (Shafer)*

**BIOMEDIN 239. Computer Based Medical Education**—Directed reading and research for graduate-level students in the use of modern web-based hypermedia and simulation techniques in education. Possible topics: replacement of a lecture or a lab session, distance learning, student models, and clinical case simulations.

*1-6 units, any quarter (Dev)*

**BIOMEDIN 250. Introduction to Health Care Systems and Health Policy**—(Enroll in HRP 205.)

*2 units, Win (Baker)*

**BIOMEDIN 251. Outcomes Analysis**—Introduction to methods of conducting empirical studies which use large existing medical, survey, and other databases to ask both clinical and policy questions. Econometric and statistical models used to conduct medical outcomes research. How research is conducted on medical and health economics questions when a randomized trial is impossible. Problem sets emphasize hands-on data analysis and application of methods, including re-analyses of well-known studies. Prerequisites: one or more courses in probability, and statistics or biostatistics.

*3 units, Spr (Bhattacharya)*

**BIOMEDIN 262. Computational Genomics**—(Same as CS 262.) For graduate students. Introduction to the applications of computer science to genomics research and to basic concepts in genomics from a computer science point of view. Topics: algorithms for sequence analysis and their applications to the most current genomics research, sequence alignments, hidden Markov models, multiple alignment algorithms and heuristics such as Gibbs sampling, and the probabilistic interpretation of alignments. Applications of these tools to sequence analysis: DNA sequencing and assembly, genomic annotation of repeats, genes, and regulatory sequences, microarrays and gene expression, and comparative genomics. Prerequisites: 161 or familiarity with basic algorithmic concepts. Basic knowledge of genetics helpful, but not required.

*3 units, Spr (Batzoglou)*

**BIOMEDIN 278. Probabilistic Models in Artificial Intelligence**—(Enroll in CS 228.)

*3 units (Staff) not given 2002-03*

**BIOMEDIN 299. Directed Reading and Research**—For students wishing to receive credit for directed reading or research time.

*1-18 units, any quarter*

**BIOMEDIN 301. Special Topics in Biomedical Informatics**

*1-6 units, any quarter (Staff)*

**BIOMEDIN 328. Protein and Nucleic Acid Structure, Dynamics, and Engineering**—(Enroll in SBIO 228.)

*3 units, Win (Staff) alternate years, not given 2003-04*

**BIOMEDIN 345. Genome Databases Seminar**—Reading/demonstrations/analysis devoted to the databases used to encode molecular biology information, including DNA sequence, protein sequence and structure, gene expression, and other functional genomics data. Emphasis is on representation and integration of the many diverse data sources, and their presentation for biomedical and pharmaceutical researchers. Issues: data structures and ontologies, cross referencing, quality control and error detection, search processes, suitability of different DBMSs, data provenance, and privacy protection for patient derived information. Introduction and discussion by instructors and presentations by experts from commercial and research organizations. May be combined with a CS 395 project.

*1 unit, Win (Karp, Wiederhold)*

**BIOMEDIN 348. Computer Graphics: Image Synthesis Techniques**—(Enroll in CS 348B.)

*3-4 units, Spr (Hanrahan)*

**BIOMEDIN 374. Algorithms in Biology**—(Same as CS 374.) Algorithms and computational models applied to molecular biology and

genetics. The most current algorithms from a variety of biological areas. Topics: protein folding, phylogeny and haplotypes, genome rearrangements, sequence comparison, DNA computing, and self-assembly. May be repeated for credit. Prerequisites: 161, 262, or 274, or equivalents.

*3 units, Aut (Batzoglou)*

**BIOMEDIN 432. Analysis of Costs, Risks and Benefits of Health Care**—(Same as MGTECON 332, HRP 392.) How do you do cost-benefit analysis when the output is difficult or impossible to measure? How do M.B.A. analytic tools apply in health services? Study/discussion of the main literature on the principles of cost-benefit analysis applied to health care. Critical review of actual studies. Emphasis is on the art of practical application.

*4 units, Aut (Staff)*

**BIOMEDIN 459. Frontiers in Interdisciplinary Biosciences**—(Cross-listed in multiple departments in the schools of Humanities and Sciences, Engineering, and Medicine; students should enroll directly through their affiliated department, otherwise enroll in CHEMENG 459.) An introduction to cutting-edge research involving interdisciplinary approaches to bioscience and biotechnology; for specialists and non-specialists. Organized and sponsored by the Stanford BioX Program. Three seminars each quarter address a broad set of scientific and technical themes related to interdisciplinary approaches to important issues in bioengineering, medicine, and the chemical, physical, and biological sciences. Leading investigators from Stanford and throughout the world present the latest breakthroughs and endeavors that cut broadly across many core disciplines. Pre-seminars introduce basic concepts and provide background for non-experts. Registered students attend all pre-seminars in advance of the primary seminars, others welcome. Prerequisite: keen interest in all of science, engineering, and medicine with particular interest in life itself. Recommended: basic knowledge of mathematics, biology, chemistry, and physics.

*1 unit, Aut, Win, Spr (Robertson)*

## OVERSEAS STUDIES

Courses approved for the Biomedical Informatics major and taught overseas can be found in the "Overseas Studies" section of this bulletin, or in the Overseas Studies office, 126 Sweet Hall.

## PARIS

**BIOMEDIN 231. Computational Molecular Biology**—(Same as BIOC 218.)

*3 units, Aut, Win, Spr, via Internet (Brutlag)*

This file has been excerpted from the *Stanford Bulletin*, 2002-03, pages 638-642. Every effort has been made to insure accuracy; late changes (after print publication of the bulletin) may have been made here. Contact the editor of the *Stanford Bulletin* via email at arod@stanford.edu with changes, corrections, updates, etc.