Edward H. Shortliffe

Section on Medical Informatics,
Stanford University School of Medicine,
Stanford, CA, USA

Education and training

Medical Informatics Training at Stanford University School of Medicine

Abstract: Stanford University School of Medicine has offered graduate degrees in medical informatics since 1982. Located approximately 50 kilometers south of San Francisco near the city of Palo Alto, the university offers both MS and PhD degrees, combining research training with formal course requirements in clinical information sciences, bioinformatics, computer science, decision science, basic biomedicine, health economics, and social and ethical issues. Requirements are adapted to the varying backgrounds of trainees. Graduates of the program work in a variety of capacities, although the majority have sought careers in academia or in industrial research settings.

Introduction

Stanford University's Medical Information Sciences (MIS) training program is an interdepartmental program offering instruction and research opportunities leading to an MS or a PhD degree in medical informatics. The program is administratively based in the School of Medicine but is overseen by the Graduate Studies Committee of Stanford University and is viewed by the Graduate Division as a free-standing department for purposes of granting degrees. The faculty of the program, which numbers over 30 participants, is drawn broadly from throughout the medical school and other parts of the university. Areas of investigation are diverse and include topics such as decision-support systems, integrated workstations, knowledge acquisition, speech input, mobile computing, medical records, computational biology, medical imaging, reasoning under uncertainty, information retrieval, medical terminology and vocabularies, technology assessment, and health-services research.

This paper summarizes the history, philosophy, and current status of the Stanford training program. More detailed information may be found on the World Wide Web under Uniform Resource Locator (URL)

http://www-camis.stanford.edu/academic.html.

History of Medical Informatics Research and Training at Stanford University

Stanford Medical School has been heavily involved in medical computing research since the late 1960s when the ACME Resource, an early experiment in time-shared computation for biomedical research, was developed by geneticist Joshua Lederberg and his colleagues. The university was already benefiting from close ties to the local computer industry in "Silicon Valley", the nickname for the surrounding region in which a large number of electronics firms, starting with the Hewlett Packard Company, had been created since the 1950s. It was natural that many portions of the university would become involved in the exploding local interest in computation, and early collaborations between biomedical scientists from the medical school and computer scientists on the main campus led to the development of such well known programs as DENDRAL [1] and MYCIN [2] . In 1973 ACME evolved into a new federally funded resource, codirected by Professor Lederberg from the medical school and Professor Edward Feigenbaum from the Computer Science Department. SUMEX-AIM was a time-shared machine, available to a national community over the early ARPANET (predecessor to today's Internet), and devoted entirely to research on applications of artificial intelligence to problems in biomedicine. For the next 18 years SUMEX-AIM provided the support for a series of research efforts in biomedical artificial intelligence [3], both at Stanford [4-6] and elsewhere [7].

By 1980 it had become clear that a new discipline was emerging at the intersection among biomedicine, information science, and computer science. The first Symposia on Computer Applications in Medical Care (SCAMC) were held in the late 1970s and soon attracted substantial audiences, medical informatics journals were beginning to grow in both circulation and influence, and we at Stanford were beginning to receive large numbers of inquiries from students who wanted to study computer science while concentrating on biomedical applications.

We have previously described the series of events that led to the development of a formal interdepartmental degree program [8]. The term "medical informatics" was not yet well accepted in the United States, and "medical information sciences" was a commonly used alternative name for the field. We first proposed a degree program in medical information sciences in January 1981 and, after deliberation by several administrative bodies, the proposal was approved by the university's faculty senate in October 1982. The first class of four students entered in 1983, and the program grew to nine students in 1984, and to fifteen in 1985. Both the training and research activities broadened during the following decade so that artificial intelligence is now only one component of a large and varied set of informatics activities at Stanford.

We soon realized the need for a comprehensive textbook in the field, and spent much of the late 1980s developing a volume to use in our introductory informatics course [9]. Today the training program has 25-28 trainees enrolled per year, with MS students requiring a minimum of two years and PhD students from 4 to 5 years. Approximately two thirds of the trainees are physicians or medical students pursuing joint degrees, but we have also trained other health professionals as well as individuals without any prior medical training. All students are housed with the core faculty and staff in shared laboratory space that includes offices, student carrels, individual machines for all trainees, and shared research facilities. We typically receive 50-60 applications per year for 5-7 new positions. The program has 40 graduates, including 16 with doctorates and 24 with master's degrees.

The principal shared computing facilities used by MIS trainees are provided by the Center for Advanced Medical Informatics at Stanford (CAMIS), a shared resource funded in part by the National Library of Medicine. Essentially all computing facilities at the medical center and computer science department are linked together by a local communications network which is also connected to most machines on campus and to national academic and research communities through gateways to the Internet. The network is also connected to a variety of servers in the program's offices, including laser printers, file servers, and telecommunication gateways.

Several trainees are supported by post-doctoral or pre-doctoral stipends through a training grant from the National Library of Medicine. Other trainees, including foreign students, tend to be supported by external fellowships or by research assistantships provided by their research preceptors.

The training program is overseen by several core faculty who serve on the administrative and admissions committees. Edward Shortliffe (who oversees research on decision-support systems, patient records, and national networking) directs the training program and Lawrence Fagan (artificial intelligence, mobile computing, integrated decision support) is codirector. Drs. Shortliffe and Fagan are joined by Mark Musen (knowledge acquisition, architectures for intelligent systems), director of the program's admissions committee and head of the Section on Medical Informatics in which the training program is administratively based. The program's computing and communications environment, plus advanced systems software research, is overseen by the Symbolic Systems Resources Group (SSRG), directed by Thomas Rindfleisch. Russ Altman oversees our research activities in the area of computational biology, directs our bioinformatics training activities, and has built associations with faculty in the Departments of Genetics, Biochemistry, Cell Biology, Mathematics, and Statistics. Additional core faculty include Parvati Dev (instructional technology, imaging, and multimedia), Gio Wiederhold (databases), Yuval Shahar (artificial intelligence and temporal reasoning), and consulting faculty Michael Walker (bioinformatics and statistics), Max Henrion (reasoning under uncertainty), and Glenn Rennels (information retrieval, critical appraisal, and clinical-trial design).

Program Philosophy

Stanford's interdisciplinary program was created in response to a recognized need for well-trained researchers and academic leaders in the expanding field of medical informatics. The design of the Stanford program reflects our belief that the newness of the field of medical informatics, the need for trained professionals, and the broad opportunities available at Stanford make it appropriate to provide a wide range of training options. We therefore offer both MS and PhD degrees and custom-tailor the classroom and research requirements to the diverse backgrounds and professional needs of our students. We require all trainees to be formal degree candidates, believing that leaders in the field will require broad formal course exposure in addition to intense research training. The curriculum provides structured but flexible exposure to topics in the areas of clinical medicine (for trainees who are not already health professionals), computational biology, computer science, decision science, statistics, operations research, psychology, health policy, ethics, technology assessment, and medical informatics itself. Trainees attend weekly journal clubs presented by students and weekly research colloquia offered by faculty, students, staff, and visitors to the university.

All trainees spend half time during their first two years in formal coursework and the remainder in focused research projects, working with one of the program faculty (or any other faculty member in the university who agrees to oversee their work and assure its relevance to the medical informatics training goals of our program). All trainees take a comprehensive MIS oral examination after two years in the program, and MS candidates are also expected to complete a master's research practicum by this time. The PhD degree adds an additional two to three years, with formal defense of a thesis proposal at the end of the third year and a completed dissertation during the fourth or fifth year.

The overriding goal of our training environment is to teach our students to become medical informatics professionals, i.e., to view the field as their professional identity, to be committed to contributing to it in the future, and to understand the technical, logistical, financial, and organizational issues that shape the field. Since the field is young, we expect our graduates to be educators who, regardless of their eventual professional setting, can communicate well about the field and its role in the worlds of biomedicine and health-care delivery. We accordingly emphasize the development of skills in scientific writing and in oral presentations, with a set of structured experiences that help assure that all trainees are experienced in presenting their own ideas, plus the work of others, to varied audiences ranging from clinicians to computer scientists. There is also an emphasis on the culture of medical informatics, with trainees expected to learn about activities at other centers and the names and accomplishments of present and past leaders who have shaped the informatics field. We also offer a series of informal lectures on such practical matters as grants preparation, the organization and functioning of federal agencies, industrial relations, and institutional barriers to the implementation of clinical information technologies. Finally, because we believe that medical informatics is inherently interdisciplinary and requires an understanding of diverse aspects of the health-care and biomedical environments, we explicitly seek diversity among our trainees, bringing together physicians, nurses, and other health professionals with students who have no prior training in the health sciences. We also seek diversity in national backgrounds, ethnicity, and gender among our students.

Program of Study

Both MS and PhD candidates take the same core courses, although the curriculum is adapted or augmented depending upon the interests and prior experience of the student. Required courses are drawn from five major areas (see Appendix for a description of the MIS course offerings, including those cited by number below):

1. *Medical Informatics:* Students are expected to understand current applications of computers in medicine and to develop a broad appreciation for research in the management of biomedical information. Required courses are MIS 210 (Computer Applications in Medical Care), 211 (Computer-Assisted Medical Decision Making), 214 (Algorithms and Representations for Molecular Biology), and 212 (Project Course).
2. *Computer Science:* The student is expected to acquire a knowledge of the use of computers, computer organization, programming, and symbolic systems. Required courses include offerings in machine architectures, programming languages, analysis of algorithms, artificial intelligence, numerical analysis, human-computer interaction, and databases.
3. *Decision Theory and Statistics:* Students are expected to learn basic probability theory, Bayesian statistics, decision analysis techniques, statistical inference, and experimental design techniques.
4. *Biomedicine:* Students are expected to acquire a basic knowledge of human physiology, anatomy, and disease. MIS course offerings in this area include MIS202 (Clinical Diagnosis), MIS205 (Introduction to Clinical Environments) and MIS204 (Physiology for Informaticians).
5. *Health Policy/Social Issues:* Trainees are expected to be familiar with key issues regarding public health policy, financing, ethics, and legal topics. A variety of courses from the medical school, computer science department, and graduate school of business meet the core requirements in this category.

The MS degree is designed for individuals who wish to undertake in-depth study of medical informatics. Normally a student spends two years in the program and implements and documents a substantial project during the second year. The first year involves acquiring the fundamental concepts and tools through course work and initial research-project involvement. Graduates of this program are prepared to contribute creatively to basic or applied projects in medical informatics.

The PhD degree is designed for individuals wishing to prepare themselves for careers as independent researchers in medical informatics. PhD students must plan and successfully complete a coherent program of study including the core curriculum, oral examination, and all requirements for the master's program. In addition, doctoral candidates are expected to complete two additional advanced courses. The master's requirements must be completed by the end of the second year in the program. Each doctoral student is also required to serve as a teaching assistant in at least one MIS course. By the end of three years, each student must orally present a thesis proposal to a dissertation committee that 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. The student is expected to demonstrate an ability to present scholarly material orally and presents his or her research in a lecture at a formal seminar. 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.

Recent revisions to our curriculum have included the addition of new faculty and emphasis in the area of bioinformatics (the study of the application of computational technologies to problems in biology, with an emphasis on molecular biology). Several trainees are already actively involved in this field and came to Stanford in part because of their bioinformatics interests. We have learned that many of the research programs of other, more clinically oriented students share important methodological problems with the field of bioinformatics. Thus we see medical informatics as a broad discipline that involves fundamental methodologies of applicability in both biological and clinical domains. We accordingly believe that all trainees in medical informatics need basic exposure to both clinical and biological applications and issues, partly to understand the full breadth of the field and partly so that shared methodologies are fully appreciated.

Research Opportunities

Participating faculty for the MIS program are drawn from departments throughout Stanford University, including anesthesia, biochemistry, computer science, economics, electrical engineering, engineering-economic systems, genetics, health research and policy, internal medicine, obstetrics and gynecology, pathology, psychology, radiology, statistics, surgery, and the graduate schools of business and of education. The varied research opportunities for trainees can not be described in detail here, but they extend across the full range of topics in medical informatics and computational biology. Descriptions of many current research activities are available on the World Wide Web under URL http://www-camis.stanford.edu/research.

Program Status and Future Directions

The training program currently (March 1995) has 29 enrolled students, of whom all but three are PhD candidates. Of the current students, four postdoctoral and eight predoctoral trainees are supported by the National Library of Medicine and the remainder are supported as research assistants on sponsored projects or by personal fellowships. Eight students are non-US citizens, eleven are women, seven are medical students pursuing MD/PhD degrees, and nine are physicians who are pursuing informatics as their field of specialization. Although no nurses are currently enrolled, we have trained nurses and pharmacists in the past and continue to encourage applications from all health professionals, including medical librarians.

Among our 40 graduates, 15 are in academic positions, 14 are working in industry, one works for a hospital, 2 work for the federal government, 3 are in clinical practice, and 5 are completing residency training. We have found that there is high demand for individuals with formal training in medical informatics [10], despite a marketplace in which doctoral graduates in many fields are finding it difficult to find research positions in academia or industry.

The program's administrative status within the School of Medicine is undergoing transition. As the program has grown and established itself as a stable and valued educational and research component at our institution [11], reconsideration of suitable structures that will help nurture further faculty growth and development as well as student opportunities, are under discussion.

Acknowledgment

Supported in part by Training Grant LM07033 and Resource Grant LM05305, both from the National Library of Medicine.

References

1. Buchanan BG, Sutherland G, Feigenbaum E. Rediscovering some problems of artificial intelligence in the context of organic chemistry. In: Meltzer B, Michie D, ed. *Machine Intelligence*. Edinburgh: Edinburgh University Press, 1970: 209-54.
2. Shortliffe EH. *Computer-Based Medi*

cal Consultations: MYCIN. New York: Elsevier/North Holland, 1976.
3. Clancey WJ, Shortliffe EH, ed. *Readings in Medical Artificial Intelligence: The First Decade*. Reading, MA: Addison-Wesley, 1984.
4. Buchanan BG, Shortliffe EH, ed. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading, MA: Addison-Wesley, 1984.
5. Hayes-Roth B, Buchanan B, Lichtarge O, et al. Deriving protein structures from constraints. In: *Proceedings of AAAI-86 -Fifth National Conference on Artificial Intelligence*. Los Altos, CA: Morgan Kaufman Publ Inc, 1986:904-9.
6. Tu SW, Kahn MG, Musen MA, Ferguson JC, Shortliffe EH, Fagan LM. Episodic skeletal-plan refinement based on temporal data. Commun ACM 1989;32:1439-55.
7. Miller RA, Pople HE, Myers JD. INTERNIST-1: An experimental computer-based diagnostic consultant for general internal medicine. New Eng J Med 1982;307:468-76.
8. Shortliffe EH, Fagan LM. Research training in medical informatics: The Stanford experience. Acad Med 1989;64:575-8.
9. Shortliffe EH, Perreault LE, Wiederhold G, Fagan LM, eds. *Medical Informatics: Computer Applications in Health Care*. Reading, MA: Addison-Wesley, 1990.
10. Greenes RA, Shortliffe EH. Medical informatics: An emerging academic discipline and institutional priority. JAMA 1990;263:1114-20.
11. Shortliffe E. Medical informatics meets medical education. JAMA 1995; 273: 1061-65.

Appendix

This appendix summarizes courses that are offered by the training program with MIS course numbers. Note that many courses in the core curriculum are offered by other departments in the university and are not described here.

200. *Medical Information Sciences Colloquium*. Series of colloquia, offered by program faculty, students, and occasional guest lecturers.
201. *Medical Information Sciences Journal Club*. Journal club for all students and several faculty. Participants report on recent relevant articles from the MIS literature.
202. *Clinical Diagnosis*. Designed for the learning of techniques of interviewing and symptom analysis, through the study of a variety of common and well-defined clinical entities and by role-playing in a problem-solving setting.
203. *Intermediate Biostatistics*. Introduction to advanced statistical procedures commonly used in health services and epidemiological research, e.g., multiple linear regression, multiple logistic regression, actuarial analysis of observations on time to event with censoring, and the analysis of frequency data by Poisson and chi-squared methods.
204. *Physiology for Informaticians*. This course introduces students with a strong computer science or engineering background to the basic principles of human physiology, starting with a review of molecular biology and then covering critical organ systems: cardiac, pulmonary, renal, hematologic, neural, endocrine and GI.
205. *Introduction to Clinical Environments*. One half day per week is spent becoming familiar with a variety of clinical settings at Stanford Medical Center and the Veterans Administration Medical Center. Selected faculty introduce assigned students to the medical wards, outpatient clinics, emergency room, operating room, intensive care unit, psychiatry ward, and clinical lab.
210. *Computer Applications in Medical Care*. Survey of use of computers in the medical field. Includes a variety of research and applied environments and the factors that influence the acceptance of these applications. Topics: integration of computer systems in the medical center, hospital information systems, ambulatory care systems, medical databases and networking, bibliographic search, aids for disabled patients, image processing, computer-aided instruction, and decision support systems.
211. *Computer-based Medical Decision Making*. Overview of concepts in medical decision making and survey of methods for the implementation of such concepts in computer-based clinical decision-support tools. Emphasis on Bayesian statistics, decision analysis, neural networks, artificial intelligence/expert systems, belief networks, and the synergies among such approaches.
212. *Medical Informatics Project Course*. For students who have completed 210, 211, and 214 and who wish to implement those ideas in a computer program.
214. *Algorithms and Representations for Molecular Biology*. Introduces the fundamental algorithms and representations used for sequence and structure analysis within molecular biology. It stresses the primary literature in bioinformatics, and is a practical programming course in which the key algorithms are recreated by the students.
228. *Influence Diagrams and Probabilistic Networks*. Theory of networks as representations for decision analysis and probabilistic inference: influence diagrams, belief networks, and Markov random fields.
230. *Seminar on Knowledge Acquisition for Expert Systems*. 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.
231. *Computer Applications in Molecular Biology*. The flow of information from genome to structure, from structure to biochemical function, and from function to phenotype. Reviews and evaluates current computer methods used in molecular biology.
239. *Computer-based Medical Education*. Directed reading and research for graduate-level students in the use of modern hypermedia techniques in education.
256. *Economics of Health and Medical Care*. Open to graduate students with training in microeconomics and some background in statistics or mathematics. Empirical, institutional, and theoretical analysis of problems of health and medical care.
299. *Directed Reading and Research*. For students wishing to receive credit for directed reading or research time.
348. *Computer Graphics: Image Synthesis Techniques*. Intermediate level, emphasizing sampling, shading, and display aspects of computer graphics.
354. *Probabilistic Reasoning in Computing*. Basics of (Bayesian) probability theory as applied to computing and artificial intelligence. Theory is illustrated by case studies. Practical problems in learning, search, approximate reasoning, data analysis, and decision-making under uncertainty.
432. *Cost-benefit Analysis in Health Care*. How do you do cost-benefit analysis when the "output" is difficult or impossible to measure? How do the MBA analytic tools apply in health services? Study and discussion of the main literature on the principles of cost-benefit analysis as applied to health care.

Address of the author:
Edward H. Shortliffe, M.D., Ph.D.,
Room MSOB X-215,
Section on Medical Informatics,
Stanford University School of Medicine,
300 Pasteur Drive,
Stanford, CA 94305-5479,
USA