Education and Training

J.J. Cimino, B.A. Allen, P.D. Clayton

Department of Medical Informatics, Columbia University, College of Physicians and Surgeons, New York, NY, USA

Medical Informatics Training at Columbia University and the Columbia-Presbyterian Medical Center

Abstract: The Department of Medical Informatics at Columbia University College of Physicians and Surgeons consists of a faculty of 17 full- and part-time faculty. The Department faculty collaborate with the Department of Computer Science and several clinical departments of the medical center. We offer courses in medical informatics, formal degrees (M.A., M.Phil. and Ph.D.) and a postdoctoral training program. In addition to academic offerings, the close affiliation with the Columbia-Presbyterian Medical Center and the primary responsibilities for clinical information systems offers trainees unique opportunities to work with and develop real-world applications. Faculty research programs include work on the Integrated Advanced Information Management System (IAIMS), Unified Medical Language System (UMLS), High-Performance Computing and Communications (HPCC), Electronic Medical Records, automated decision support and technology transfer through the Center for Advanced Technology.

Keywords: Medical Informatics, Education, Training, Training Program, Graduate Program

1. The Department of Medical Informatics

The Center for Medical Information Science was established at the Columbia University College of Physicians and Surgeons in 1987. In 1994 it became the Department of Medical Informatics, the second such university department in the United States. It is the only formal medical informatics educational program in the New York Metropolitan Area. The department has seventeen faculty, including eleven full-time faculty. A number of major research projects are ongoing with sponsorship from a variety of federal,

state and commercial sources. During its short history, the Department (and the Center before it) has made several significant contributions to national medical informatics research projects such as the Arden Syntax for Medical Logic Modules, the Unified Medical Language System, and the Health-Level 7 standard for medical data interchange. An appendix included below lists references to papers describing research of the Department. Current projects include clinician data capture, automated decision support, high performance computing and communication, electronic medical records, controlled medical vocabu-

lary, natural language processing, clinical data systems architecture, security and confidentiality, outcomes evaluation, care plans and practice guidelines, database design, voice recognition, computer-based curricula, dental informatics, and community information networks.

An important characteristic of the Department is its close ties to the information systems at the Presbyterian Hospital and the Health Sciences campus of Columbia University, which together form the Columbia-Presbyterian Medical Center (CPMC). Through a special relationship, the Department is responsible for the hospital's de-

partment of Clinical Information Services (CIS). By working with the other information systems departments in the hospital, the CIS is building a state-of-the-art clinical information system. This partnership provides opportunities to apply the Department's research to the development of applications which can be deployed in real-world settings where they can have concrete impact on the health care process and undergo meaningful evaluation. Current clinical applications include a centralized clinical database, a decision support engine, in-patient and out-patient physician workstations, and ancillary systems for pharmacy, laboratory and radiology departments. The central clinical information system also includes data from cardiology, gastroenterology, neurology, surgery and discharge summary systems. The campus-wide network provides access to over 3,000 workstations in 31 buildings at 21 geographic locations.

The Department's ties to the broader health sciences community, especially the Health Sciences Library, were established through its Integrated Academic Information Management System (IAIMS) project which culminated in a five-year implementation grant. Through this grant, the Department was able to initiate a campus-wide local area network and provide the means for coordinating and sharing on-line information resources for use by health care providers, researchers, educators, students and administrators. An additional grant from New York State has allowed us to establish a Center for Advanced Technology which supports collaborative technology transfer projects with the Computer Science Department of the University

These four aspects of the Department (medical informatics research, clinical computing, academic information systems and technology transfer), create a special environment for the training of medical informatics

professionals. Trainees have the opportunity to test their innovative uses of information in the real world and see the fruits of their own research incorporated into the dynamic computing environment found at CPMC. In its educational role, the Department offers medical informatics courses, as well as programs leading to Masters and Doctoral degrees, and it administers a medical informatics postdoctoral training program.

2. Medical Informatics Courses

As might be expected of any new educational unit, the Department of Medical Informatics is experiencing a period of growth in its course offerings. Because of the interdisciplinary nature of the field, many of the topics in medical informatics are adequately covered by other University departments such as Computer Science, Epidemiology, Biostatistics and Physiology. Our department focuses on teaching those subjects which occur at the intersections of different fields. Our offerings include the regular teaching of Introduction to Medical Informatics, Quantitative Methods of Medical Decision Making and a Research Seminar in Medical Informatics. The faculty also teach a variety of courses based on demand and interest (see Table 1). For example, in the Spring of 1994, the Department offered courses in Medical Language Processing and Topics in Hospital Information Systems. This latter course explored in detail the ways in which medical data are encoded, stored, retrieved, exchanged and manipulated in real hospital information systems. Students may also pursue independent research projects under the guidance of a faculty member. Recent student projects have included natural language processing of ECG reports, linking clinical systems to on-line knowledge

sources, comparison of knowledg representation schemes for use in medi cal applications, voice data input an encoding, and translation of user que ries into conceptual graphs to facil tate automated information retrieval

3. The Degree Programs

The graduate training program designed to train students both fr academic careers as researchers ar teachers, as well as for profession positions in the field of Medic Informatics. Students in the progra study the form and function of fac about patients or aggregates of p tients. They examine the influence knowledge on the processing of su facts when making scientific hypot eses and diagnostic or therapeutic d cisions. They also learn about met ods for measuring the impact of infc mation, as well as approaches for ef ciently acquiring and displaying rt evant information in a meaningful wa Candidates are considered based their prior academic and experient preparation, motivation, intellectu capacity, and ability in quantitati and verbal areas. Financial support provided for those students accept beyond the MA degree program.

M.A. Degree

The M.A. degree can pursued a terminal degree in its own right, or a prerequisite to the M.Phil. degr The basic requirement for admiss: as a student in the Medical Informat program is a bachelor's degree, whi should ordinarily be in computer, ma ematical, physical, engineering, or b logical science. M.A. candidates m demonstrate competence in the bas of computer science. Students are mitted on either a full-time or p. time basis to the M.A. degree ɛ gram, but must complete the deg requirements within a two year peri The exact program of study is plan

on the basis of the student's background and interests, in consultation with members of the Medical Informatics Degree Program Committee, who assign each student a faculty advisor. Except where the student has previously studied them, the coursework must cover linear algebra, probability theory (or mathematical statistics), data structures, fundamental algorithms, software laboratory and scientific computation. Table 1 provides a partial listing of courses which can be used to meet the M.A. requirements. In addition to coursework, students submit an original essay in the form of a thesis or publication quality manuscript before completion of the second year of study.

M.Phil. Degree

The M.Phil. degree is generally obtained as a prerequisite to the Ph.D. program. Students are admitted on a full-time basis only for the M.Phil degree program. The program carries its own course requirements beyond the M.A. degree in a selection of courses approved by the Program Committee. Students must complete an essay, as for the M.A. degree (they may petition to use their M.A. essay toward this requirement). The program requires teaching experience: all M.Phil. students are expected to participate in the education activities of the Department. Students must also pass a three-part exam, consisting of a written exam, an oral exam (in the form of a presentation and defense of a research proposal) and public presentation of a seminar. The M.Phil. degree program is to be completed by the end of the third year of study (including the two years needed for the M.A. requirements), except for those students granted advanced standing, who must complete the degree by the end of the second year of study.

Ph.D. Degree

Ph.D. candidates must complete all requirements for the M.Phil. degree. The Ph.D. program consists of the preparation, defense and deposition of a doctoral dissertation. The Ph.D. degree should be completed within five years, including earning the M.A. and M.Phil.

---

4. The Postdoctoral Training Program

The Department is fortunate to be the recipient of a Post-Doctoral Training Program grant, sponsored by the National Library of Medicine. The training program provides a stipend, tuition, health benefits and travel funds. A prospective candidate must hold a doctoral degree (M.D., Ph.D. or equivalent) and be a U.S. citizen or permanent resident. Beyond these absolute requirements, candidates are considered based on their knowledge and experience in the health and information sciences. Strong candidates are generally those with (a) a health-related (typically M.D.) or library science-related background who have some demonstrated aptitude in computer programming and computer science or (b) strong computer science background and some demonstrated aptitude in a health-related field. Candidates are generally expected to make a commitment for three years of training, with academic years running from July to June. There are three components to the program: faculty supervision and direction, proficiency requirements, and research.

Faculty Supervision and Direction

At the start of the training program,

Medical Informatics
- Introduction To Medical Informatics
- Quantitative Models For Medical Decision Making
- Medical Language Processing
- Topics in Hospital Information Systems
- Research Seminar In Medical Informatics
- Readings In Medical Informatics
- M.Phil. Teaching Experience
- Doctoral Research
- Doctoral Dissertation

Computer Science
- Database Systems
- Programming Languages and Translators
- Operating Systems
- Computer Networks
- Software Engineering
- Computer Graphics
- Artificial Intelligence
- Natural Language Processing
- Knowledge-Based Expert Systems

Physiology
- Human Physiology

Biostatistics
- Introductory Probability
- Design Of Medical Experiments
- Analysis Of Categorical Data
- Generalized Linear Models
- Research Methodology

Table 1 - Partial listing of courses which can be used to meet degree requirements.

each trainee is assigned a member of the Department faculty who serves as an advisor. The advisor assists in orienting the trainee and helps the trainee to assess areas of pre-existing proficiency where coursework may not be necessary as well as areas where additional coursework beyond the requirements would be helpful. The advisor guides the trainee in formulating a research project and proposal, assists the trainee with problems that may arise, provides general support, monitors progress, and gives the trainee appropriate feedback. There are meetings of the advisors and the Program Directors at least once each semester to discuss the progress and performance of each trainee.

Proficiency Requirements

The academic program is intended to provide trainees with a basic understanding of the wide range of subjects encompassed by medical informatics. Trainees generally meet proficiency requirements through formal coursework, although some trainees come to the program with backgrounds and abilities that address some of these requirements and are given waivers. Course requirements for each trainee are structured based on individual need and direction, but generally follow those of the Masters degree programs described above. All NLM trainees are required to complete a course in medical research ethics.

Research Requirement

Each trainee formulates and conducts a research project. The trainee submits a written research proposal by the end of the first year of the program. The proposal follows the outline of an NIH grant proposal and addresses areas such as required resources, specific aims, significance, methods, references and human subjects (if appropriate). The role of the trainee is to conduct his or her own

research, not to function as a programmer on the team of one of the faculty. The research project is expected to lead to publication in a leading peer reviewed journal. Trainees present their research proposals and, later, their results in regular Department research seminars.

The "Typical" Experience

Due to variability in the background and goals of the trainees, the training program is customized for each trainee. However, some generalizations about the experience can be made. Each trainee needs to find a balance between coursework and research. Most trainees devote a high proportion of their time to courses in the first year with less in the second and little or none in the third, with a corresponding increase in research work. The first year is often an exploratory year, with the trainee doing some preliminary project or a small part of an on-going faculty project. During the second year, the trainee submits the formal research proposal and commences work on it. Presentation of results generally occurs in the third year. A large proportion of our trainees are physicians who have completed their residency training. Those who wish to maintain their clinical skills can do so on a limited basis as long as it does not interfere with their ability to meet the program requirements.

In addition to formal courses, trainees attend (and eventually present at) the weekly medical informatics seminar, the monthly journal club, an annual scientific meeting (such as SCAMC) and the annual NLM Training Directors meeting.

5. Discussion

The medical informatics training opportunities at Columbia are many and varied. There are several characteristics which serve to strengthen the learning experience. The educational resources of Columbia University are among the best in the

world, particularly in the health sci ences. The Presbyterian Hospital is on of the largest teaching and research hos pitals in the country, with 1,500 bed 600 residents and 1,200 attending phy sicians, and provides a broad range c health care settings in which to establis medical informatics research. Trainee can learn by studying the clinical con puting environment that has been deve oped thus far but can also gain valuabl hands-on experience by choosing re search projects which can be incorpc rated in the existing environment, to b tested by real users and affect the car of real patients. They can also partic pate with the faculty in research e forts such as the New York State Cer ter for Advanced Technology and se eral national research projects. Th variety of opportunities and flexibilit of the program almost guarantees th each trainee will have a unique expe rience. The program is quite new, bu the graduates thus far have establishe successful careers in medica informatics research, hospital infoi mation systems and commercial sof ware publishing.

Additional Information

For more information about the de gree programs, contact our secretar for educational affairs at (212) 30: 1816 or Dr. Barry Allen at (212) 30: 3815, fax (212) 305-3302 or via e-ma at allenba@cucis.cis.columbia.ed For more information about tl postdoctoral training program, cor tact Dr. James Cimino at (212) 30: 8127, fax (212) 305-3302 or via e-ma at cimino@columbia.edu. To apply a degree program, contact the Offic to Graduate Affairs, 701 West 168 Street, Room 406, New York, N 10032. Telephone: (212) 305-805 Fax: (212) 305-1031.

Sample of Faculty Publications

1. Allen BA, Levinthal C. CARTOS II semi-automated nerve tracing: three-dimensional reconstruction from serial section micrographs. Comput Med Imag Graph 1990;14:319-29.
2. Barrows RC Jr, Allen B, Fink DJ. An X Window system for statlab results reporting. In: Safran C, ed. *Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care.* New York: McGraw-Hill, 1993:331-5.
3. Cimino JJ, Aguirre A, Johnson SB, Peng P. Generic queries for meeting clinical information needs. Bull Med Lib Assoc 1993;81:195-206.
4. Cimino JJ, Clayton PD, Hripcsak G, Johnson SB. Knowledge-based approaches to the Maintenance of a large controlled medical terminology. J Am Med Inform Assoc 1994;1:35-50.
5. Cimino JJ. Data storage and knowledge representation for clinical workstations. Intern J Bio-Med Comput 1994;34:185-94.
6. Cimino JJ, Johnson SB. Use of the unified medical language system in patient care. Meth Inform Med 1995: in press.
7. Clayton PD, Sideli RV, Sengupta S. Open architecture and integrated information at Columbia-Presbyterian Medical Center. MD Comput 1992;9:297-303.
8. Clayton, PD. Integrated advanced medical information systems (IAIMS): payoffs and problems. Meth Inform Med 1994;33:351-7.
9. Friedman C, Sideli R. Tolerating spelling erros during patient validation. Comput Biomed Res 1992;25:486-509.
10. Friedman C, Alderson PO, Austin JHM, Cimino JJ, and Johnson SB. A general natural language text processor for clinical radiology. J Am Medic Inform Assoc 1994;1:161-74.
11. Friedman C, Cimino JJ, Johnson SB. A schema for representing medical language applied to clinical radiology. J Am Medic Inform Assoc 1994;1:233-48.
12. Friedman C, Huff SM, Hersh WR, Pattison-Gordon E, Cimino JJ. The Canon effort: working toward a merged model. J Am Medic Inform Assoc 1995;2:4-18.
13. Johnson SB, Cimino JJ, Friedman C, Hripcsak G, Clayton PD. Using metadata to integrate medical knowledge in a clinical information system. In: Miller RA, ed. *Proceedings of the Fourteenth Annual Symposium on Computer Applications in Medical Care.* 1990:340-4.
14. Johnson SB, Aguirre A, Peng P, Cimino JJ. Interpreting natural language queries using the UMLS. In: Safran C, ed. *Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care.* New York: McGraw-Hill 1993:294-8.
15. Hripcsak G. Using connectionist modules for decision support. Meth Inform Medic 1990;29:167-81.
16. Hripcsak G. Arden syntax for medical logic modules. MD Comput 1991;8:76-8.
17. Hripcsak G, Cimino JJ, Johnson SB, Clayton PD. The Columbia-Presbyterian Medical Center decision-support system as a model for implementing the Arden Syntax. In: Clayton PD, ed. *Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care.* 1991:248-52.
18. Hripcsak G, Clayton PD, Cimino JJ, Johnson SB, Friedman C. Medical decision support at Columbia-Presbyterian Medical Center. In: Timmers T, Blum BI, eds. *Software Engineering in Medical Informatics.* Amsterdam: North-Holland, 1991:471-9.
19. Hripcsak G, Johnson SB, Clayton PD. Desperately seeking data: knowledge base-database links. In: Safran C, ed. *Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care.* New York: McGraw-Hill 1993:639-43.
20. Hripcsak G. Monitoring the monitor: automated statistical tracking of a clinical event monitor. Comput Biomed Res 1993;26:449-66.
21. Hripcsak G, Friedman C, Alderson PO, DuMouchel W, Johnson SB, Clayton PD. Unlocking clinical data from narrative reports: a study of natural language processing. Ann Intern Med 1995;122:681-88.
22. Johnson S, Friedman C, Cimino JJ, Clark T, Hripcsak G, Clayton PD. Conceptual data model for a central patient database. In: Clayton PD, ed. *Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care.* 1991:381-5.
23. Pryor TA, Hripcsak G. The Arden syntax for medical logic modules. Intern J Clinic Monitor Comput 1993;10:215-24.
24. Sengupta S, Clayton PD, Molholt P, Sideli RV, Cimino JJ, Hripcsak G, Johnson SB, Allen B, McCormack M, Hill C. IAIMS and sharing. Intern J Bio-Medic Comput 1994;34:339-48.
25. Sherman EH, Shortliffe EH. A user-adaptable interface to predict users' needs. In: Schneider-Hufschmidt T, Kuhme T, Malinowski T, eds. *Adaptive User Interfaces: Principle and Practice.* Amsterdam: North-Holland, 285-315.
26. Zimmerman JL, Lang WP, Williams JN, Spohn EE. DENTIN (Dental Information Network) a communications and information resource for dental education and practice. J Dent Educ 1994;58:241-5.
27. Zimmerman J. Dental informatics: strategic issues for the dental profession. J Dent Educ 1990;54:600-1.

Addresses of authors:
James J. Cimino, MD,
Barry A. Allen, PhD,
Paul D. Clayton, PhD,
Department of Medical Informatics,
Atchley Pavilion, Room 1310,
Columbia-Presbyterian Medical Center,
New York, NY 10032,
USA