Education and Training

The Harvard-MIT-NEMC Research Training Program in Medical Informatics

Robert A. Greenes[1]
G. Octo Barnett[2]
Stephen G. Pauker[3]
Peter Szolovits[4]
Milton C. Weinstein[5]

[1]Brigham and Women's Hospital,
[2]Massachusetts General Hospital,
[3]Tufts-New England Medical Center,
[4]Massachusetts Institute of Technology,
[5]Harvard School of Public Health,
Boston, MA, USA

Abstract: The Harvard-MIT-NEMC Research Training Program in Medical Informatics brings together five separate research groups to provide a diversified training experience in Boston with sponsorship by the National Library of Medicine. This program offers predoctoral and postdoctoral programs, supplemented by many opportunities for collegial exchange and interaction among the groups.

Keywords: Medical Informatics, Tranining Programs

A. History of Medical Informatics Training in the Boston Area

The Boston area is home to a number of medical informatics research and development groups, located at Harvard University, Massachusetts Institute of Technology, and Tufts University, as well as in several affiliated hospitals of Harvard and Tufts. Educational and training opportunities in the field have been offered by some of these groups for many years.

Among the earliest training opportunities was that provided by the Laboratory of Computer Science, at Massachusetts General Hospital (MGH), in the mid-1960s, in which physicians, other professionals, or students were able to affiliate on a somewhat informal basis, and work on one or more ongoing research projects, focused primarily on hospital information system development. In the mid-1970s, graduate training in computer science with an emphasis on medical artificial intelligence began at Massachusetts Institute of Technology (MIT). In the 1980s, with National Library of Medicine (NLM) support, a training program emphasizing clinical decision analysis was initiated at Tufts-New England Medical Center (NEMC); the program also involved Dartmouth Medical School, and had a loose affiliation with MIT. In 1985, a Harvard training program was also started with NLM support, involving the Decision Systems Group at Brigham and Women's Hospital (BWH), the Laboratory of Computer Science at MGH, and the Departments of Biostatistics and Health Policy and Management at the Harvard School of Public Health (HSPH).

In 1992, with the aid of continuing NLM support, the separate training programs of MIT, NEMC, HSPH, MGH, and BWH were organized into a combined Harvard-MIT-NEMC research training program, involving five participating groups:

- The Clinical Decision Making Group (CDMG) group, at MIT, directed by Peter Szolovits, Ph.D.,
- The Health Decision Science (HDS) program at HSPH, directed by Milton C. Weinstein, Ph.D.,
- The Decision Systems Group (DSG), at BWH, directed by Robert A. Greenes, M.D., Ph.D.,
- The Laboratory of Computer Science (LCS) at MGH, directed by G. Octo Barnett, M.D.,
- The Division of Clinical Decision Making (DCDM) at NEMC, directed by Stephen G. Pauker, M.D.

In the remainder of this paper, we describe this combined program, its design, research and development emphases, and experience. We then reflect on current challenges in medical informatics and the way we expect to address them.

Supported in part by grant LM 07037 from the National Library of Medicine

B. Design of the Combined Harvard-MIT-NEMC Training Program

The combined Harvard-MIT-NEMC program is administered by the Division of Health Science and Technology (HST), a joint division of Harvard Medical School (HMS) and MIT. This provides considerable flexibility in terms of course offerings from both institutions, and administrative mechanisms for handling the financial aspects of the program. While there is no formal academic Department or Center for Medical Informatics in the Boston area, the coordination within HST enables this Division to serve as an academic focal point for the program. Faculty in the program can have appointments in the HST Division, although their primary academic appointments are in other Departments at Harvard, MIT, or Tufts University.

The co-directors of the program, who are directors of their own research laboratories, jointly review all applicants and may encourage applicants to apply to their tracks; the applicant make a formal application to one or more of the specific tracks of the training program, since we believe this promotes the best match between the interests and capabilities of the applicant and the unique opportunities of each of the tracks.

The Harvard-MIT-NEMC training program provides support through NLM funding for 22 trainees annually, 6 of whom are at the predoctoral level and 16 at the postdoctoral level. Additional trainees with other sources of support also participate in the program.

Predoctoral training is offered for Ph.D. candidates in computer science, through the CDMG program a MIT, or in health decision science, through the HDS program at HSPH. Courses in these programs are also available to postdoctoral fellows, some of whom are also in degree programs as described below.

In the MIT predoctoral program, students study and perform research leading to a Ph.D. in computer science from the MIT Department of Electrical Engineering and Computer Science. Most students choose research topics closely related to the CDMG's overall research interests. Students entering without an advanced degree (e.g., M.S. or M.D.) must complete a Master's on the way to the Ph.D., which involves coursework across a variety of aspects of computer science and a modest research thesis. Students entering directly from their Bachelor's training typically take six years to complete the Master's and Ph.D. programs.

The HDS predoctoral program is part of the Health Policy and Management Department and the Department of Biostatistics at HSPH. It represents both a formal degree program and a concentration of research interests. The Doctoral program in Health Decision Sciences at Harvard University is designed to train researchers in the methods and applications of decision science to health problems.

Postdoctoral training is provided largely at the three hospital-affiliated laboratories, the DSG, LCS, and DCDM. While trainees often take courses and occasionally concurrently work on Master's degrees, this is primarily a research track. Most postdoctoral trainees are MDs, who have completed their residency training in any of a variety of specialties. Occasional trainees enter with other professional degrees or with a prior Ph.D. Some MDs have entered directly after medical school, or before completing their residencies, or after several years of clinical practice. Traineeships are usually for a period of two to three years. MDs are encouraged to continue patient care responsibilities on a limited basis during their training.

Two additional postdoctoral degree positions are offered, typically for MDs who are seeking a degree in the computer science Ph.D. program, or in the HDS Ph.D. program. The program is ideal for the physician who decides to specialize in medical informatics or health decision science research and wishes to pursue this specialty at the highest academic levels. Some postdoctoral candidates may wish to spend less time than the full Ph.D. program typically requires, and may plan, instead, to end their studies with a Master's degree.

Short-term traineeships for medical students are also available, and are used to introduce medical informatics research to qualified students at our medical schools, with the intent of attracting them to this field.

Educational opportunities include course offerings at MIT in many scientific and technical fields; at HSPH in biostatistics, decision science, epidemiology, health policy and management, and computer science. Courses are also available in other schools of Harvard University including the Business School, the Kennedy School of Government, and the Division of Applied Science, as well as at Tufts University. In addition, students may find helpful the facilities of the Harvard/MIT Division of HST, an interdisciplinary training and research division that focuses on the biomedical applications of technology. The Department of Brain and Cognitive Science at MIT offers useful contacts for students specifically interested in cognitive issues, and MIT's Biology department and the Whitehead Institute specialize in research problems in molecular biology and genetics.

Both Harvard and MIT are home to many seminar series, guest lectures, discussion groups, and journal clubs. Cooperation with the various Boston-

area hospitals and medical research centers provides opportunities for collaboration on a wide variety of theoretical and practical projects.

While we have maintained the format of a confederation of individual tracks, several academic and social events provide the opportunity for fellows and faculty at the various laboratories to get to know each other. A common e-mail list announces the large number of seminars, journal clubs, symposia, short courses, and other events at the various laboratories as well as collective activities of the program as a whole. The fellows at the different laboratories play an active role in organizing speakers and programs on topics of interest to them.

Other features

The Harvard-MIT-NEMC training program is somewhat unique, in three respects:

1. The program represents a confederation of medical informatics groups with a long tradition of independence and a wide spectrum of essentially non-overlapping research interests that tend to complement each other well. By design, therefore, the program continues to reflect the character and diversity of the individual laboratories. As a consequence, it accommodates a diverse range of trainees, with differing research and career goals.
2. Given that medical informatics professionals must function at the intersection of a number of fields, we believe that a major goal as well as challenge for training is to combine both breadth and depth. In this program, depth is achieved through involvement and assumption of responsibility for project work in individual laboratories. Breadth is achieved by expanding one's horizons beyond the particular project, in terms of core knowledge and skill development, and continuing opportunities for becoming acquainted with the diversity of the field and its challenges. Through the combined program, we aim to give fellows the opportunity for a broader and richer training experience, wider range of collegial interactions, and increased appreciation of the range of the field than were possible when the programs were separate.
3. Subsequent career paths of trainees vary from purely academic to highly applied, and with jobs spanning the range from researcher to educator to manager to commercial entrepreneur. By virtue of the range of activities in the participating laboratories, and in the information systems departments of the affiliated hospitals, we are fortunate to be able to provide a mix of theoretical, applied, and operational training experiences.

C. Participating Laboratories

1. Clinical Decision Making Group (CDMG), MIT

The primary focus of the Clinical Decision Making Group is the development and application of new methods of artificial intelligence for use in medical systems for diagnosis and therapeutic management. Since the mid-1970s, the CDMG has developed various conceptual models, and computer programs that embody them, e.g., for diagnosis of kidney diseases, diagnosis and therapy of heart disorders, design of cancer therapy protocols. The CDMG has also developed technical approaches to allowing programs to explain their knowledge, strategies and conclusions, and to teach based on that same knowledge.

The complex patients that most require assistance and monitoring from computers demand the ability to describe the normal functioning of the human body, the causes and mechanisms of disease, the body's response to external and internal derangements, the abilities of tests to probe aspects of the internal reality, and the expectable reactions of the body to complex interventions. Simply to encode the required knowledge in a computer requires the ability to represent knowledge whose character is at once temporal, spatial, probabilistic, and taxonomic.

An expert system must also be able to represent typical and routine plans and actions, and preferences of both the patient and the treating physician for different interventions and outcomes. The wealth of knowledge that is potentially relevant to making a medical decision makes it imperative that reasoning strategies help focus and narrow a reasoner's attention to just those issues and that knowledge that is in fact likely to be useful in a particular setting.

Medical AI (AIM) programs may be designed to serve as consultants on complex problems where the medical practitioner knows ahead of time that help may be necessary. They may also serve as more limited reasoning tools to help the user explore particular consequences and models. Alternatively, they can be used as background monitors to detect unusual, possibly dangerous conditions that can then be brought to the attention of the health-care providers. THE CDMG has explored a range of possible applications spanning this spectrum. In addition, it has begun to build personal medical agents (called Guardian Angels), that provide life-long health record collection, interpretation, and guidance to individual patients.

2. Health Decision Science Program (HDS), HSPH

Health decision science is defined to include decision analysis, math-

ematical modeling, behavioral decision theory, cost-effectiveness and cost-benefit analysis, and other aspects of statistics, economics, psychology, epidemiology, management science, and computer science applicable to the study of decision making in health care. Research interests in HDS are particularly focused on cost-effectiveness of clinical practices and technologies, population modeling for chronic disease, utility analysis related to quality of life measurement and valuation, risk analysis with particular focus on the role of expert judgment and on outcome valuation, statistical methods in clinical decision analysis including diagnostic technology assessment, and Markov modeling of public health decisions.

Some representative recent projects include:

a. A working group on utilities and outcome measurement relates decision-theoretic and psychometric approaches, considering both theoretical investigations and survey research for assessing outcome preferences.
b. The Coronary Heart Disease Modeling Project provides opportunities for computer model building, synthesis of data for model input, and applied simulation studies to evaluate interventions to prevent or treat coronary heart disease.
c. A study comparing cost-effectiveness of health interventions across medicine, environmental health, and public safety, has resulted in a number of methodological questions about the appropriate basis for comparison of programs for purposes of priority setting.
d. The Acute Myocardial Infarction Patient Outcome Research Team (AMI/PORT) includes a decision analytic component.
e. A study group in the Harvard AIDS Institute, is developing models for decision making for research and development. This is part of a larger set of interests in methodologies for allocation of resources to research.

3. Decision Systems Group (DSG), BWH

The Decision Systems Group (DSG) is a research unit in medical informatics at Harvard Medical School, located in the Brigham and Women's Hospital. Primary interests of the DSG are in means to support health care professional and student education and decision making. Trainees participate actively in the projects of the DSG, and are encouraged to take on project responsibility.

Primary areas of activity of the DSG include:

a. Information architectures to support collaborative work. This is a software engineering project aimed at development of infrastructure tools and methods to facilitate construction of applications via composition from diverse, independently created, network-distributed information resources. A new software development environment called Arachne has evolved from the DSG's previous work on DeSyGNER. Features include: (1) a cross-platform GUI, (2) a generic object classification scheme, and (3) tools for remote procedure calls and client-server communication.
b. An Internet-based collaboratory. The DSG and other Harvard collaborators are working with Stanford and Columbia on a new project involving development of a set of distributed, object-oriented tools and resources that can be shared by all, and can be incorporated into various applications being implemented by individual participants.
c. Structured taxonomies and clinical data capture. As part of the NLM's "Unified Medical Language System" project, the DSG has been pursuing the problem of structured data entry for capture of physicians' notes. Related goals include: (1) development of an acceptable user interface for direct capture of reports; and (2) refinement of a semantic-net model for findings and diagnoses that represents the range of granularity of perceptual vs. interpretive findings.
d. Clinical management frameworks and guidelines. A long-term interest of the DSG has been the representation of clinical guideline knowledge in a form that facilitates its integration into the process of patient care. This has included work in navigation and display of guidelines, and in verification of consistency and completeness of guideline knowledge. A new project seeks to combine guidelines with structured data entry. The DSG is also exploring the use of portable pen-based devices as a means for accessing information resources.
e. Medical student education, continuing medical education, and evaluation of clinical competence. The DSG has built a number of tools for incorporating multimedia objects into educational software for Harvard Medical School's curriculum. More recently, it has also been building tools for clinical competence assessment for the American Board of Radiology and the American Board of Family Practice. A primary approach has focused on case simulations.
f. Integrated clinical workstation. As an application of many of the infrastructure tools and components described above, the DSG has been exploring the development of an integrated workstation that combines clinical data access, image access and manipulation, and access to supporting information and knowledge (e.g., expert systems, guidelines, textbook knowledge, MEDLINE, Gopher and WWW resources, and others). An initial fo-

cus is a workstation for radiology consultation and review.

4. Laboratory of Computer Sicence (LCS), MGH

The doctoral fellowship program at the Laboratory of Computer Science at Massachusetts General Hospital is primarily an intense experience in developing and implementing a project in the application of computer technology to problems in medical care, medical education, or medical research. Each trainee works closely with LCS senior personnel in specifying a problem, designing a project, and carrying out the programming development. The one requirement is that the project of a trainee be closely related to a project which is of interest to the Laboratory, and is in an area which has independent funding support. The trainee is expected to become proficient in computer programming so that he/she can carry out relatively independent project development. The research training is similar to a graduate thesis, in that the trainee is expected to participate in the development and specification of the project, to present it at national meetings, and to prepare manuscripts for publication.

The Laboratory of Computer Science was established 30 years ago as a unit of the HMS/MGH Department of Medicine to carry out research and development in Medical Informatics. This unit was one of the first computer science divisions to be established in the nation in a medical school/hospital environment. Among the accomplishments of the Laboratory of Computer Science have been: (a) the design and programming of one of the first comprehensive hospital information systems using modular definition as the specification paradigm; (b) the development and national dissemination of MUMPS - the most widely used programming language for medical applications; (c) the development and national dissemination of COSTAR - one of the most comprehensive and widely used automated ambulatory medical record systems; (d) the development of a large library of medical education programs (the RxDx series) which are distributed through a medical publisher (Williams & Wilkins); and (e) the development and dissemination of a diagnostic decision-support system (DXplain) to medical schools and teaching hospitals. The LCS has been involved in the training of postdoctoral medical fellows for over two decades, and many of the graduates of this program now hold positions of senior academic rank in a variety of medical schools in the USA and Canada.

The LCS has active research programs in the application of computer technology to medical education, clinical research, expert systems for medical diagnosis, medical record systems, knowledge management, and clinical problem solving. The tradition at LCS is to emphasize creation, implementation and evaluation of practical real-world applications which are used in patient care and medical education. The selection of research areas of LCS is based on a strategy of developing and implementing computer-based systems that meet high-priority information needs for medical care. The selection of research areas is intensely problem-based, and guided by windows of opportunity - both technological and institutional. A successful project is evaluated in part by the degree to which it is accepted and utilized in patient care, both locally and nationally.

The major projects at present include: a system which provides expert knowledge relevant to problems in the intensive care unit; a workstation-oriented ambulatory medical record system; a diagnostic decision-support system; the creation of a software platform that allows the easy entry of clinical queries and retrieval of relevant information from a variety of computer-based knowledge resources (LCS has been a prime contractor since the beginning of the UMLS activities); and development of patient simulation programs for medical education.

5. Division of Clinical Decision Making (DCDM), NEMC

The DCDM, in the Department of Medicine at the NEMC, has been involved in research and training of physicians in the applications of computers, information theory and quantitative reasoning to medicine for about two decades. Medical informatics training at NEMC has been funded by NLM for 14 years. The Clinical Decision Making Consultation Service forms a unique core of clinical material for research and teaching activities.

The training program focuses on developing skills for performing clinical decision analyses in individual patients, health policy and cost-effectiveness analyses for populations and providing decision support to clinicians. Trainees obtain didactic exposure to these disciplines, and each trainee undertakes an array of projects under the supervision of a faculty member. Trainees typically produce several manuscripts and submit several abstracts to national meetings. The Division includes four faculty members.

One unique feature of the DCDM program is its clinical decision-making consultation service which receives requests for analyses about clinical conundra involving either tradeoffs among quantity and quality of life, various morbidities, or patients with unusual combinations of disease. Trainees rotate through the Service and have the opportunity to participate

in a variety of projects involving that patient base. Each case is analyzed by a trainee in collaboration with a faculty member. Each analysis involves a literature review, the construction of an explicit decision- analytic model, and extensive sensitivity analyses. A weekly case discussion conference allows constructive criticism and a selection of this material is published in Clinical Decision Conferences (CDCs), with the trainee as the first author.

This work has stimulated development of a variety of decision tree analysis support programs, including the original program, DECISION-MAKER, and its successors, D-MAKER and SMLTREE, all of which have involved trainees in their development. Current application domains for health policy analyses include cardiovascular disease, diabetes, gastrointestinal disease, infectious disease, neonatal screening, neurology and oncology, but this array responds flexibly to the interests of trainees.

A sampling of current and recent projects undertaken by trainees includes: a general model comparing neonatal screening programs for different disease; developing a database to combine decision models with the patient record; studies of cognitive errors and biases of clinicians; models of using radiofrequency ablation for symptomatic and asymptomatic WPW and of using implantable defibrillators for survivors of cardiac arrest; developing a generic platform for assessing patient utilities; models concerning the treatment of hepatitis and cirrhosis; models of screening and treatment of HIV disease; models comparing bypass surgery, angioplasty, and medical therapy for angina; developing new approaches for analyzing time-varying state transitions models; models of the cost-effectiveness of CMV immune globulin in renal and liver transplant recipients; and exploring the limitations of simple models of survival.

D. Experience of the Program and Future Directions

Considering only NLM funding since 1980, for both the previously separate and recently combined programs, over 70 NLM-supported trainees have matriculated to date.

Predoctoral trainees who have completed their doctoral studies typically pursue careers as professors at universities or as researchers at academic, government, or industrial research laboratories. Students who end their studies with the Master's degree have gone to work in industry for large corporations as well as start-up firms, or at medical centers.

Postdoctoral trainees usually continue to pursue their medical careers, mostly in faculty positions in medicine where they pursue their medical informatics research interests. Trainees tend to subsequently devote a significant portion of their professional career in the development and support of computer-based applications in medical care, medical education, and medical research.

Regarding the training program as a whole, we are beginning to see new joint projects evolving that involve the collaboration of participating laboratories. This has resulted in part from the increasing robustness of communication networks and universal involvement of faculty and trainees in e-mail exchanges. Increasing interest is also being focused by our medical schools and hospitals on ways to exchange patient data, outcomes data, guidelines, and other information resources, as well as the ability to provide access to clinical images and other multimedia content. In addition, emphasis in several of our laboratories on software engineering aimed at distributed software architectures and client-server development have provided the opportunity for some of the specialized capabilities of the participating laboratories to be made available as remotely accessible components, such as expert system, information resources, and analytical tools.

We are excited by the future directions in medical informatics and about our ability to capitalize on them both in our own continuing research and development, and in order to provide an ever richer training experience in this field.

Addresses of the authors:
Robert A. Greenes, M.D., Ph.D.[1]
G. Octo Barnett, M.D.[2]
Stephen G. Pauker, M.D.[3]
Peter Szolovits, Ph.D.[4]
Milton C. Weinstein, Ph.D.[5]

[1]Brigham and Women's Hospital, Boston, MA
Decision Systems Group,
75 Francis Street, Boston, MA 02115, USA
[2]Massachusetts General Hospital, Boston, MA
[3]Tufts-New England Medical Center, Boston, MA
[4]Massachusetts Institute of Technology, Cambridge, MA
[5]Harvard School of Public Health, Boston, MA