AI Magazine Volume 20 Number 3 (1999) (© AAAI)

AI in Medicine

The Spectrum of Challenges from Managed Care to Molecular Medicine

Russ B. Altman, M.D., Ph.D

■ AI has embraced medical applications from its inception, and some of the earliest work in successful application of AI technology occurred in medical contexts. Medicine in the twenty-first century will be very different than medicine in the late twentieth century. Fortunately, the technical challenges to AI that emerge are similar, and the prospects for success are high.

When I was asked to make this presentation, the organizers specifically asked me to review a bit of the history of AI in medicine (AIM) and to provide an update of sorts. I have therefore taken the liberty of dividing the last 30 years of medical AI research into three eras: the era of diagnosis, the era of managed care, and the era of molecular medicine. A description of these eras allows me to review for you some of the early and current work in AIM and then tell you about some of the exciting opportunities now emerging.

Why is AI in medicine even worth considering? In the late 1950s, medicine was already drawing the attention of computer scientists principally because it contains so many stereotypical reasoning tasks. At the same time, these tasks are fairly structured and so are amenable to automation. Every medical student learns that when one thinks about a disease, one thinks in an orderly way about epidemiology, pathophysiology, diagnosis, treatment, and then prognosis. These are the bins into which medical information is parsed. These sorts of structured reasoning methods made medicine an attractive application area. In addition, medicine is clearly knowledge intensive, and so at places like Stanford (where knowledge was power [Feigenbaum 1984]), it was very tempting to try to encode knowledge for the purposes of reproducing expert performance at diagnosis and treatment. The working hypothesis was that rich knowledge representations would be sufficient, with only relatively weak inference algorithms required. There was (and is) considerable debate about how complex inference should be for expert performance, but it is clear that medicine is a field in which there is a lot of knowledge required for good performance. It is also clear that physicians constantly feel a sense of overload as they deal with the individual data associated with their patients as well as the content knowledge of medicine that they are trying to apply to the treatment of these patients. I can try to provide a feel for the information-processing load on a physician: A full-time general practitioner is currently expected to longitudinally follow a panel of 2000 to 2500 patients. Of course, the severity of illness varies, but it is clear that physicians need systems (computer or otherwise) to track the data pertaining to these patients and turn it into working hypotheses for diagnosis, treatment, and long-term prognosis.

The other appeal to working in AI in medicine is that the field is large, and so virtually all aspects of intelligent behavior can be studied in one part or another of medicine. You can study issues of image processing, automated management of database information, robotic automation of laboratories, computer-assisted diagnosis, multimedia for physician and patient education, virtual and telesurgery, and many other issues. For some, AI in medicine provides a kinder, gentler, "greener" application area in which to apply their techniques.

Three Eras for AI in Medicine

The first era of AI in medicine was the "Era of Diagnosis." The first aspect of medical reasoning that caught the imagination of AI researchers was the process of collecting clinical data and applying inference rules to make the diagnosis of a disease. This is the common image of the doctor as sleuth, determining what disease is causing the patient's symptoms. The second era of AI in medicine was what I have called the "Era of Managed Care of Chronic Disease." This era has approached a set of problems quite distinct from those tackled in the preceding period, as I will discuss. Finally, we are on the precipice of the "Era of Molecular Medicine," which is once again going to raise issues that are different from those occupying researchers during the first two.

The Era of Diagnosis

In 1959, Ledley and Lusted (1959) published a paper in *Science* entitled "The Reasoning Foundations of Medical Diagnosis." This classic paper has the feature of many classic papers: It puts forth a series of statements that are now taken as almost self-evident. Ledley and Lusted pointed out that medical reasoning was not magic but instead contained well-recognized inference strategies: Boolean logic, symbolic inference, and Bayesian probability. In particular, diagnostic reasoning could be formulated using all three of these techniques. Their paper mapped a research program for the next 15 years, as investigators spun out the consequences of applying these inference strategies to medical domains.

The research that followed was varied and excellent, and I cannot properly review all the contributions but instead will pick some exemplary efforts. For example, in 1965 Lawrence Weed introduced a computer system called PROMIS to support a problem-oriented medical information methodology (Tufo et al. 1977). Weed's work was among the first to demonstrate a truly electronic medical record. Moreover, this record was a highly structured, strongly typed data structure (in many ways similar to our modern frame-based systems) that even today is rarely matched in its insistence on structured data input. Weed's work was limited by the absence of standard terminologies to use within his data structure, but his belief in structured data is still a major goal within the medical informatics community.

In the late 1960s the National Library of Medicine (NLM) (www.nlm.nih.gov/) was established as one of the National Institutes of Health (NIH). This was remarkable for many reasons, not least of which was that most institutes within the NIH are associated with an organ or a disease (for example, The National Institute of Heart, Lung, and Blood or The National Cancer Institute). The NLM is still in search of its organ or disease. Nevertheless, the extramural research program of the NLM has been a principal source of research funds for AI in medicine. The principal intramural contribution from the NLM was the creation of an online database of the published biomedical literature, MEDLINE. Having gone through a number of transformations, the MEDLINE database was recently made available to the general public via the PUBMED resource on the World Wide Web (www.ncbi.nlm.nih.gov/PubMed/). For better or worse (I believe for the better), physicians and patients now have unprecedented access to a literature that is growing exponentially. The challenges in indexing, searching, and parsing this literature represent a major challenge to AI investigators.

The 1970s brought the push for diagnostic performance. De Dombal et al. in 1972 showed that you could make clinically accurate diagnoses using Bayesian inference (de Dombal et al. 1972). Also in 1972, Kulikowski and the team at Rutgers created the CASNET system in which they explored methods for using causal models for somewhat deeper diagnostic tasks (Kulikowski and Weiss 1982). The models were deeper because physiological models were now being used to explain symptoms and describe diagnostic possibilities. Shortliffe, Buchanan, and coworkers showed soon afterwards (with the MYCIN system in Buchanan and Shortliffe [1984]) that production rules could be used to make expert-level diagnosis of infectious diseases. Pauker and coworkers created the PIP system (Presenting Illness Program) in which the cognitive processes associated with short-term and long-term memory were modeled in order to create programs that could consider multiple diagnoses but then focus on the few most likely solutions quickly (Szolovits and Pauker 1976). Figure 1 shows a figure from one of the PIP papers, in which an associative memory structure is modeled. As particular concepts are activated and drawn into the river representing active memory, they drag into the river with them associated ideas that then come to the attention of the inference engine.

A magnum opus during this period was the INTERNIST knowledge base and inference program published by Miller, Myers, and coworkers (Miller, Pople, and Myers 1982). INTERNIST had the goal of diagnosing any problem within general internal medicine—basically any systemic disease or disease of the organs between the neck and the pelvis. INTERNIST was based on a very large knowledge base that was transferred to a PC-based program called QMR, which now forms the basis for a commercial product (Miller, Masarie, and Myers 1986). The INTERNIST/QMR knowledge base associated diseases with findings using two numbers: a frequency of association and an evoking strength. There was then an algorithm created for col-

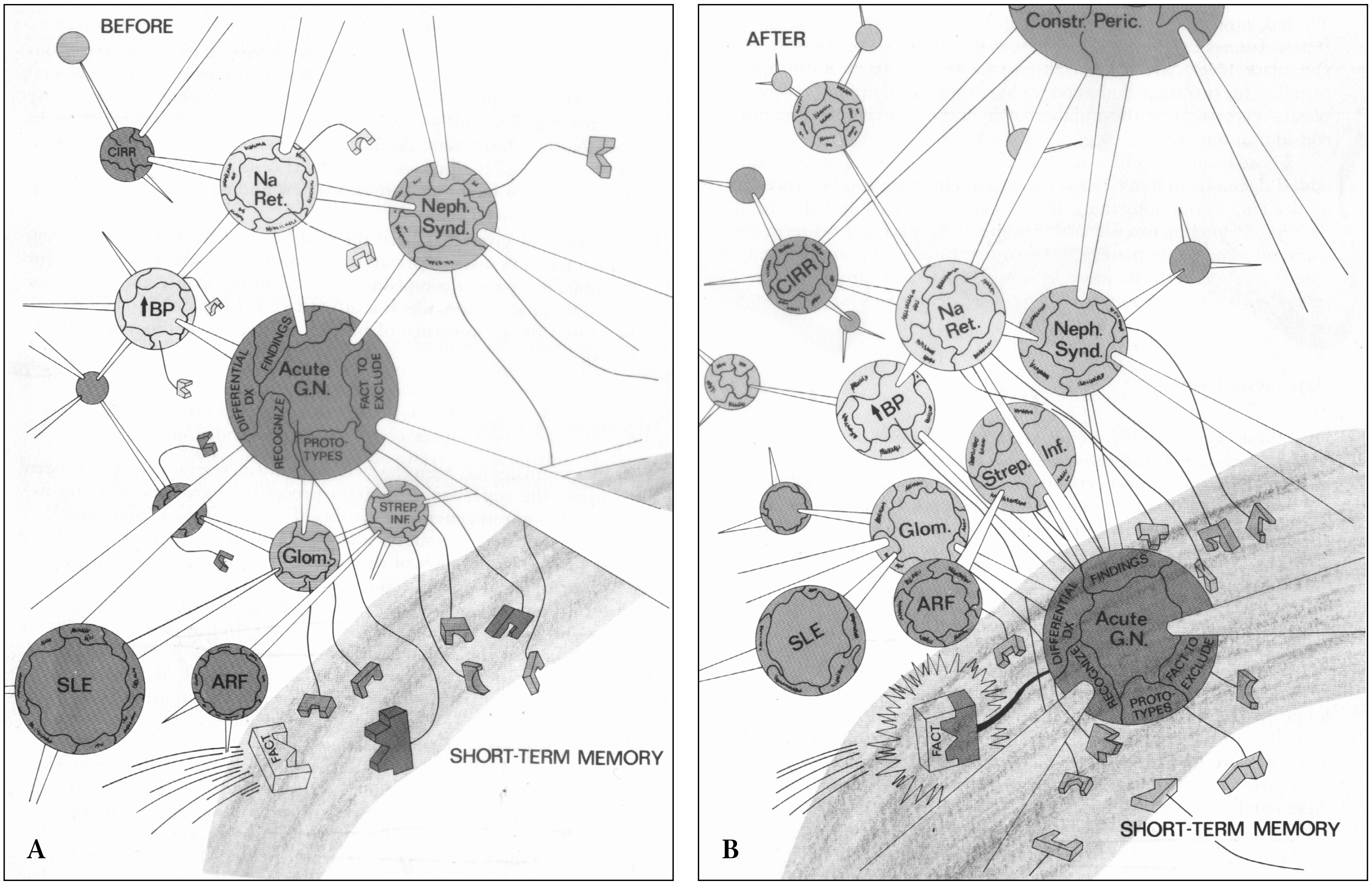


Figure 1. A Representation of the Associative Memory Required for Medical Diagnosis from the PIP *Work (Szolovits and Pauker 1976).*

A. A "fact" is searching for a connection to the network, so that the appropriate concepts can be pulled down into the short-term memory. B. A "fact" has found a matching concept and thus pulls the appropriate associated concepts into the short-term memory.

lecting findings and computing the most likely diagnoses. Since the introduction of this program, others have been introduced that are based on similar ideas, including DXPLAIN (Barnett et al. 1987) and ILIAD (Bouhaddou et al. 1995). The performance of these programs has been evaluated and compared by running them on some challenging case reports (called *clinicopathological cases,* or CPCs) such as those that appear each week in the *New England Journal of Medicine* (www.nejm.com) (Berner, Jackson, and Algina 1996; Wolfram 1995; Feldman and Barnett 1991). In most cases, the performance of the programs is comparable to expert diagnostic performance (as judged by a blinded review of diagnoses produced by both experts and the programs, or unblinded evaluation of the performance using defined performance criteria for success). The programs routinely outperform medical students and physicians in training.

In the mid- to late 1980s, Heckerman and coworkers showed that the preliminary work of De Dombal could be extended using Bayesian networks for diagnosis, in which the conditional dependencies between variables could be modeled in a somewhat natural manner (Heckerman, Horvitz, and Nathwani 1992). They also were able to recast some of the assumptions behind the other (apparently nonprobabilistic) systems (MYCIN and INTERNIST) to create a unified probabilistic "map" of the space of diagnostic algorithms (Dan and Dudeck 1992; Middleton et al. 1991; Shwe et al. 1991). So by the end of the 1980s, there was a large and distinguished literature on medical diagnosis. This literature has continued and expanded to nonmedical areas such as the diagnosis of faults in electronic circuitry and other engineering applications.

The Era of Managed Care of Chronic Disease

So what happened to the Era of Diagnosis? All of these systems were evaluated, and all of them seemed to perform near the level of human experts. Well, there were a few problems. First, physicians did not embrace these technologies. Clinical data, unlike billing data, were not routinely available in a digital form, so when you ran these programs there were these very awkward interfaces that asked you lots of questions in order to get the information necessary to do the diagnosis. Clinicians simply did not want to spend time entering data that were already written into a note using natural language. The AI in medicine community realized that they needed electronic medical records as a prerequisite infrastructural element to allow the deployment of these technologies. Thus, issues of knowledge representation, automatic data acquisition, federation of databases, and standard terminologies became quite important. The second problem for diagnostic programs was that physicians did not want help with diagnosis. Diagnosis is fun, and physicians are trained to do it well in medical school and only improve with years of practice. They did not want to give up that fun to a computer. The most significant problem, however, was that diagnosis is a actually very small part of what physicians do in the delivery of medicine. Most visits to a physician are for an existing, previously diagnosed problem. The challenge to the physician is to follow the problem and respond to its evolution intelligently. Diagnosis is a relatively rare event, probably accounting for less than 5 percent of physician time. What physicians really need is help following chronic and slowly evolving disease in 2500 patients that are seen in brief episodes but require expert interventions. So we have the era of chronic care driving AI in medicine research. This problem is compounded by an aging population with more chronic diseases.

The AI in medicine community realized that they needed electronic medical records as a prerequisite infrastructural element to allow the deployment of these technologies.

There is one other element of medicine that has changed the imperatives for AI research, and this is the emergence of new economic models for funding medicine (Selby 1997; Detsky and Naglie 1990). The traditional model has been fee for service: A physician performs a service and gets paid an agreed-upon amount. If the physician performs lots of services, the physician makes more money. The new model of medical funding is based on a standard rate per patient that is paid to a physician, regardless of the usage of services by the patient. Now, the financial incentives are reversed. If the physician provides a service, then its cost in time and resources is taken out of the pot of money that represents potential profit. Now physicians still want to treat illness, but there is now a huge incentive to deliver cost-effective, high-quality care. Systems for supporting these activities become the mandate.

One of the ways to reduce the cost of health care is to move it out of expensive hospital rooms and into outpatient clinics. So instead of intense episodes in the hospital, we have these much more frequent less intense episodes in the clinic where similar things are being done but in a more fragmented manner. The fragmentation may cause confusion as we ask physicians to track the progress of 2500 patients with periodic interactions.

One way to capture the look and feel of AI in Medicine today is to look at the contents of a recent meeting. The AI in Medicine Europe (AIME) conference was held in Grenoble in 1997 (Shahar, Miksch, and Johnson 1997). An examination of the table of contents reveals three subjects, in particular, that reflect current concerns: (1) the representation and manipulation of protocols and guidelines, (2) natural language and terminology, and (3) temporal reasoning and planning. Other areas of importance include knowledge acquisition and learning, image and signal processing, decision support, and (our old friend) diagnostic reasoning.

Protocols and guidelines have become an important way to standardize care and reduce variance. Guidelines are created by panels of physicians who assess available data and recommend treatment strategies. For example, how should a newly discovered breast lump be evaluated? The AI challenges follow directly: How do we develop robust and reusable representations of process? How do we create adaptive plans that respond to changes in available information? How do we distinguish between high-level plan recommendations and their specific local implementation? How do we modify guidelines in response to data collected during their execution? How do we model the effects of guidelines on organizations? There is an increasing interest in the representation and simulation of organization systems in order to predict the effects of interventions in medical care. One recent development in this area has been the development of a guideline interchange format (GLIF) (Ohno-Machado et al. 1998). GLIF is a syntax for specifying clinical protocols. It contains a language for representing actions, branch steps, and synchronization steps (among others) needed to specify a clinical guideline (figure 2).

Natural language and standardized terminologies remain a critical issue in medical computing. The medical goal is to create standards for communication that move away from hand-written natural language. Medicine is the only major industry still relying on hand-written documentation. How do we define formal semantics so that when we create these electronic medical records, we can populate them with clean data? What is the underlying ontology for clinical medicine? How do you map natural language into standard terminologies? How do we accommodate local and global changes to these terminologies? How do we integrate legacy databases with newer, semantically clean databases? How can we have machine-learning techniques for extracting new medical knowledge from our semantically clean databases? It is important here to mention the Unified Medical Language System (UMLS), a project at the National Library of Medicine with the goal of integrating a number of existing medical vocabularies using a common semantic structure (Bodenreider et al. 1998). The existing terminologies include those for specifying diagnoses, medical procedures, and bibliographic indexing (Cote and Robboy 1980; Slee 1978). The UMLS is based on a semantic network and has about 500,000 terms that have been classified into about 150 semantic types with specified relationships. A fragment of its semantic network is shown in figure 3.

Temporal reasoning and planning become critical in a setting where diseases are chronic, and interactions are episodic. The challenges are to integrate database and knowledge base technology with temporal inferencing capabilities. How do we actually modify medical databases so that we can do effective temporal inference with them? How can we recognize and abstract temporal trends in clinical data? Nonmonotonic reasoning becomes essential: As new data are collected, we retract old inferences and assert different ones. How do we create "smooth" models of patient state based on episodic data collection? Finally, how can we create plans for treatment over time?

My colleague at Stanford, Yuval Shahar, has done excellent work in the area of temporal abstraction and has a system that is able to automatically take a set of discrete data points and transform them into sensible intervals that can, in turn, be grouped together into even higher-level abstractions (Shahar, Miksch, and Johnson 1998) (as summarized in figure 4).

There are some other application areas within medicine that deserve mention, including *telemedicine,* how to deliver medical care at a distance using multimedia; *intensive care medicine,* with emphasis on reasoning with limited resources; and *clinical trials,* methods to automatically recognize that a patient is eligible for a trial and to enroll them.

```
Guideline the_breast_mass_guideline
{   name = "Breast Mass Guideline";
    authors = SEQUENCE 1 {"Max Borten, MD, JD";};
    eligibility_criteria = NULL;
    intention = "Evaluation of breast mass.";
    steps =
        SEQUENCE 40
    {       (Branch_Step 1);
            (Action_Step 101);
            (Action_Step 102);
            (Action_Step 103);
            (Synchronization_Step 1031);
            (Conditional_Step 104);
            (Conditional_Step 105);
Action_Step 102
{   name = "Elicit Risk Factors for Breast Cancer in Personal History";
    action = Action_Spec 102.1
        {   name = "Elicit Personal History";
            description = "Elicit Personal History";
            patient_data = SEQUENCE
{ Patient_Data 102.1
    { name = "Personal risk factors for breast cancer";
      type = "" ;
      possible_values = Sequence 0 {};
      temporal_constraint = "valid thru all time";
      didactics = SEQUENCE 0 {};};
};
            didactics = SEQUENCE 0 {};
      };

subguideline = NULL;
next_step = (Synchronization_Step 1031);
didactics = SEQUENCE 0 {};}
```

Figure 2. The Guideline Interchange Format.

This is a segment of a representation of the process of evaluating a breast mass. The language has a syntax for sequences of events, branches, and other structured metainformation about the process of evaluation.

The Era of Molecular Medicine

Although the management of chronic disease under conditions of capitated payment are likely to continue, I believe that there is an even more revolutionary set of changes coming to medicine. These changes will arise from the work being done in basic biology in determining the complete DNA sequence of both the human organism as well as most major disease-causing organisms. There is an excellent paper in the IAAI-98 proceedings by Rick Lathrop and coworkers (Lathrop et al. 1998) that is an example of the opportunities in linking molecular concepts with medical care and AI research.

Some Biology First, it is appropriate to give some background about the genome sequencing efforts. The entire development, structure, and function of an organism is specified by a sequence of four DNA letters: A, T, C, and G are the abbreviation of their chemical names

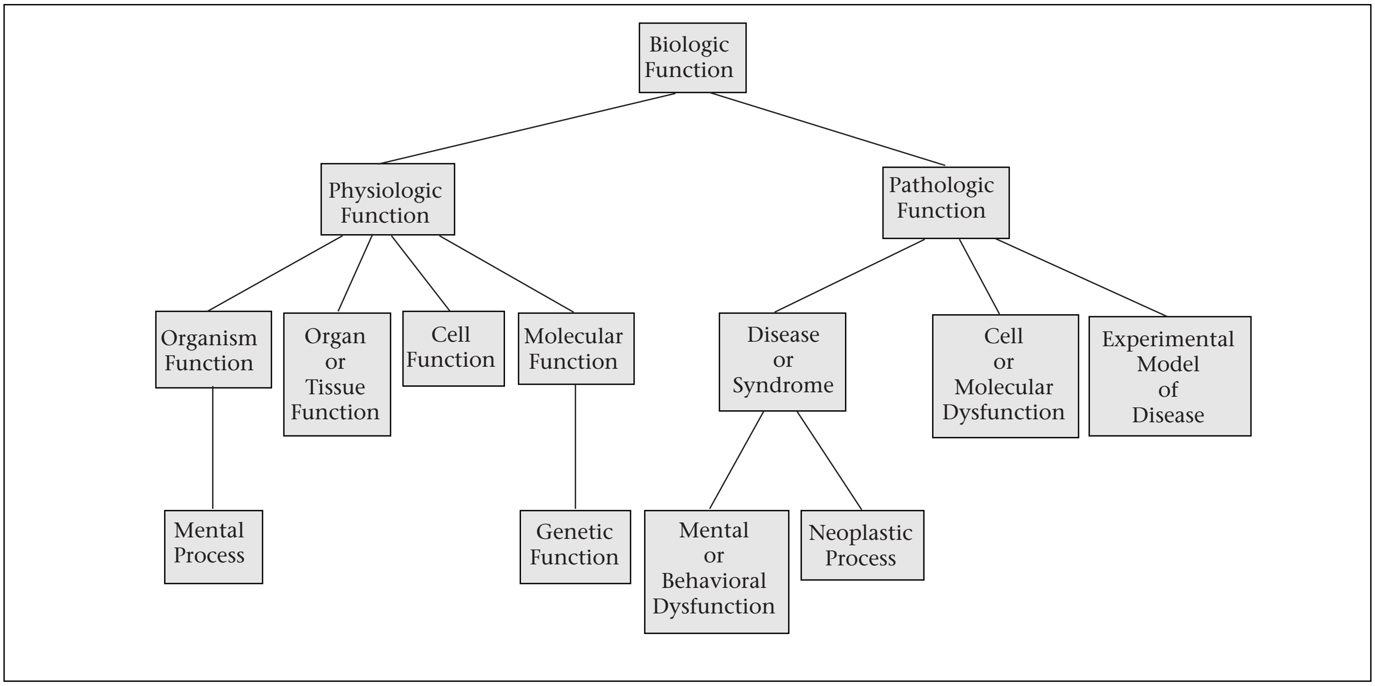


Figure 3. A Subset of the Semantic Net Created for the Unified Medical Language System (UMLS), in Which the Concept of Biological Function Is Specialized into Subsets.

The semantic network is used to organize about 500,000 concepts in the UMLS.

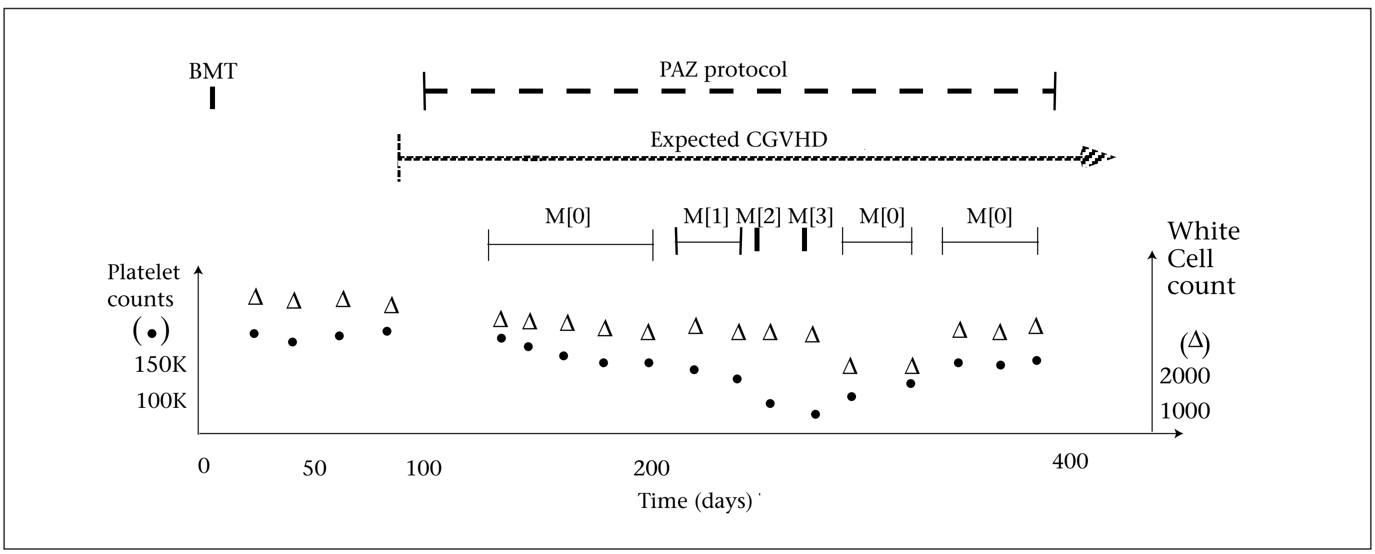


Figure 4. Temporal Abstraction in a Medical Domain.

Raw data are plotted over time at the bottom. External events and the abstractions computed from the data are plotted as intervals above the data. BMT = a bone-marrow transplantation event; PAZ = a therapy protocol for treating chronic graft-versus-host disease (CGVHD), a complication of BMT; ⊢ ⊣ = event; • = platelet counts; Δ = white cell counts; ∞ = context interval; ⊢⊣ = abstraction interval; *M*[*n*] = bone-marrow–toxicity grade *n*.

(energy). A human organism is specified by three billion letters, arranged serially, that constitute its *genome*. With 2 bits per DNA letter, it takes about 750 megabytes of data to specify a human. There are 23 *chromosomes* that divide these 3 billion into subsegments for logistical reasons, with an average length of 256 million DNA letters. *Genes* are subsequences within the sequence of 3 billion that encode for particular functions or structures that exist in your body.

There are about 100,000 genes within a human genome. More than 99.9 percent of the genome is identical for all humans. And so all the diversity of human life is contained in the 0.1 percent that is different. One of the human genes encodes a channel that allows a chloride ion to pass from the outside of a cell to its inside. This channel sometimes has a mutation that leads to the disease cystic fibrosis. An understanding of how the DNA letters differ in patients with cystic fibrosis allows biologists to begin to understand the mechanism of the disease and ways to alter its course. The cystic fibrosis gene was isolated in a relatively expensive, focused manner before the genome sequencing project was under way. The logic behind a genome project is to isolate all genes for an organism and catalog them using economies of scale. The primary genome project is the Human Genome Project, but there are also genome projects for important model organisms and organisms that cause human disease. The principal funding agencies for the genome project are the National Institutes of Health (via the National Human Genome Research Institute, www.nhgri.nih.gov/) and the Department of Energy (www.er.doe.gov/facepage/hug.htm).

Associated with the genome sequencing projects are a number of other new technologies in biology that will allow data to be collected on a large scale, never before possible. Soon it will be possible to assess the complete set of genes that are active in a cell and compare this set with the genes that are active in a diseased version of the same cell (Marshall and Hodgson 1998). Thus, we can find out which genes are active in a normal prostate cell as well as which genes are active in a prostate cancer cell. The differences are the obvious places to look for new treatments of prostate cancer. The differences may also provide new ways to make a more sensitive and specific diagnosis of prostate cancer in a patient. Finally, the differences may be used to determine the likely prognosis of a particular prostate cancer, based on its constellation of genes, and whether they are associated with an indolent or aggressive type of cancer.

Having defined all the genes in a biological system, there are incredible opportunities for information storage, retrieval, and analysis technologies. First, the epidemiology of disease will now have a molecular basis. We will track the spread of infections by identifying the unique sequences of the offending bacteria and using this as a signature to follow its spread through the population. For example, investigators have tracked the spread of tuberculosis with these technologies (Behr et al. 1998; Blower, Small, and Hopewell 1996). Second, clinical trials will have patients who are stratified by the differences and similarities in their genes. We will be able to relate particular clinical syndromes to particular treatments based on a molecular understanding of the basis of these syndromes. The diagnosis of many diseases will become a simple lookup in the genome of the patient to see which variant is present (Winters et al. 1998). We will be able to focus treatments using this information, once we have learned from the data the best drugs to use against different disease/gene variations. Finally, we will have prognostic information beyond anything currently available because we will have access to the full genetic endowment of a patient and, when relevant, the infectious pathogens causing disease. In some cases, in fact, we may know decades before a disease is evident that a patient is at high risk for that disease. At this point, it is important to mention the ethical, social, and legal issues associated with the Human Genome Project. A certain fraction of the annual genome project budget is spent on grants addressing these issues, including issues of privacy, ethical use of medical information, patients rights to information, and the like (www.nhgri.nih.gov/About_NHGRI/Der/Elsi/).

What's the status of the genome sequencing projects? Although this is not AI per se, it is useful to get a feeling for the amounts and types of data that are being generated. Consider the GENBANK database of DNA sequences (Benson et al. 1998). A recent release of that database contained 1.6 billion bases. (Remember, there are 3 billion bases in a human). However, this database contains DNA sequences from all organisms, and not just humans. Figure 5 shows the growth in the size of this database since its inception in 1982. All these data are available on the World Wide Web, and one of the remarkable aspects of the explosion of biological data is the ease in which it can be accessed, and so it becomes something of a playground for information scientists who need to test ideas and theories. Table 1 shows the ranking of species in the DNA databases by the values of sequenced bases for each sequence. For example, we have roughly 700 million bases of human genome sequence. The human genome is currently scheduled to be completed around 2003. Other organisms include important laboratory test organisms (for example, mouse, rat, or fruit fly) or human pathogens (for example, the HIV virus, malaria, syphilis, or tuberculosis). One of the most exciting challenges that arises as we learn the

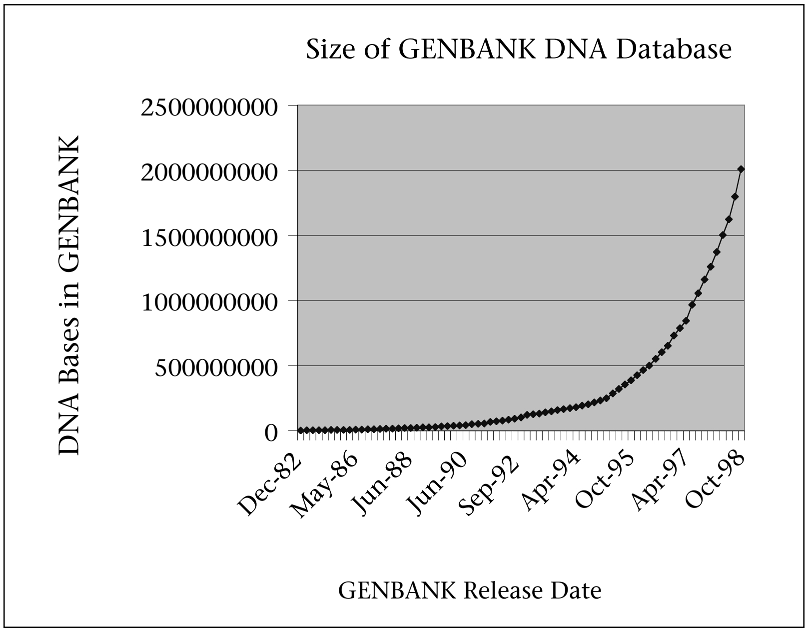


Figure 5. The Rapid Increase in Genetic Data Contained in a Major DNA Sequence Database.

complete genetic background of bacteria is to develop comparative methods for understanding how differences in genetic endowment create differences in preference for mode of infection, type of infection, and virulence. Table 2 shows some organisms whose genomes are completely known.

An international society (The International Society for Computational Biology, www.iscb.org/) has been formed to create a community for researchers in biocomputing. In many ways, it is a spin-off from the AAAI community. A conference entitled Intelligent Systems for Molecular Biology was first held in 1993 in association with AAAI (//ismb99.gmd.de/). It subsequently was disconnected from AAAI for scheduling reasons but has been remarkably successful. The proceeds from this meeting provided the core funds to spawn the society. The name of the field (computational biology, bioinformatics, biocomputing, intelligent systems for molecular biology) has been the matter of some debate and often reflects the disciplinary training of the debate participants (Altman 1998).

AI Contributions to Molecular Medicine I would like to briefly review some of the successes in the AI and molecular biology community. In many cases, existing technology has been transferred effectively, or new variants have been created in response to the needs of the field.

First, Hidden Markov Models, developed originally for natural language processing, have become a powerful tool for analyzing linear sequences of the DNA and of protein molecules (Durbin et al. 1998).

Second, the technologies for defining ontologies, terminologies, and their logical relationships have been used to create formal theories for areas within biology (Schulze-Kremer 1998).

Third, genetic algorithms and genetic programming have been used to create solutions that are in some cases superior to solutions created by hand (Koza 1994).

Fourth, neural networks have been used, as they have in many other fields, to achieve classification performance that is often quite impressive. The work in predicting aspects of three-dimensional structure from sequence information alone has received considerable attention (Rost and Sander 1994).

Fifth, unsupervised cluster analysis of biological sequences and structures, including Bayesian approaches, has been successful in creating sensible categories within biological data sets (States, Harris, and Hunter 1993).

Sixth, case-based reasoning has become very important in areas that are still data poor. For example, information about the three-dimensional structure of biological molecules is still lagging behind the associated DNA sequence information. Thus, of the 100,000 proteins in a human, we only know the structure of about 700 of them. These examples represent valuable "cases" that are constantly being used to extrapolate new information about the remaining 99,000 proteins (Jones 1997).

Seventh, knowledge representation techniques have been used to represent the entire set of metabolic capabilities of the bacteria *Escherichia coli*. The resulting ECOCYC knowledge base has been used to infer metabolic capabilities and compare these capabilities across organisms (Karp et al. 1998).

Eighth, new knowledge representation and digital library techniques have been used to represent the complete literature of a subdiscipline within molecular biology using ontologies for biological data, biological structure, and scientific publishing (Chen, Felciano, and Altman 1997). We have created a collaborative resource, RIBOWEB, that allows scientists to interact with these data and compute with it over the web.

Ninth, intelligent agents are being designed to assist biologists in understanding and mining the data that are accumulating from the new high-throughput biological experiments. The National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/) is the clearing house for many sources of useful biological data. Their collection includes data about DNA

Genes	DNA Bases	Species (Common Name)
1573906	1000128755	Homo sapiens (human)
403552	191558011	Mus musculus (mouse)
76540	142527757	Caenorhabditis elegans (soil nematode)
71079	78218600	Arabidopsis thaliana
56199	63799526	Drosophila melanogaster (fruit fly)
10581	28685645	Saccharomyces cerevisiae (baker's yeast)
45849	28537572	Rattus norvegicus (rat)
4953	18023376	Escherichia coli
41866	17672014	Rattus sp.
32190	16498151	Fugu rubripes (puffer fish)
36345	16196521	Oryza sativa (rice)
9610	12068959	Schizosaccharomyces pombe
25383	11280798	Human immunodeficiency virus type 1 (HIV)
1094	9985595	Bacillus subtilis
4734	7009140	Plasmodium falciparum (malaria)
16688	6331052	Brugia malayi (filariasis)
5379	5922144	Gallus gallus (chicken)
685	5711838	Mycobacterium tuberculosis (tuberculosis)
5136	4648144	Bos taurus (cow)
10847	4413291	Toxoplasma gondii (toxoplasmosis)

Table 1. Number of Genes and Total Bases of DNA Sequenced for Various Organisms as of 10/98.

Aquifex aeolicus (bacteria that grows at 85° to 95° C!)
Archaeoglobus fulgidus (bacteria that metabolizes sulfur, lives at high temperatures)
Bacillus subtilis (ubiquitous soil bacteria)
Borrelia burgdorferi (causes Lyme Disease)
Chlamydia trachomatis (causes blindness in developing countries)
Escherichia coli (can cause urinary tract infections, dysentery)
Haemophilus influenzae (causes upper respiratory infections)
Methanobacterium thermoautotrophicum (bacteria that produces methane, lives at 70° C)
Helicobacter pylori (causes ulcers, maybe cancer)
Methanococcus jannaschii (bacteria that produces methane)
Mycobacterium tuberculosis (causes tuberculosis)
Mycoplasma genitalium (smallest genome of known independent organisms)
Mycoplasma pneumoniae (causes "walking pneumonia")
Pyrococus horikoshii (grows best at 98° C!)
Saccharomyces cerevisiae (baker's yeast)
Treponema pallidum (causes syphillis)

Table 2. Some Completed, Fully Sequenced Genomes.

sequences, human genetic diseases, protein and nucleic acid structures, the biomedical literature, and other important sources. The good news is that this information is easily available. The bad news is that it is not yet integrated in a manner that allows rapid discovery and inferencing. Thus, a biologist with lots of time can eventually find the answer to a question using these databases, but they have not yet been distilled to the point where a physician in a clinical practice can use the data to guide clinical decision making. The intelligent integration of biological information thus becomes one of the major bottlenecks in progress for information processing (Markowitz and Ritter 1995).

Ten Challenges I want to end my presentation with the 10 grand challenges to medicine I have previously proposed (Altman 1997). These can be divided into infrastructure, performance, and evaluation goals and are summarized here:

Infrastructure Challenges

1. We need an electronic medical record based on semantically clean knowledge representation techniques.

2. We need automated capture of clinical data, from the speech, natural language, or structured entry, in order to provide the data required to move forward.

3. We need computable representations of the literature. Both clinical and basic biology data should be structured and universally accessible for automated data analysis.

Performance Challenges

4. We still need to do automated diagnosis. Despite the passing of its era, it is still worth understanding because there are times it is useful.

5. We need automated decision support for providers who interact with patients episodically and need help in making decisions about the treatment trajectory.

6. We need systems for improving access to information and explanation for patients.

7. We need systems to provide and document continuing education for physicians.

Evaluation Challenges

8. We need demonstrations of the cost-effectiveness of advanced information technology.

9. We need to create new medical knowledge with machine-learning and/or data-mining techniques. Having established the data infrastructure for clinical data and biological data, there will be unprecedented opportunities for gaining new knowledge.

10. Finally, we need to ensure that there is equitable access to these technologies across patient and provider populations.

Conclusions

The Era of Diagnosis got things rolling and created excitement as existing inferencing strategies were tested in the real-world application domain of medicine. The current Era of Chronic Disease and Managed Care has changed the focus of our efforts. The coming Era of Molecular Medicine contains challenges that can keep information technologists busy for decades.

Acknowledgments

RBA is supported by grants from NIH (LM-05652, LM-06422), NSF (DBI-9600637, agreement ACI-9619020), a Faculty Scholar grant from IBM, and a gift from Sun Microsystems Inc.

References

Altman, R. 1998. A Curriculum for Bioinformatics: The Time Is Ripe. *Bioinformatics* 14(8): 549–550.

Altman, R. B. 1997. Informatics in the Care of Patients: Ten Notable Challenges [see comments]. *Western Journal of Medicine* 166(2): 118–122.

Barnett, G. O.; Cimino, J. J.; Hupp, J. A.; and Hoffer, E. P. 1987. DXPLAIN. An Evolving Diagnostic Decision-Support System. *Journal of the American Medical Association* 258(1): 67–74.

Behr, M. A.; Hopewell, P. C.; Paz, E. A.; Kawamura, L. M.; Schecter, G. F.; and Small, P. M. 1998. Predictive Value of Contact Investigation for Identifying Recent Transmission of Mycobacterium Tuberculosis. *American Journal of Respiratory and Critical Care Medicine* 158(2): 465-469.

Benson, D. A.; Boguski, M. S.; Lipman, D. J.; Ostell, J.; and Ouellette, B. F. 1998. GENBANK. *Nucleic Acids Research* 26(1): 1–7.

Berner, E. S.; Jackson, J. R.; and Algina, J. 1996. Relationships among Performance Scores of Four Diagnostic Decision Support Systems. *Journal of the American Medical Informatics Association* 3(3): 208–215.

Blower, S. M.; Small, P. M.; and Hopewell, P. C. 1996. Control Strategies for Tuberculosis Epidemics: New Models for Old Problems. *Science* 273(5274): 497–500.

Bodenreider, O.; Burgun, A.; Botti, G.; Fieschi, M.; LeBeux, P.; and Kohler, F. 1998. Evaluation of the Unified Medical Language System as a Medical Knowledge Source. *Journal of the American Medical Informatics Association* 5(1): 76–87.

Bouhaddou, O.; Lambert, J. G.; and Morgan, G. E. 1995. Iliad and the Medical House Call: Evaluating the Impact of Common Sense Knowledge on the Diagnostic Accuracy of a Medical Expert System. In Proceedings of the Nineteenth Annual Symposium on Computer Applications in Medical Care, 742–746. Salt Lake City, Utah: Applied Medical Informatics.

Buchanan, B., and Shortliffe, E., eds. 1984. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristics Programming Project*. Menlo Park, Calif.: Addison-Wesley.

Chen, R. O.; Felciano, R.; and Altman, R. B. 1997. RIBOWEB: Linking Structural Computations to a Knowledge Base of Published Experimental Data. *Intelligent Systems for Molecular Biology* 5:84–87.

Cote, R. A., and Robboy, S. 1980. Progress in Medical Information Management. Systematized Nomenclature of Medicine (SNOMED). *Journal of the American Medical Association* 243(8): 756–762.

Dan, Q., and Dudeck, J. 1992. Certainty Factor Theory and Its Implementation in a Medical Expert System Shell. *Medical Informatics (London)* 17(2): 87–103.

de Dombal, F. T.; Leaper, D. J.; Staniland, J. R.; McCann, A. P.; and Horrocks, J. C. 1972. Computer-Aided Diagnosis of Acute Abdominal Pain. *British Medical Journal* 2(5804): 9–13.

Department of Energy. 1996. Human Genome Project. To Know Ourselves. Available at www.ornl.gov/hgmis/tko/. Also available at www.ornl.gov/TechResources/Human_Genome/publicat/primer/intro.html.

Detsky, A. S., and Naglie, I. G. 1990. A Clinician's Guide to Cost-Effectiveness Analysis [see comments]. *Annals of Internal Medicine* 113(2): 147–154.

Durbin, R.; Eddy, S.; Krogh, A.; and Mitchison, G. 1998. *Biological Sequence Analysis*. Cambridge, U.K.: Cambridge University Press.

Feigenbaum, E. A. 1984. Knowledge Engineering: The Applied Side of Artificial Intelligence. *Annals of the New York Academy of Sciences* 246:91–107.

Feldman, M. J., and Barnett, G. O. 1991. An Approach to Evaluating the Accuracy of DXPLAIN. *Computer Methods and Programs in Biomedicine* 35(4): 261–266.

Heckerman, D. E.; Horvitz, E. J.; and Nathwani, B. N. 1992. Toward Normative Expert Systems: Part I. The PATHFINDER Project. *Methods of Information in Medicine* 31(2): 90–105.

Jones, D. T. 1997. Progress in Protein Structure Prediction. *Current Opinion in Structural Biology* 7(3): 377–387.

Karp, P. D.; Riley, M.; Paley, S. M.; Pellegrini-Toole, A.; and Krummenacker, M. 1998. ECOCYC: Encyclopedia of *Escherichia coli* Genes and Metabolism. *Nucleic Acids Research* 26(1): 50–53.

Koza, J. R. 1994. Evolution of a Computer Program for Classifying Protein Segments as Transmembrane Domains Using Genetic Programming. *Intelligent Systems for Molecular Biology* 2:244–252.

Kulikowski, C., and Weiss, S. 1982. Representation of Expert Knowledge for Consultation: The CASNET and EXPERT Projects. In *Artificial Intelligence in Medicine,* ed. P. Szolovits. Boulder, Colo.: Westview.

Lathrop, R. H.; Steffen, N. R.; Raphael, M.; Deeds-Rubin, S.; Pazzani, M. J.; Cimoch, P. J.; See, D. M.; and Tilles, J. G. 1998. Knowledge-Based Avoidance of Drug-Resistant HIV Mutants. In Proceedings of the Tenth Annual Innovative Applications of Artificial Intelligence Conference, 1071–1078. Menlo Park, Calif.: American Association for Artificial Intelligence.

Ledley, R. S., and Lusted, L. B. 1959. Reasoning Foundation of Medical Diagnosis: Symbolic Logic, Probability, and Value Theory Aid Our Understanding of How Physicians Reason. *Science* 130(9): 9–12.

Markowitz, V. M., and Ritter, O. 1995. Characterizing Heterogeneous Molecular Biology Database Systems. *Journal of Computational Biol*ogy 2(4): 547–556.

Marshall, A., and Hodgson, J. 1998. DNA Chips: An Array of Possibilities. *Nature Biotechnology* 16(1): 27–31.

Middleton, B.; Shwe, M. A.; Heckerman, D. E.; Henrion, M.; Horvitz, E. J.; Lehmann, H. P.; and Cooper, G. F. 1991. Probabilistic Diagnosis Using a Reformulation of the INTERNIST-1/QMR Knowledge Base. II. Evaluation of Diagnostic Performance. *Methods of Information in Medicine* 30(4): 256–267.

Miller, R.; Masarie, F.; and Myers, J. 1986. Quick Medical Reference (QMR) for Diagnostic Assistance. *MD Computing* 3(5): 34–48.

Miller, R. A.; Pople, H. E.; and Myers, J. D. 1982. INTERNIST-1: An Experimental Computer-Based Diagnostic Consultant for General Internal Medicine. *New England Journal of Medicine* 307(8): 468–476.

Ohno-Machado, L.; Gennari, J. H.; Murphy, S. N.; Jain, N. L.; Tu, S. W.; Oliver, D. E.; Pattison-Gordon, E.; Greenes, R. A.; Shortliffe, E. H.; and Barnett, G. O. 1998. The Guideline Interchange Format: A Model for Representing Guidelines. *Journal of the American Medical Informatics Association* 5(4): 357–372.

Rost, B., and Sander, C. 1994. Combining Evolutionary Information and Neural Networks to Predict Protein Secondary Structure. *Proteins* 19(1): 55–72.

Schulze-Kremer, S. 1998. Ontologies for Molecular Biology. Paper presented at the Pacific Symposium on Biocomputing, 4–9 January, Maui, Hawaii.

Selby, J. V. 1997. Linking Automated Databases for Research in Managed Care Settings. *Annals of Internal Medicine* 127(8 Pt 2): 719–724.

Shahar, Y.; Miksch, S.; and Johnson, P. 1997. A Task-Specific Ontology for the Application and Critiquing of Time-Oriented Clinical Guidelines. In *Proceedings of the Sixth Conference on Artificial Intelligence in Medicine, ed.* E. Keravnou, 51–61. New York: Springer-Verlag.

Shahar, Y.; Miksch, S.; and Johnson, P. 1998. The ASGAARD Project: A Task-Specific Framework for the Application and Critiquing of Time-Oriented Clinical Guidelines [in process citation]. *Artificial Intelligence in Medicine* 14(1–2): 29–51.

Shwe, M. A.; Middleton, B.; Heckerman, D. E.; Henrion, M.; Horvitz, E. J.; Lehmann, H. P.; and Cooper, G. F. 1991. Probabilistic Diagnosis Using a Reformulation of the INTERNIST-1/QMR Knowledge Base. I. The Probabilistic Model and Inference Algorithms. *Methods of Information in Medicine* 30(4): 241–255.

Slee, V. N. 1978. The International Classification of Diseases: Ninth Revision (ICD-9) [Editorial]. *Annals of Internal Medicine* 88(3): 424–426.

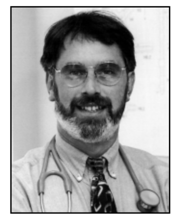
States, D. J.; Harris, N. L.; and Hunter, L. 1993. Computationally Efficient Cluster Representation in Molecular Sequence Megaclassification. *Intelligent Systems for Molecular Biology* 1:387–94.

Szolovits, P., and Pauker, S. G. 1976. Research on a Medical Consultation System for Taking the Present Illness. In Proceedings of the Third Illinois Conference on Medical Information Systems, 299–320. Chicago: University of Illinois at Chicago.

Tufo, H. M.; Bouchard, R. E.; Rubin, A. S.; Twitchell, J. C.; VanBuren, H. C.; Weed, L. B.; and Rothwell, M. 1977. Problem-Oriented Approach to Practice. I. Economic Impact. *Journal of the American Medical Association* 238(5): 414–417.

Winters, M. A.; Coolley, K. L.; Girard, Y. A.; Levee, D. J.; Hamdan, H.; Shafer, R. W.; Katzenstein, D. A.; and Merigan, T. C. 1998. A 6-Basepair Insert in the Reverse Transcriptase Gene of Human Immunodeficiency Virus Type 1 Confers Resistance to Multiple Nucleoside Inhibitors [in process citation]. *Journal of Clinical Investigation* 102(10): 1769–1775.

Wolfram, D. A. 1995. An Appraisal of INTERNIST-1. *Artificial Intelligence in Medicine* 7(2): 93–116.



Russ Altman is an assistant professor of medicine (medical informatics, and computer science by courtesy) at Stanford University. He is interested in the application of advanced computational techniques to problems in biomedicine and molecular biology. He has an AB from Harvard University in biochemistry and molecular biology, a Ph.D. from Stanford University in medical information sciences, and an MD from Stanford. He is a member of the Association of Computing Machinery, the American College of Physicians, the American Association for Artificial Intelligence, and the Institute for Electrical and Electronics Engineers and is a founding member of the International Society for Computational Biology. He is a fellow of the American College of Medical Informatics. He is a past recipient of a U.S. Presidential Early Career Award for Scientists and Engineers and a National Science Foundation CAREER Award. His web page is www.smi.stanford.edu/people/altman/. His e-mail is russ.altman@stanford.edu.