

Virtual Mentor

American Medical Association Journal of Ethics
March 2001, Volume 3, Number 3: 76-82.

STATE OF THE ART AND SCIENCE

Toward Striking a Balance in Bioinformatics

Kenneth Goodman, PhD

One could make a very strong argument that information technology and genetics are the sciences that will have the greatest effect on 21st century healthcare. Although it is already clear that informatics is profoundly reshaping the health professions, we are only beginning to come to terms with the extraordinary risks and potential benefits of progress in the human genome sciences.

We need to determine what to make of the intersection of health informatics and genomics. Add in ethical and social issues, and we confront one of the greatest intellectual and practical challenges in the history of science.

Bioinformatics--the use of information technology to acquire, store, manage, share, analyze, represent, and transmit genetic data--has blossomed in the past several years. The term is most often used by scientists who sequence and otherwise analyze the genomes of humans and other species with computer technology. If we like, we can stipulate that bioinformatics also includes applications as pedestrian as using a personal computer to store the results of genetic tests ("Patient X has the BRCA1 gene"), as well as using intelligent machines to link physiological traits with a database in order to diagnose genetic maladies, predict clinical correlations, conduct research, and so forth.

Although bioinformatics raises many issues for human subjects research, we will confine ourselves here to more clinical concerns and group them under the label "clinical bioinformatics." (It is well to note, though, that once human genetic information is stored on a computer, it is much easier to study; in some cases the distinction between clinical and research issues will narrow dramatically.) Let us organize the ethical and social issues raised by clinical bioinformatics into the following categories: (1) accuracy and error, (2) appropriate uses and users of digitized genetic information, and (3) privacy and confidentiality.

Accuracy and Error

Health informatics has taught us that accuracy and error avoidance raise ethical issues that are often related to evolving standards of care. If there are emerging or established standards for database management, for instance, then a system that relies on a database will be more or less useful, reliable, and safe, depending upon whether or not the database is appropriately maintained, tested, augmented, and so on. The reason to link error and ethics is that errors, however unintentional, can

produce harm. Determining whether a harm constitutes a wrong is one of the main challenges of ethics. Our specific challenge here is to nurture the growth of an exciting new science while simultaneously ensuring that patients are not harmed or wronged.

Several current and future issues related to the accuracy of bioinformatics systems follow:

Risks to persons. To the extent that we can expect more and more frequent computer-aided discoveries of the genetic loci of human diseases, errors can pose or increase risks to public health and even the wellbeing of individuals. Patients may also be at risk when computers are used to predict the expression of future genetic maladies. The risks may be psychological and will likely vary depending on whether there is a treatment or cure for a given malady. The role of genetic counselors will loom large here.

Recanted linkage studies. Preliminary or unreplicated linkage studies are sometimes recanted or re-evaluated. Erroneous linkage analyses can throw colleagues off the track and, perhaps more importantly, cause unnecessary psychological trauma for individuals who fear they may be affected. In the case of purported linkages that correlate with race or ethnicity, there is the added risk of producing social stigma—perhaps especially in the case of neurogenetics and psychiatric genetics.

Meta-analysis. It is exciting to observe the emergence of meta-analysis in genomics. In this research technique, the results of previous studies are aggregated and reanalyzed by statistical software with the aim of achieving statistical significance or adequate sample sizes. This technique raises ethical issues by virtue, in part, of doubts about the quality of included data and the validity of inferences based on diversity of data. These doubts are important when meta-analytic results are applied to patient care.

Decision support. Even though diagnostic and decision support systems are well known to raise ethical issues in clinical medicine, there is as yet no critical analysis of decision support for genetic diagnoses in which, for example, clinical information, photographic material, pedigree, and gene localization data are analyzed by computers. The growth of genomic data bases and the increasing availability of genetic information at the clinical level suggest that decision support systems are a ripe source for ethical and social inquiry.

Appropriate Uses and Users

Questions concerning who should use clinical information systems and in what contexts have been shown to raise interesting and important ethical issues; we should expect that genetic data processing will elicit related concerns and pose new problems.

For instance, suppose a physician or nurse begins including genetic data in patient charts, uses those data to predict the likelihood of clinical manifestations and correlations, and employs those analyses to refer patients to genetic counselors. The first question is basic: Was this novel use undertaken with the patient's consent? Because genetic information can frighten or alarm patients in ways that other health and medical data do not, we need to ask whether the patients knew that genetic data was being gathered and stored for clinical purposes. In the absence of a treatment or cure for a particular genetic malady, it is not unreasonable for a patient to prefer not to know a genetic diagnosis or prognosis. Consent seems to be a crucial gate through which the physician or nurse must pass before using these data "for the patient's sake." The weight of valid or informed consent seems greater here - that is, in the area of genetics - than for more familiar kinds of clinical decision support.

To raise another concern, suppose that individuals' genetic data were being collected by governments, managed care organizations, or other third-party payers with the goal of shaping or adjusting risk pools or coverage eligibility. The difference between evidence-based actuarial calculations and discrimination can be very slight, indeed. To the extent that computers are used for these tasks, it will be essential for individuals, institutions, and society to decide on ethically optimized strategies for clinical bioinformatics applications.

Now we must ask who should use a genetic diagnostic or prognostic system. For example, does the possibility that determining health benefits raises a problem imply that bioinformatics tools should never be used by certain entities?

Consider that individual physicians, nurses, genetic counselors, or psychologists might use computer systems not only to improve patient care but also for less worthy purposes. Does it follow that certain users - in addition to uses - might be problematic?

One way to approach the question is to ask whether the user is employing a computer in a task not normally within his or her competence. In other words, if you are unable or untrained to perform certain tasks without a computer, then it is inappropriate to suppose that the computer can somehow imbue you with those skills. For instance, if a physician or nurse does not normally render genetic diagnoses, it is unwise to suppose that she or he acquires competence via the machine.

In fact, it is more than unwise - it is a patent mistake. Computers can improve our skills at many tasks but rarely, if ever, give us new professional skills or abilities. Therefore, an appropriate use of a genetic decision support system, for instance, will be to assist adequately trained professionals, not to replace them or to bring them "up to speed" in domains in which they lack basic skills.

This point must be clearly understood: Computers can be outstanding educational tools in bioinformatics as elsewhere, but there is a difference between acquiring a

skill and presuming its existence. We have learned from "ordinary" clinical computing that humans practice medicine and nursing but computers do not. This is a lesson well worth applying to bioinformatics.

Privacy and Confidentiality

The electronic storage of genetic information replicates a tension already familiar in health informatics: the tension between (1) the need for appropriate or authorized access to personal information, and (2) the need to prevent inappropriate or unauthorized access. Striking a balance between these two imperatives is an exciting but sometimes vexing challenge.

Privacy and confidentiality are potentially threatened when individual genetic data are maintained or transmitted using computers. The threats include bias and discrimination, personal stigma (as opposed to population or subgroup stigma), psychological stress, and tensions within families, among other risks. The difficulties posed by expectations of privacy and confidentiality are well explored in regard to the electronic patient record, but we do not yet know whether the inclusion of genetic data adds to or alters those difficulties.

Specifically, our objective is to determine if and in what way bioinformatics raises ethical issues that are distinct from ethics and genetics and, depending on what is found, either to adapt existing conceptual and pedagogic tools or provide new ones.

The key means by which we plan to meet these objectives are the successful development of ethically optimized guidelines (for organizations that maintain data bases, for IRBs, etc.) and model curricula in ethics and bioinformatics (for students and professionals).

Striking a Balance

The need for organizational policies, best-practice standards, and/or guidelines is widespread in the human sciences. Because the thrust of the proposed research is at the intersection of three vast areas of inquiry and practice - genetics, computing, and ethics - the challenge we face is extraordinary: Guidelines and standards often fail because they are either so broad or simplistic that they cannot adequately guide behavior, or are so specific or detailed that they are too inflexible to be useful in diverse and unexpected cases. There is therefore a need to strike a balance between these two shortcomings. Striking that balance would provide a very useful tool for organizations.

As to educational materials, it is worth observing that research ethics curricula usually overlook issues in bioinformatics. If we are correct in anticipating that the future of genetic research will be inextricably linked to information-processing technologies, then this oversight is, or will be, quite serious. Indeed, we may well conclude from our inquiry that there are larger nets to cast and that we should extend our emphasis on genetics to include all biology and medicine and the changes mediated by information technology.

Case 1: Genetic Information in Centralized Databases

Individual genetic information is increasingly stored in public, private and governmental health databases. The databases are or could be used for clinical practice, epidemiologic research, pharmaceutical investigations and other purposes. The World Medical Association (WMA) is drafting guidelines for the use of genetic information in such databases.

According to WMA Chair Anders Milton, "The public is rightly concerned about whether their right to privacy and confidentiality is threatened by these databases and whether information about them as individuals could be misused. Centralized health databases can make a tremendous contribution to the improvement of health. But the public's right to privacy and consent are essential to the trust and integrity of the patient/physician relationship and we must ensure that these rights are properly protected. Any guidelines must address the issues of privacy, consent, individual access and accountability."

Questions for Discussion:

1. Genetic information has been included for years' without clear regulation or rule - in electronic databases. If there were guidelines, should they apply to information collected retrospectively, prospectively, or both? Might it ever be too late for guidelines?
2. To what extent can informed or valid consent requirements be loosened if genetic information is anonymous or not linked to identifiable persons? How should the problem of racial or ethnic stigma be addressed in any guidelines?
3. Does - or how does - the purpose of a database have ethical consequences for its use? That is, does it matter if a database is (i) owned by a for-profit corporation, (ii) public health organization, (iii) government, etc.?

Case 2: Web-mediated Paternity Testing

Paternity testing has always raised difficult questions. Now, though, Web sites <http://www.dnanow.com/>, <http://www.genetestlab.com/>, and <http://dnatesting.com/> offer a chance to test a child's paternity - without the consent or knowledge of a woman or her child, a man or his (putative) child - or, indeed, a child, perhaps as adult, with a hair sample from whom what one site terms the "alleged father."

The companies offer genetic analysis of hair or buccal mucosa samples to determine paternity. A man might therefore complete an on-line form, submit a sample of his and a child's hair and, in a few days, learn via e-mail if he is the father. A woman unsure of which of several potential candidates is the father of her child might obtain a hair sample or samples and submit them along with her child's. If anyone has questions about the process, one firm offers the following: "For instant answers or advice, chat to a DNA expert online!"

A British Department of Health spokesperson was quoted in one report as saying of one vendor, "There is nothing illegal about the Web site. We are aware of concerns raised by advances in DNA testing and we are in the process of drawing up a voluntary code of practice on the way companies work."

Questions for Discussion:

1. The standard of care in genetic counseling generally requires pre- and post-test counseling for individuals and couples. Paternity is among the issues raised during such sessions, and couples or individuals are often warned that they might acquire information that could significantly alter or damage relationships. Web-based paternity testing diminishes or eliminates counseling and/or such disclosure, or provides these services via e-mail. Should Web-based testing be required to hew to standards elsewhere in genetic testing? How so, given that firms providing these services operate across international boundaries?
2. What kinds of caveats or disclaimers are appropriate for such Web-based services?
3. In the absence of laws that might regulate such Web-based testing, could voluntary policies or guidelines have an adequate effect?

Case 3: Errors in Genetic Databases

A molecular pathologist in California thought he had finally identified a gene he had been working on. He submitted it to GenBank, the public database that contains every published DNA sequence. GenBank can identify similar genes and so is useful in trying to infer a new gene's function. But the data base turned up more than 100 matches - a sign that something had gone terribly wrong. Indeed, each of those matches had in common a sequence that had been introduced by the commercial kit he had used to clone his gene.

The pathologist says he found the error "entirely by accident" and that "there's a huge number of public sequences that are incorrect."

Questions for Discussion:

1. We know well that databases are dependent on those who build and maintain them, and that database design, construction and maintenance raise ethical issues. What special issues are raised when databases store biological or health information?
2. Who should be responsible for errors in very large and/or complex databases? What is to be done when errors are perpetuated? An error might be caught or missed, have no effect or have a tragic effect - independently of the action that introduced the error. Does the consequence of the error have moral significance?
3. Is database size or complexity an adequate excuse for errors?

References

1. Annas GJ, Glantz LH, Roche PA. Drafting the Genetic Privacy Act: science, policy, and practical considerations. *J Law Med Ethics*. 1995;23(4):360-366.
2. Annas GJ. Privacy rules for DNA databanks: protecting coded 'future diaries.' *JAMA*. 1993;270(19):2346-2350.
3. Arena JF, Lubs HA. Computerized approach to X-linked mental retardation syndromes. *Am J Med Genet*. 1991;38(2-3):190-199.
4. Boguski MS. Hunting for genes in computer data bases. *N Engl J Med*. 1995;333(10):645-647.
5. Bork P, Koonin EV. Predicting functions from protein sequences: where are the bottlenecks? *Nat Genet*. 1998;18(4):313-318
6. Dickson D. Open access to sequence data "will boost hunt for breast cancer gene." *Nature*. 1995;378(6556):425.
7. Goodman KW. Ethics, genomics and information retrieval. *Comput Biol Med*. 1996;26(3):223-229.
8. Gostin, LO. Genetic privacy. *J Law Med Ethics*. 1995;23(4):320-330.
9. Hilgartner S. Biomolecular databases. *Science Communication*. 1995;17:240-63.
10. Waldrop MM. On-line archives let biologists interrogate the genome. *Science*. 1995;269(5229):1356-1358.
11. Wolf CR, Smith G, Smith R. Pharmacogenetics. *BMJ*. 2000;320(7240):987-990.

Kenneth Goodman, PhD is founder and director of the Forum for Bioethics and Philosophy at the University of Miami in Florida and adjunct professor in the UM Departments of Medicine and Philosophy. He has written numerous books and articles on the matter of medicine, computing, and ethics and is a contributing editor to *Physician and Computers* and *MDComputing*. Dr. Goodman has a master's degree in theoretical linguistics and a doctorate in philosophy.

The viewpoints expressed on this site are those of the authors and do not necessarily reflect the views and policies of the AMA.

Copyright 2001 American Medical Association. All rights reserved.