

The Needs for Database Research and for Privacy Collide

Large medical databases are a critical resource for research that has an important impact on public health, but at the same time, use of these databases raises important issues with respect to patient confidentiality. The collision between the methods of research—even research with lofty public health goals—and the protection of individuals has been increasingly a matter of public debate. Recently, for example, the National Bioethics Advisory Commission spent a year focused on the ethics of research involving individuals with impaired decision-making capacity. Workshops have been held (and others planned) on the ethics of placebo controls in clinical trials of psychotropic medications, and the National Institute of Mental Health has developed new procedures to improve the protection of potentially vulnerable populations who volunteer for research. Population-based medical databases are an additional arena for the balancing of public health goals and individual protections—in the case of databases, the right of privacy.

In their article on the use of large medical databases in this issue, Simon et al. attempt to achieve a reasoned middle ground between the needs of research and the protection of patient confidentiality. The Department of Health and Human Services (DHHS) has been addressing this very topic, but, as a member of DHHS, I am not in a position to comment on their deliberations. It is safe to say, however, that prior to the drafting of these regulations, much useful discussion focused on the goal of trying to find the right balance between research needs and privacy. Simon et al. make a convincing case that large medical and insurance databases are irreplaceable as a resource for population-based studies on diverse and important topics that affect the public health. Examples include the impact of different insurance benefit designs on access to care, the overall economic benefits of newer psychotropic medications despite high initial formulary costs compared with older drugs, the gap between knowledge and practice in the identification and treatment of depression and other mental disorders, and the powerful effect that depression exerts on overall medical utilization. Simon et al. argue cogently that large representative databases will become, if anything, more valuable as the organization and financing of health care change ever more rapidly in the United States. Without access to such databases, we will not be in a position to formulate policy based on evidence, but will be forced to rely on theory and guesswork.

Despite the clear need for data, both the public and policy makers have been increasingly concerned about personal privacy. Concerns about privacy have been heightened by the rise of corporations that assemble data about individuals, largely for the purpose of targeted marketing efforts, and by the extent to which the Internet permits mining of personal data. In this world of decreasing privacy, there is little that an individual might want to hold more in confidence than health information. For example, information about diagnoses and other aspects of health status can influence hiring and promotion decisions in the workplace and may have an impact on educational and other opportunities. Concerns about the privacy of health information are exacerbated when the conditions in question are mental illnesses, which, lamentably, remain misunderstood and stigmatized in our society. Many individuals have legitimate reason to worry that information about mental illness diagnoses, psychotropic drug use, or psychotherapy could be used to discriminate against them. Because of security weaknesses on the Internet

“The research community must enhance its support of IRBs.”

and because unauthorized individuals may gain access to confidential information, consumers of health care may have concerns even about databases that have clinical care as their sole use. These concerns are heightened when databases are shared with others who are not involved with the clinical care of the patient, including researchers.

The most difficult situation arises when a database that is needed for research cannot be effectively “anonymized.” Situations in which patient identifiers may be needed include research that must link separate inpatient and outpatient databases or clinical care databases with pharmacy databases. I believe that finding the balance between the needs of research and the need for human subject protections—at least as we look toward the future—must depend on the judgments of local institutional review boards (IRBs), as is the case in all research. Unfortunately, this is easier said than done.

Throughout their paper, Simon et al. attempt to find ethical positions that work both for research, by not being excessively restrictive or cumbersome, and for privacy. In some instances, their compromise positions might not stand up. For example, the authors suggest that controls on proprietary use of clinical data should differ from controls on data used in peer-reviewed public domain research. Unfortunately, this distinction is not always so straightforward given the tangle of public-private collaborations in research and given the fact that peer-reviewed research that begins in the public domain can readily be commercialized after the fact and used for proprietary purposes. In all likelihood, we will need clear consent standards for all medical database research in which patient identifiers are included.

First, looking to the future, I believe organizations (e.g., provider organizations, practitioners, pharmacies) that collect patient data should provide information in clear and plain language on the possible uses of such data, including clinical care, research, and commercial purposes. Of course, it is impossible to anticipate all of the uses to which future science will want to put data collections, and this should be noted. Second, when it comes to research, local IRBs are the cornerstone of human subject protections in the United States. IRBs or other appropriate peer review groups (e.g., National Institutes of Health study sections) that make information available to IRBs must be in a position to ensure that proposed research on medical databases is of high quality and therefore warrants any risk to subjects. IRBs must ensure that consent issues are considered appropriately, and that patient confidentiality is protected effectively. Here, I would underscore a point that the authors have made very well. The research community must enhance its support of IRBs. The variability in the staffing, expertise, and ability of IRBs to attend to many issues, including the use of medical databases, demands attention from the research community and from policy makers at all levels. The Institute of Medicine has pending a study on IRB roles in database research that hopefully will provide guidance, but at a minimum we must be sure that for all research that they oversee, IRBs have appropriate levels of financial and staff support and that the research community honors and supports the contributions of those who volunteer for IRB service. Simon et al. also make the important point that patient or consumer representation is essential when sensitive issues arise—including the issue of research use of medical databases. If we cannot demonstrate that we have a strong, consistent IRB system in our country, no amount of computer security or other “after-the-fact precautions” will save the important research that depends on medical databases.

STEVEN E. HYMAN, M.D.

Address reprint requests to Dr. Hyman, National Institute of Mental Health, 31 Center Drive, Room 4A-5, Bethesda, MD 20892-2475; shyman@mail.nih.gov (e-mail).