Privacy, Confidentiality, and the Use of Medical Records in Research

Leon Gordis and Ellen Gold

In recent years widespread public concern has developed in the United States over encroachments on privacy and confidentiality in many areas of American life. This concern has led to a proliferation of regulations and legislative proposals at local, state, and federal levels. Although the medical and epidemiologic research community has been highly successful in protecting the confidentiality of personal information obtained for research purposes, attention has nonetheless focused on possible threats to privacy that may be associated with research involving human subjects. The Privacy Act of 1974 (P.L. 93-579) was designed to ensure that personal information about individuals collected by federal agencies would be limited to that which is legally authorized and necessary and that the information would be maintained in a manner that would preclude unwarranted intrusions upon individual privacy. In addition, the act provided for creation of the Privacy Protection Study Commission, which subsequently made recommendations to Congress and to the President regarding various record-keeping practices in the private sector and the use of such records for research purposes (*1*).

In this article, we present an overview of the implications of the privacy and confidentiality issue for future medical and epidemiologic research in the United States. We first describe briefly the importance of population-based biomedical research; second, show how essential the use of medical records with individually identifiable information has been in the past; third, describe some of the health problems for which knowledge based on use of medical records is desperately needed; and finally, describe the safeguards that are in effect for protecting the rights to privacy of human research subjects.

The Society for Epidemiologic Research and the Association of American Medical Colleges are committed to protecting the confidentiality of the medical and personal data obtained in the course of research activities. We believe that privacy protection can best be accomplished through the regulations of the Department of Health, Education, and Welfare that are now in force and through new legislation based largely on the recommendations made by the Privacy Protection Study Commission in its report to Congress.

Population-Based Health Research

Epidemiology is the study of the distribution and dynamics of disease in human populations. Its purpose is to identify specific agents or factors, related to people and their environment, that may cause disease or identify people who are at high risk for developing a disease. In so doing, epidemiology provides the basis for public health programs designed to prevent and control disease. Prevention can be effected by reducing or eliminating exposure to a specific factor once its importance in producing disease has been demonstrated. There are at least two important reasons why identification of people at high risk for disease is desirable: (i) measures can be adopted to prevent such people from developing disease and (ii) medical supervision and screening tests can be provided when appropriate, so that if they do develop disease, their illnesses can be identified at an early stage.

Among the public health programs aided by knowledge resulting from epidemiologic investigations are those directed at preventing and controlling conditions such as infectious and cardiovascular diseases, cancer, and stroke. Epidemiologic methods are also essential for evaluating the efficacy of new preventive and therapeutic measures and any possible harmful side effects they may have. For example, the possible harmful effects of immunization against swine flu required epidemiologic investigations which, indeed, were responsible for demonstrating the relation between immunization and development of the Guillain-Barre syndrome (*2*).

Major Health Studies That Required Use of Medical Records

The major contributions of epidemiology to the understanding of disease have been based on studies in which medical and other vital records of many people were used, both for their data and for identifying those individuals appropriate for subsequent study. The studies were often conducted many years after the information was recorded. Their importance can be demonstrated by the following examples.

1) *Cancer.* Studies which demonstrated (i) the relationship of cigarette smoking to lung cancer, coronary heart disease, bladder cancer, and other conditions (*3–6*); (ii) an increased cancer risk associated with occupational exposure to substances such as asbestos and vinyl chloride (*7–14*); (iii) the increased risk of several types of cancer after exposure to radiation (*15–24*); (iv) that the daughters of women who received the hormone diethylstilbestrol (DES) during pregnancy have an increased risk of developing vaginal cancer (*25*); and (v) that women taking estrogens for menopausal symptoms have an increased risk of developing endometrial or uterine cancer (*26–28*).

2) *Cardiovascular diseases.* Studies which demonstrated (i) that high blood lipids, high blood pressure, and smoking shorten life expectancy, particularly through coronary heart disease (*29–34*); (ii) that women taking oral contraceptives have an increased risk of developing thromboembolism or stroke (*35–39*); and (iii) that administration of anticoagulants to patients with myocardial infarctions is associated with lower postinfarction mortality rates (*40, 41*).

3) *Infectious diseases.* Studies which led to the development of vaccines for poliomyelitis, measles, and other infectious diseases (*42–45*), and studies which showed that cases of polio subsequent to polio immunization in 1955 resulted from a vaccine lot contaminated with live virus (*46*).

4) *Child health.* Studies which demonstrated (i) that the administration of high

Leon Gordis is professor and chairman of the Department of Epidemiology and Ellen Gold is assistant professor of Epidemiology at the Johns Hopkins School of Hygiene and Public Health, Baltimore, Maryland 21205. This article is based on testimony presented before the Government Information and Individual Rights Subcommittee of the Committee on Government Operations of the U.S. House of Representatives on 9 April 1979 on behalf of the Society for Epidemiologic Research and the Association of American Medical Colleges.

0036-8075/80/0111-0153$01.00/0

concentrations of oxygen to premature infants results in blindness (*47–49*); (ii) that rubella (German measles) or other viral infection of the mother during pregnancy can produce congenital malformations in the infant (*50–55*); (iii) that radiation exposure of the mother during pregnancy is associated with an increased risk of congenital malformations and childhood cancer in her offspring (*15–17*); (iv) that Rh disease (erythroblastosis fetalis) in newborns can be prevented (*56, 57*); and (v) that comprehensive care programs for inner-city children and youth are effective in reducing rates of rheumatic fever (*58*).

These are but a handful of the studies which have produced important direct benefits for human health by identifying the factors associated with increased risk of disease, facilitating the development of preventive methods, and evaluating new ways of providing medical care and organizing health care delivery. (Some of these studies are considered in greater detail later.) It would be tragic if the potential benefits to society of such research were lost as a result of restrictions placed on the information available to researchers.

Information That Identifies Individuals

In order to carry out epidemiologic research, it is often necessary to identify individuals with specific diseases or disabilities, or individuals who share some common environmental exposure. Medical records are essential for identifying populations with specific diseases and for obtaining detailed historical, clinical, and laboratory information about individual patients. Access to these records is only a first step in identifying and evaluating patients with the particular disease under study so that they can subsequently be contacted and, with their informed consent, interviewed and studied further. Identification of individuals during the time the research is conducted is also essential to link records from different sources for a given person—records such as those kept by physicians, hospitals, and employers as well as birth and death certificates when appropriate. Identifying information about healthy persons (nonpatients) must also be made available for making comparisons when determining the cause of a disease. Without individually identifying information, which is essential for follow-up and record linkage, it would be virtually impossible to carry out epidemiologic studies of the etiologic characteristics, risk factors, natural history, and prognosis of a disease. Nor could studies be conducted in which new approaches to prevention, early detection, treatment, and delivery of health services are developed.

Patient Consent

Much population-based research would be very difficult to carry out if prior patient consent were required in order for the investigator to have access to medical records. The studies we described earlier were frequently conducted many years after the medical information was originally recorded, since the state of knowledge at the time the information was obtained may not even have permitted the study to be conceived. Thus patients' consents could not possibly have been obtained. In addition, reviewing medical records is often the first step used to identify patients with a given disease so that they may subsequently be traced, contacted, and, with their permission, studied further. Any requirement that patient consent be obtained before any medical record is reviewed would, therefore, be extremely destructive to medical and epidemiologic research.

Redisclosure

Provisions that allow redisclosure or rerelease of information for research or health statistics purposes (while maintaining safeguards of confidentiality and privacy) are extremely important, as the following example will illustrate. Because of the increasing importance of the problem of cancer, many cities and states are establishing cancer registries. These are lists of newly identified cancer patients, and they are designed to facilitate the long-term care of cancer patients, to alert health officials quickly to clusters of new types of cancer (which may reflect a new exposure to a carcinogen), and to permit investigators to identify all patients with a particular cancer so that the cause of that cancer can be investigated through well-designed epidemiologic studies.

Cancer registries generally obtain their data from hospitals and pathology laboratories. A registry would be of very limited use, however, were it not possible for the registry to go the next step and make its data available to legitimate cancer investigators. Thus rerelease of information with appropriate safeguards is essential.

A second example of the importance of rerelease concerns a large-scale study of women who had been exposed to x-rays during pregnancy. The study addressed the question of whether children from such pregnancies had an increased risk of developing cancer. Some years later, an investigator who had not been part of the original study team became interested in the reproductive histories of the children (who were by then fully grown) and in the status of their children's health. The children of the first study were contacted and asked a variety of medical and social questions. This was possible only because the data from the original study, including identifying information about individual subjects, were rereleased to the investigator. Throughout, rigid safeguards were maintained to protect the confidentiality of all data.

Research Use of Medical Records: Further Examples

Diethylstilbestrol and vaginal cancer. A few years ago, investigators in Boston demonstrated through an epidemiologic study that when mothers had been given DES during pregnancy to prevent miscarriage, female offspring from these pregnancies had an increased risk of developing a rare type of vaginal cancer when they reached adolescence (*25*). (This finding has important implications, since for many years DES was added to livestock feed in the United States.) Three features are particularly noteworthy: (i) the cancer did not appear in the person taking the medication but only in her female offspring; (ii) the cancer appeared some 15 to 20 years after exposure to DES, so it was necessary to go back many years to determine exposures and to identify the drugs taken in pregnancy; and (iii) the girls and young women who had this cancer were first identified through their medical records—only then could their mothers be contacted and studied further. If such use of medical records had been prohibited, or had been permitted only with the consent of the patient, this study—perhaps the first demonstration in human beings of transplacental carcinogenesis—would have been extremely difficult or impossible to carry out.

There may be other carcinogens that mothers should avoid during pregnancy because of the hazard to their children. To identify these agents, rigorous epidemiologic investigations in which medical records are used are needed to pro-

tect the health of women and their children.

Occupational cancers. Workers in certain industries are often subjected to high concentrations of potentially toxic substances. For example, workers exposed to vinyl chloride were shown to have an increased risk for liver cancer (*12–14*). This finding, which has now been confirmed, was obtained by reviewing and correlating the medical, work, and death records of large groups of employees in specific industries. Without access to these records it would have been impossible to identify and confirm vinyl chloride as a cause of cancer. Also, had there been a requirement that patient consent be obtained before the records were made available, these studies could not have been carried out because many patients had died by the time the study was done or had moved and could not be traced.

We have only begun to scratch the surface with respect to the toxic and cancer-producing potentials of substances to which workers are exposed during their daily labor. Any restriction that would preclude the identification of harmful substances and the documentation of their effects would be a major setback.

Oral contraceptives (the "pill"). Although the pill has been demonstrated to be a highly effective and convenient form of birth control and has been adopted by many women, many epidemiologic studies have demonstrated that women who take the pill for long periods of time increase their risk for blood clots, strokes, heart attacks, high blood pressure, liver tumors, gallbladder disease, congenital malformations in their offspring, and other conditions (*35–39*). These important findings were primarily the result of large-scale studies in which hospital and medical records were used—studies which, again, would have been impossible to carry out had patient consent been required. Epidemiologic studies of the effects of drugs like the pill are critical for protecting the health and well-being of the public.

Radiation. One of the questions posed by the accident at Three Mile Island in Pennsylvania has been, How serious is the potential risk to residents of the area who may have been exposed to radiation from the reactor? We know that high levels of radiation are extremely hazardous to human beings, but what is not known with any certainty is the extent of the hazard posed by low levels of radiation. In order to generate data on the hazards presented by low levels of radiation, it is necessary to collect information on a population that was exposed to low-level radiation. If such a population were identified, we would attempt to trace its members and obtain any relevant physician records, hospital records, or death certificates. For comparison purposes, it would be necessary to identify a similar but unexposed population and obtain similar records for its members to determine the rate of disease in that population. Only in this way could we determine whether the exposed group has more disease than the group that was not exposed. For the conclusions to be valid, complete records must be available for both groups. It would be necessary to know names and addresses, to have access to their medical and vital records with personal identifiers included, and to establish procedures for tracing, recontacting, and following up the members of both populations to determine all episodes of serious illness and death. If access to these records is restricted, Americans will be denied information on the hazards of low-level radiation. It is, therefore, essential that Congress ensure that legitimate medical and epidemiologic researchers have unhindered access to medical records. Such access, naturally, must be conditional on the demonstration by the investigator to his institutional review board (IRB) that he has provided adequately for protection of the privacy of the subjects in his study, as described below.

Existing Safeguards for Protecting Confidentiality

All epidemiologists and medical researchers have a major professional and personal responsibility to minimize invasion of privacy as much as possible and to protect vigorously the confidentiality of the data in their possession. The provisions of the National Research Act (P.L. 93-348) (especially its implementing regulations on protection of human subjects) codify an elaborate system of safeguards, currently in operation within the scientific community, to prevent violations of the rights of patients for purposes of research. Each investigator must justify to the IRB the rationale for subjecting any human research subject to any risk—including invasion of privacy—and must demonstrate the measures he or she is taking to ensure the confidentiality of all personal and medical data obtained in the course of the research (*59–63*). Investigators must assure the IRB that research data will be kept under lock and key, and must specify who will have access to the data, how and at what point in the research personal information will be effectively separated from other data, and whether or not the data will be retained at the close of the study, and if so, why. The IRB thoroughly reviews interview instruments and questionnaires, the consent statement, and any accompanying material—all of which must be sufficiently informative and understandable to enable the subjects to make an informed decision about participating. If the subjects are patients, they are regularly assured that their care will not be jeopardized in any way by their failure to participate. All subjects are assured that they are free to withdraw from a study at any time. Many of these provisions are spelled out in current regulations of the Department of Health, Education, and Welfare.

Immunity from Subpoena

It should be apparent that epidemiologists and other medical investigators are keenly sensitive to the challenge of ensuring confidentiality of data and protection of the privacy of research subjects. However, one important issue remains to be resolved, namely, the protection from subpoena of data gathered for research purposes. At present, no such protection is guaranteed in most research areas. Although an investigator may make the maximum effort to safeguard the confidentiality of information he obtains, such information remains subject to subpoena by a court of law. Therefore absolute assurance of confidentiality cannot be made to research subjects. Provision must be made for data obtained for research purposes to be immune from subpoena except under precisely stated extreme circumstances.

Current Challenges

Among the major public health problems in the United States are cancer, cardiovascular disease, and infectious diseases such as hepatitis, venereal disease, and influenza. Many cancers today are probably environmentally determined. In an interview some time ago, Dr. Arthur C. Upton, director of the National Cancer Institute, responded to a question about research needs in the cancer field by saying, "We need a lot more good epidemiology. It can tell us not only about environmental factors but also about genetic influences, and we really

do need to know about both." Upton's comments could also apply to research in cardiovascular diseases, high blood pressure, stroke, neurological diseases (including epilepsy), diabetes, arthritis, digestive diseases—indeed, to virtually all chronic and many infectious conditions.

The effects on human health of new drugs and other chemicals in the environment can only be identified through epidemiologic and other investigations, most of which depend on the availability of medical records. The maintenance of health and improvement of protection from environmental hazards requires the facilitation of epidemiologic research and the continued availability of medical records.

Society has a vital stake in epidemiologic and other medical research. We must ensure that the dignity and privacy of subjects will be protected without hindering the advancement of knowledge of disease. The social contract that facilitates the existence of individuals within social groups requires that each individual occasionally yield some of his rights, including privacy and freedom of action, for the benefit of society as a whole. Compliance with traffic regulations and with income tax laws are but two examples of the interactive workings of the social contract. Each society must decide when a limited compromise of individual rights is justified by the potential benefits to be derived by the community as a whole. Investigations of the etiology and natural history of disease and of the effectiveness of preventive and therapeutic interventions are of great potential benefit to society, but the conduct of such studies requires that, with proper safeguards, individually indentifiable data from medical records continue to be made accessible for medical and epidemiologic research.

References and Notes

1. Privacy Protection Study Commission, "Personal privacy in an information society" (Government Printing Office, Washington, D.C., 1977).
2. L. B. Schonberger, D. J. Bregman, J. Z. Sullivan-Bolyai, R. A. Keinlyside, D. W. Ziegler, H. S. Retailliau, D. L. Eddins, J. A. Bryan, *Am. J. Epidemiol.* **110**, 105 (1979).
3. H. A. Kahn, Ed., in *Natl. Cancer Inst. Monogr. 19* (1966), p. 1.
4. E. C. Hammond, *ibid.*, p. 127.
5. R. Doll and A. B. Hill, *Br. Med. J.* **1**, 1399 (1964).
6. ———, in *Natl. Cancer Inst. Monogr. 19* (1966), p. 205.
7. P. Enterline, P. Decoufle, V. Henderson, *Br. J. Ind. Med.* **30**, 162 (1973).
8. P. M. Stell and T. McGill, *Lancet* **1973-I**, 678 (1973).
9. J. Peto, R. Doll, S. V. Howard, L. J. Kinlen, H. C. Lewinsohn, *Br. J. Ind. Med.* **34**, 169 (1977).
10. G. Theriault and L. Grand-Bois, *Arch. Environ. Health* **35**, 15 (1978).
11. N. J. Vianna and A. K. Polan, *Lancet* **1978-I**, 1061 (1978).
12. R. J. Waxweiler, W. Stringer, J. K. Wagoner, J. Jones, *Ann. N.Y. Acad. Sci.* **271**, 40 (1976).
13. J. L. Creech, Jr., and M. N. Johnson, *J. Occup. Med.* **16**, 150 (1974).
14. R. H. Wyatt, J. M. Kotchen, D. L. Hochstrasser, J. W. Buchanan, B. R. Campbell, J. C. Slaughter, A. H. Doll, *Ann. N.Y. Acad. Sci.* **246**, 80 (1975).
15. A. Stewart, J. Webb, B. Hewitt, *Br. Med. J.* **2**, 1495 (1958).
16. A. Stewart and G. W. Kneale, *Lancet* **1968-I**, 104 (1968).
17. B. MacMahon, *J. Natl. Cancer Inst.* **28**, 1173 (1962).
18. B. MacMahon and G. B. Hutchison, *Acta Unio Int. Contra Cancrum* **20**, 1172 (1964).
19. B. Modan, D. Baidatz, H. Mart, R. Steinitz, S. G. Levin, *Lancet* **1974-I**, 277 (1974).
20. S. Graham, M. L. Levin, A. M. Lilienfeld, L. N. Schuman, R. Gibson, J. E. Dowd, L. Hempelmann, in *Natl. Cancer Inst. Monogr. 19* (1966), p. 347.
21. R. E. Albert and A. R. Omran, *Arch. Environ. Health* **17**, 899 (1968).
22. R. E. Shore, R. E. Albert, B. S. Pasternack, *ibid.* **31**, 17 (1976).
23. L. H. Hempelmann, W. J. Hall, M. Phillips, R. A. Cooper, W. R. Ames, *J. Natl. Cancer Inst.* **55**, 519 (1975).
24. M. J. Favus *et al.*, *N. Engl. J. Med.* **294**, 1019 (1976).
25. A. L. Herbst, H. Ulfelder, B. C. Poskanzer, *ibid.* **284**, 878 (1971).
26. D. C. Smith, R. Prentice, D. J. Thompson, W. L. Herrmann, *ibid.* **293**, 1164 (1975).
27. T. M. Mack, M. C. Pike, B. E. Henderson, R. I. Pfeffer, V. R. Gerkins, M. Author, S. E. Brown, *ibid.* **294**, 1262 (1976).
28. C. M. F. Antunes *et al.*, *ibid.* **300**, 9 (1979).
29. W. B. Kannel and T. Gordon, *Bull. N.Y. Acad. Med.* **54**, 573 (1978).
30. J. Truett, J. Cornfield, W. Kannel, *J. Chronic Dis.* **20**, 511 (1967).
31. L. Kuller and S. Tonascia, *ibid.* **24**, 111 (1971).
32. J. O. Humphries, L. Kuller, R. S. Ross, J. C. Freisinger, E. E. Page, *Circulation* **49**, 489 (1974).
33. J. Stamler, in *Diet and Atheroscerlosis*, C. Sirtori, G. Ricci, S. Gorini, Eds. (Plenum, New York, 1973), p. 125.
34. W. Winkelstein, Jr., A. Kagan, H. Kato, S. T. Sacks, *Am. J. Epidemiol.* **102**, 502 (1975).
35. Collaborative Group for the Study of Stroke in Young Women, *J. Am. Med. Assoc.* **231**, 718 (1975).
36. J. I. Mann and W. H. W. Inman, *Br. Med. J.* **2**, 245 (1975).
37. Royal College of General Practitioners, *An Interim Report from the Oral Contraception Study of the Royal College of General Practitioners* (Pitman Medical, New York, 1974).
38. P. E. Sartwell, A. T. Masi, F. G. Arthes, G. R. Greene, H. E. Smith, *Am. J. Epidemiol.* **90**, 365 (1969).
39. M. Vessey, R. Doll, R. Peto, B. Johnson, B. Wiggins, *J. Biosocial Sci.* **8**, 373 (1976).
40. M. Szklo, J. A. Tonascia, R. Goldberg, H. L. Kennedy, *J. Am. Med. Assoc.* **242**, 1261 (1979).
41. J. Tonascia, L. Gordis, H. Schmerler, *N. Engl. J. Med.* **292**, 1362 (1975).
42. R. E. Weibel, E. B. Buynak, J. Stokes, M. Hilleman, *Pediatrics* **49**, 334 (1972).
43. E. Abrutyn, K. L. Herrmann, A. W. Karchmer, J. P. Friedman, E. Page, J. J. Witte, *Am. J. Dis. Child.* **120**, 129 (1970).
44. M. G. Grand, S. A. Wyll, S. H. Gehlbach, P. J. Landrigan, R. J. Judelsohn, S. A. Zendel, J. J. Witte, *J. Am. Med. Assoc.* **220**, 1569 (1972).
45. T. Francis, Jr., *et al.*, *Final Report, Evaluation of the 1954 Field Trial of Poliomyelitis Vaccine* (Poliomyelitis Vaccine Evaluation Center, Ann Arbor, Mich., 1957).
46. N. Nathanson and A. D. Langmuir, *Am. J. Hyg.* **78**, 16 (1963).
47. K. Campbell, *Med. J. Aust.* **2**, 48 (1951).
48. A. Patz, L. E. Hoeck, E. DelaCruz, *Am. J. Ophthalmol.* **35**, 1248 (1952).
49. V. E. Kinsey, J. T. Jacobus, F. M. Hemphill, *Arch. Ophthalmol.* **56**, 481 (1956).
50. W. P. D. Logan, *Med. Illustr.* **8**, 502 (1954).
51. B. A. Hill, R. Doll, T. M. Galloway, J. P. Hughes, *Br. J. Prev. Soc. Med.* **12**, 1 (1958).
52. R. Lundstrom, *Acta Pediatr. Scand.* **41**, 583 (1952).
53. M. Siegal and M. Greenberg, *Am. J. Obstet. Gynecol.* **77**, 620 (1959).
54. M. Greenberg, O. Pellitteri, J. Barton, *J. Am. Med. Assoc.* **165**, 675 (1957).
55. M. Siegal, H. T. Fuerst, V. F. Guinee, *Am. J. Dis. Child.* **121**, 469 (1971).
56. P. Levine, E. M. Katzin, L. Burnham, *J. Am. Med. Assoc.* **116**, 825 (1941).
57. P. Levine, L. Burnham, E. M. Katzin, P. Vogel, *Am. J. Obstet. Gynecol.* **42**, 925 (1941).
58. L. Gordis, *N. Engl. J. Med.* **289**, 331 (1973).
59. J. H. U. Brown, *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **38**, 2049 (1979).
60. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *Report and Recommendations Institutional Review Boards* (Government Printing Office, Washington, D.C. 1978), Publ. (OS) 78-0008, pp. 39–54.
61. U.S. Department of Health, Education, and Welfare, *The Institutional Guide to DHEW Policy on Protection of Human Subjects*, Publ. (NIH) 72-102 (1972).
62. U.S. Department of Health, Education, and Welfare, *Fed. Reg.*, p. 18917 (1974).
63. ———, *ibid.*, p. 11854 (1975).