Protection of Medical Data

by J.H. van Bemmel

Department of Medical Informatics, Erasmus University Rotterdam, the Netherlands

1. INTRODUCTION

There is hardly any area of health care where the computer is not yet visible: in primary care, pharmacies, clinical laboratories, outpatient clinics, departments in hospitals – the influence of computers is seen everywhere. One of the applications that are most frequently installed is the storage of patient-related data for easy and fast retrieval at any place and time for which a user is authorized.

Medical data play a key role for diagnosis, therapy, prognosis, prevention, research, and management. The transport of such data is nowadays also supported by electronic means (*EDI*, electronic data interchange). Modern health care is no longer feasible without computers for data storage and retrieval, electronic transportation, processing, and interpretation.

Databases

One of the problems we are confronted with today, is whether this electronic storage, transmission, and processing of medical data should be further stimulated or whether it should be limited because of potential disadvantages and negative aspects, such as possible misuse of the data. Such considerations do, incidentally, equally apply to data storage in computers of the police, the civil service, or for fiscal purposes. In all areas it should be carefully considered whether the advantages of electronic data storage are in balance with present and future drawbacks. In any case, when using database systems for medical

purposes we should realize that the developments in information technology will proceed at an increasingly faster pace.

Relationships

Computers may assist in assessing the relationships between medical data and diseases. There exists a wide variety of medical data (pictures, biological signals, alphanumeric data); some of these data are dynamic and time-dependent, but others are permanent or stationary and remain with us for the rest of our life, such as gender, blood group, allergies and our genetic profile. For example, the relationship between genetic data and many diseases or handicaps is gradually becoming more evident because of progress in the field of molecular biology and genetics. Genetic data are not only of interest for individuals, but possibly for their relatives as well (see, e.g. [1,2]).

Different requirements are simultaneously fulfilled with the structured storage of medical data:

- the patient or client requires medical advice;
- the care provider wishes to control the process of medical care;
- the researcher wants to obtain more insight;
- insurance companies need to calculate risks;
- the government is interested in long-term planning and the setting of priorities;
- employers want to use data for planning and to be safeguarded against financial claims.

Since stationary medical data (read, e.g.: genetic data) are of increasing importance and have large impacts on privacy and data protection, we will pay special attention to this category. First of all, we will deal with the specific nature of medical data, including those of genetic origin. Subsequently we will deal with the purposes of computer storage of medical data. In doing so, we will consider some consequences for data protection.

2. MEDICAL DATA

Humans have the ability to discern new situations and are capable of new and creative tasks. The computer does not possess these faculties, but has complementary properties that supplement human shortcomings such as a finite memory, and slow and inaccurate data processing, e.g., calculating. However, the computer causes a twofold reduction of reality: computer processing should be highly structured and formalized, and computer-stored data can only be expressed in symbolic form, i.e., characters and numbers which are subsequently coded in binary form. Computers do not know how to handle singular events or individual persons, or issues that cannot be quantified or coded. Computers also have no capability for feelings, concepts, or intentions. All this has major consequences for the structuring of processes in medicine and the storage of medical data.

Use of Medical Data

The acquisition and processing of data for the solution of some technical problem may be complicated, but is in principle feasible. Technical processes can often be modelled and described in structural terms, and the data to be derived from such processes can often be expressed quantitatively. In several other areas, such as medicine, this is different. For instance, processes that deal with health and disease can only be partly described in a formal manner, and far from all data can be expressed in quantitative terms. The patient's disease is generally highly individual and unique, and in many cases the treating physician does not only rely on 'hard' facts such as, e.g., laboratory analysis or results from organ function analysis, but also on 'on soft' data from the patient history. When treating the patient there frequently is no fully formalized strategy; diagnostic findings and therapeutic possibilities have to be carefully balanced with the prognosis, life expectancy, possible risks, patient reaction, acceptance by relatives, etc.

In using medical data for patient care we have to consider one further aspect: very often there is no one-to-one relationhip between medical data and diseases. This is partly due to the limited formal description of medical processes and quantifiability of medical data, but is also caused by the large variability of all biological processes and the fact that all required data are often not fully available. These limitations lead to inaccuracies and uncertainties in the diagnosis and even to inevitable errors, denoted as false positive (FP, the disease is positively concluded but in reality not present) and false negative errors (FN, the disease is present, but not diagnosed). Errors of these types may also be caused by incomplete knowledge or improper use of diagnostic methods.

Types of Medical Data

Medical data can be categorized into different types, but this treatise will focus only on classification into permanent and variable data. The first category is perhaps the most privacy-prone. Use of these two types of data is generally also different: variable medical data (to which belong alpha-numeric data, biological signals, and pictures) are primarily used for the diagnosis and treatment of 'transient' diseases, whereas permanent, e.g., genetic data are often strongly related to one's life, possibly far in the future, to the prognosis. The last category is, as remarked earlier, also of interest for one's next of kin: parents, children, brothers, sisters. For instance, if someone in a family appears to have a genetically determined disease, then different relatives may also be carriers of the disease c.q. of the abnormal gene, either in a dominant or a recessive form.

Because of its importance, the circle of people interested in genetic data is at least as large as that for variable medical data. Genetic data do not change or age; they are of interest for people during their entire lifetime. Knowing one's own genetic profile and risks could possibly lead to a certain lifestyle; it

can also result in unbearable psychological suffering so that not knowing is sometimes a preferable option.

Data in Computers

What makes storage of medical data in computers so special? This is related to the nature of the computer; the fact that processes have to be formally structured and data have to be coded or quantitatively described. This indicates both the power and the limitation of the computer. All data have to be described in symbolic form, but in real life, including medical care, not all observations can be expressed in such form (think of patient feelings and expectations). The medical record as expressed in a computer, therefore, represents only a limited view of the patient's disease. If someone other than the treating physician uses only computer-stored patient data, he would not always be able to obtain a complete and reliable picture of the patient's complaints and observations. However, for many reasons, this often also holds for the written medical record. Both the patient and the physician have to be protected against an improper, let alone illegal, use of computer-stored medical data.

There is another very important difference between written and computer-stored medical data. Essentially, written data only serve direct patient care whereas computerized data can furthermore be used for epidemiology, medical research, the evaluation of medical care, or education. Computers enable us to use medical data to investigate, for instance, the relations between symptoms and diseases, or the effect of different therapies on patient outcome.

3. USERS OF MEDICAL DATA

Different groups of users are interested in medical data. Since especially the permanent medical data are particularly privacy-prone, we will investigate which people are interested in such data (e.g., [3,4]). Consecutively we will deal with the following groups of interested users: the patient or client and their relatives (parents, children, brothers, sisters), the treating physician (general practitioner (GP), specialist), the medical researcher, insurance companies (pension funds, health or life insurance companies), employers, and the government (e.g., the Ministries of Health or Justice).

The Patient or Client

Now that there is an exponentially increasing amount of medical data in computers it is of utmost importance for patients and physicians to have such data stored as reliably and as completely as possible. A complicating factor is that data from the same patient are often stored in many different databases for different purposes: general practice, clinical use, occupational medicine, usage by insurance companies, etc. These different views on the same patient may cause conflicts of interest and have impact on the patient's privacy.

It is not always desirable, or even permissible, that all care providers have access to these different databases, even when they intend to use the data solely

for the patient under treatment. For example, the patient will not appreciate when a radiologist has access to or is informed of the fact that once in the past he consulted a psychiatrist, or when a medical examiner has knowledge of his entire medical record. Even the patient himself may not always want full knowledge when, on the basis of such data, the possibility of getting some future disease might be predicted. A conflict of interest of a different type may arise when, in contrast, a relative is highly interested in such data or, the reverse, when some relative is informed about hereditary diseases that someone else in the family prefers not to know.

The patient should be able to understand as fully as possible the implications of a request from a physician to allow storage of his medical data – either permanent or variable – in a computer. He should maintain the right to freely decide his response and, based on some expert advice, to know what are the consequences of such a request [5]. It should also be made clear to the patient what the implications of such a request would entail for his relatives. It is a complex ethical and legal problem whether the patient has the right to prohibit this physician from informing his relatives in the case that his medical data might have major consequences for his next of kin. Protection of the privacy of an individual may, in some circumstances, hamper the interest of others. This consideration could even be extrapolated to society as a whole.

From the foregoing it becomes evident that, especially genetic data, are of wider interest than for the individual alone. In fact this aspect has not just become apparent due to the recent progress in molecular biology and human genetics. For many generations it was already known that problems such as diabetes, cardiac diseases or certain allergies were associated with certain families [6] and that genetic properties, once acquired, are transmitted to the offspring. But nowadays, with computer storage of medical data, a much wider field for proper use, as well as abuse, has been disclosed. For that reason, all members of the same family are interested in the protection of genetic data of one member against unlawful use. When someone puts his genetic data at the disposal of an insurance company this may also have consequences for his relatives.

The Treating Physician

It is not yet clear whether the treating physician – and especially the GP – should literally be the key person to open medical data to third persons, preferably in close agreement with the patient. Apart from that it is important that someone other than the patient himself stands up for his interest and guards access to the different medical databases in which the data are stored. In case the GP has knowledge of the risk someone is running based on his or her medical or genetic data, the GP may be confronted with the dilemma of whether to inform the person concerned, especially when the patient has not requested such information. It is even more difficult to decide whether that person's relatives should be informed. This decision becomes harder as the level of risk increases and the consequences become more far-reaching. Both the

passing on and withholding of information has legal and moral consequences for the physician, which requires that his mode of conduct should be regulated.

The Scientific Researcher

Both medical care and medical research make use of the same collected patient data. Such research is inconceivable without access to computer-stored medical data that should, preferably, be made anonymous. At the same time, however, such data should be properly documented, for instance with related diagnoses and it should be possible to collect such data over time so that trends and relationships can be investigated. The 'pooling' of data concerning rare diseases can only be accomplished if data are stored in large databases, in order to obtain further knowledge about the prognosis of diseases. This implies that it should in principle be possible to trace certain patients, which could prove to be in sheer contrast to the protection of privacy.

Also, early recognition of changes in the disease profiles of the entire population can only take place when data are collected in large databases and are analyzed by epidemiologists. If this had happened at the time of the so-called Softenon case, in principle this disease would have been discovered one year earlier. No wonder that computers are today the necessary tools for medical researchers.

Insurance Companies

Pension funds, health or life insurance companies, and sickness funds are, understandably, highly interested in all data that concern the person or patient for which they have to calculate the future risks. It is understandable that insurance companies might decide to increase premiums, or even refuse to cover some risk, if on the basis of the medical data the risk appears to be too high and/or the patient has an unfavorable life expectancy. On the other hand, the patient will do his utmost to obtain an insurance that is as beneficial as possible if only he, and not the examining physician, has knowledge of his future risks, e.g., perhaps based on his knowledge of hereditary diseases within his family.

If insurance companies would calculate the risk factor also taking into consideration genetic data, this could imply high expenses for certain groups of the population, or even their total exclusion from health insurance. In the latter cases legislators should provide the rules how to handle these circumstances so that access to genetic data should not imply the end to solidarity in health care, where healthy people carry part of the burden for the less fortunate ones [3,7]. Therefore, if genetic data would become accessible to insurance companies, there is a real danger for a new type of undesirable discrimination, i.e., against those people or families who have a genetic profile with an increased genetic risk [4,6,8].

To preclude such developments in society, it should be stipulated by law that certian medical data are not accessible to others than the physician who has the full confidence of the patient, and that the insurance premiums should not be related to genetic risks.

The Employer

Following in the wake of the insurance companies, employers are also interested in the medical data of their (future) employees. Here, however, matters are more complicated, since some types of work entail a higher risk for people with specific medical conditions (e.g., the combination of chemicals or dust with certain allergies or lung diseases), which could be detrimental to the health of the employee. This could lead to earlier onset of disease or unfitness for work resulting in financial disadvantage for the employer. But also third persons could become involved in such risks, for example airline pilots who have an established risk for cardiac diseases. The remark on solidarity is of relevance in such cases as well, but it should be realized that nobody should be challenged to undertake a higher risk if it can be avoided by choosing some other profession or employer. Also here, the physician examining for fitness for some type of work should be obliged to follow legal regulations regarding his access to medical data. The discussions around HIV and genetic data are illustrative for the problems in this respect.

The Government

Societies and governments are nowadays confronted with steadily increasing costs of health care, which amount to 9% (The Netherlands) or even over 12% (U.S.A.) of the GNP. For that reason there is a tendency to decrease the number of patient-days in hospitals or nursing homes, in some instances leading to the closure of entire hospitals, and to stimulate primary care and home care. Foremost, governments prefer to stimulate prevention and health care planning and to determine priorities in health care. For those reasons governmental authorities are highly interested in the prevalence and prevention of genetically determined diseases. It should be realized that there is a long distance, but also a gradual transition between interest for reasons of prevention and measures based on eugenetic intentions.

4. PRIVACY

Privacy implies several different issues. For instance, it means the right to be left alone, but it also signifies that everyone is entitled to decide for himself how, when and to what degree others may dispose of his (medical) data. In many countries, this right has been described in the law; in The Netherlands, this has even been laid down in the Constitution [9], in Europe it has been anchored in the Treaty for the Protection of Human Rights and Fundamental Freedom [10].

The Dutch Constitution specifies that all persons have 'the right that their privacy shall be respected' (article 1), also 'related to the recording and the provision of personal data' (article 2). 'The Law regulates the rights of persons regarding the cognizance of data that have been recorded about them and their usage, together with the correction of such data' (article 3). It may be evident from this that the privacy of a patient also concerns his bodily integrity, which is described in article 11.

Essentially, the privacy of the patient is guaranteed by the professional secrecy of the physician, which is at the same time a right of the patient. A patient should be able to transfer all his medical information to his physician, without fearing that the physician will pass these data to third parties without the patient's approval. This professional secrecy has been regulated in several articles of the Law in The Netherlands.

If the physician wants to fulfill his responsibility to guard patient data, he should take measures that data are well protected. This entails measures against loss, theft, or damage; against (unintended) abuse and/or false interpretations; also against intended use or misuse. The latter also includes that medical data would be unjustly used for purposes other than for which they were collected, without the provider of the data knowing this. To accomplish this, proper scientific (syntactic and semantic), technical (e.g., spatial safeguarding against damage or fire), software (such as passwords, auditing trails, confined functionality, encryption), and hardware measures (e.g., back-ups and double installation of essential parts such as disks) are required.

Because of the fact that modern health care provision very often requires teamwork instead of care by a single physician, and as a consequence of information technology, the individual physician is no longer capable to personally guarantee the patient's privacy. For that reason, after the regulation of the professional secrecy, modern society has also legally laid down the right to privacy. This means that for all automated registrations of personal data written regulations are required, to be supervised by a Privacy Committee. These regulations should contain descriptions of the purpose of the registration, the disposal of data to third parties, the right of all persons concerned to inspection, alteration and destruction of their data. In principle, these regulations do not concern anonymous data.

The sensitivity of medical data to privacy is very much dependent on the context in which they are used. For instance, psychiatric data are often indicated as being highly sensitive to privacy. Nevertheless, there are data on many other diseases that could damage someone's career and that are also privacy-prone. The mere fact that it is known that someone once had a medical consultation in a psychiatric clinic or was hospitalized in a cardiological department, might influence decisions of employers or insurance companies. Similar consequences apply to the use of certain drugs, documented in somenone's medical record.

Regretfully, the present privacy regulations in The Netherlands offer few guarantees that take into account the special character of genetic data that are of interest for more people other than the single person about whom they were recorded; perhaps being relevant for different generations of families. For that reason, in all systems in which genetic data are to be stored, the purpose of the data collection should be properly described; outside this scope it should be forbidden that such data are used. The same applies to coupling of different databases. Preferably, the data should be stored anonymously and the key to the data should be in the hands of a physician who is fully trusted by the patient, e.g., his general practioner. All data that are stored in any such system

have to be maximally reliable and objective; one should also be very careful in storing subjective data or personal opinions. All databases should be subjected to a periodic auditing; the Privacy Committee should take care of the observance of all such requirements. In these committees patient and consumer organizations should also be represented.

A physician who has (perhaps accidentally) knowledge of the risk of his patient or client, for instance based on this genetic data, now faces the difficult legal and ethical problem whether, in some circumstances, he should inform the patient - or even his relatives. Most experts in such matters have the opinion that the physician should only transfer information if the patient requests so and is able to carry the burden of knowing. In some circumstances people prefer to live further without having knowledge of their future destiny. In other circumstances, however, people may decide to be fully informed, for instance when they consider to marry and/or to have children. True prevention may ultimately imply the renouncement of offspring.

REFERENCES

1. Harris R. – Genetic counselling and the new genetic. TIG 1988; **4**: 52-6.
2. Loppe M. – The limits of genetic inquiry. Hastings Center Report 1987; **17**: 5-10.
3. Holtzman N.A. – Public interest in genetics and genetics in the public interest. Am J Med Genet 1980; **5**: 383-9.
4. Special issue. Am J Med Genet 1987; 26.
5. Rowley P.F. – Genetic discrimination: rights and responsibilities of tester and testee. Am J Hum Gen 1988; **43**: 105-6.
6. Lappé M. – Ethical issues in genetic screening for susceptibility to chronic lung disease. J Occup Med 1988; **30**: 493-501.
7. Kenen R.H., Schmidt R.M. – Stigmatization of carrier status: social implications of heterozygote genetic screening programs. Am J Publ Health 1978; **68**: 1116-20.
8. Holtzman N.A. – Recombinanant DNA technology, genetic tests and public police. Am J Hum Gen 1987; **42**: 624-32.
9. Constitution of the Kingdom of The Netherlands.
10. Treaty for the Protection of Human Rights and Fundamental Freedom. Rome; 1954, art. 8.