SECURITY OF TRANSPORTABLE COMPUTERIZED FILES

A. Bouckaert, M.D., Ph.D.
School of Public Health
Catholic University of Louvain
B-1200 Brussels, Belgium

1. SCAMPI (Starting on conception automatic medical profile of individual)

The SCAMPI system is a generalization of the personal medical file. It includes features like transportability, continuous update, protection of confidentiality, unerasability. As a distinctive character, it includes a summary of pregnancy and birth as its first subset. Other points of interest, together with the main specificities of the system are described below.

A) The file is stored on a floppy disk belonging to the patient (ID). Since the data were coded using a binary conversion, they are difficult to understand without a specific program. This specific program is not the patient's property. It includes a double key requiring for each access the simultaneous agreement of the patient and of the "master of the data". The meaning of "master of the data" and "owner of the program" will be made explicit later on.

B) ID is entirely made of records made by persons who are generally medical doctors, or medical personnel. Those data producers will be considered as "masters" of the data they have produced (and recorded), i.e. of a sub-ID.

For example, the gynecologist is the master of the data he fed into the file during pregnancy. To be master is to be responsible and co-owner.

Consequently, any ID can be broken down into sub-IDs with different masters, each one being co-owned by its master and by the ID owner. For example, in the sub-ID dealing with birth, data were produced by the gynecologist but also by the anesthetist and

by the pediatrician. These three persons share the medial secret of the sub-ID together with their collective co-ownership of the sub-ID according to the traditional rules of shared medical secrets.

- C) An integral copy of the ID is produced together with the ID itself. While it seems to be wise to ask to the family doctor to keep the copy, a more complicated system will actually be described in the next section.
The family doctor has no right to access the whole ID but only a subset defined by the validity extent of the double key patient-master.
- D) The only program allowing to read the whole ID is the property of the designers of the SCAMPI Software. However, even that person is not likely to be able to read the ID completely very often since this would require prior communication of the key of the patient and of all the keys of the masters.
- E) Data entry, display and update require the use of a PASCAL code. The PASCAL source is never communicated to data producers or owners. Hence; no paper or other visual memo will allow to break the rules of double keys by additional programming. Each sub-ID is accessed by a different program. Programs belong to three main classes :
 - 1° data entry, cheking, copy (for medical practitioners)
 - 2° administrative forms production (for public services and hospital administrators)
 - 3° statistics (for epidemiologists and public health managers).

Classes 1° and 2° are subjected to the traditional rules of medical ethics. Class 3° falls under the jurisdiction of the statistical secret. The latter is no less exacting than the medical secret since it requires for example a programmed deletion of all identifications, as well as special measures of protection against indirect identification that will be discussed below.

- F) Since January 1st, 1984, the first program of class 1° with ID production is run in the obsetrics department of the St. Luke Hospital in Brussels. By the end of 1984, that program will be

complemented by others belonging to classes 2° and 3°. A postneonatal pediatric follow-up will then become mandatory (immunization timetable, growth charts). It seems that SCAMPI can be expected to result into significant consequences in at least three fields :

- 1) From a medical point of view, a pediatric follow-up and a feedback between obstetrics and pediatrics can be obtained without any additional clerical work.
- 2) For administrators, the system could mean considerable improvement since the six to eight forms that are routinely produced now by hand for each birth will now be produced automatically with no limit on the number of copies. Important also for public health, the compulsory linkage between birth premium and a minimum number of prenatal visits can now be more easily and realistically enforced.
- 3) As a social consequence, the ownership of the data about her own pregnancy is a modest but real increase of the mother's responsibility and autonomy.

2. DATA ENTRY

The SCAMPI development was initiated by a PASCAL program used on the micro-computer Apple II for the obstetrics department of the St. Luke Hospital. The program runs interactively for data entry, storage, retrieval, corrections.

Two kinds of data are fed into the micro-computer during birth :

- a) A pregnancy record (about 300 pieces of information) extracted from the hand record just before it leaves the department to be included in the Central Archives of the hospital.
- b) A birth record (about 100 pieces of information) whose data are entered in real-time, and including an anesthesiological sub-record.

While the child stays in the neonatal department, the past part of the data is entered. Neonatal data entry cannot proceed later than the departure of mother and child from the hospital. Informations communicated by neonatal metabolic screening (PKU, hypothyroidism) will usually not be available at the time and will have to be entered later on.

Interactivity means e.g.

- a) That logical mismatches or impossibilities are diagnosed and their entry into the file is prevented.
- b) No missing data should be tolerated. Partial filling of a file for one birth precludes data entry for the next one (but for special procedures unknown to common operators).
- c) Data are stored on floppy disks. They can be retrieved, displayed, altered.
- d) The same program uses the same data in text-processing mode to print a letter to the family doctor.

This program (PETERPAN) was produced by a general PASCAL program (GENFORM) used to generate interactive questionnaires defined parametrically. GENFORM (except some minor improvements) was written in 1982. It requires comments, controls and some other parameters to be able to produce any computer questionnaire. GENFORM was used for a variety of medical problems (leprosy, school medicine, dental health). Data stored on floppy disk are first grouped into the PASCAL "record" structure. The record includes all data stored in a "packed array of CHAR" except anesthesiological data (REAL) and identifications (STRING). About 100 records can be stored on a single-density Apple II diskette. Replacement of STRINGS by numeric codes for identification allows the content of the floppy disk to reach 200 records. A floppy disk with PASCAL formatting and one single record is given to the mother when she leaves (ID of child).

3. FOLLOW-UP

The ID follows the child for all its life, growing at the same pace as it grows. It conforms the first day of its production to the two rules of patient ownership and shared use by programming.

Shared use means :

- 1) To read the ID, one needs as many programs as there are sub-IDs.
- 2) As a rule, any data producer owns the program that allows him to visualize the data he entered by himself.
- 3) Moreover he is entitled to read the data entered by other producers so far as this does not require another program in addition to his.

Traditionnally, and leaving aside any technical consideration, it is accepted that the medical secret is shared if this can be of any usefulness to the patient. The pediatrician is entitled to know what happened during pregnancy to the child he treats. This is no more a break of confidentiality than the usual rule of medical practice demanding that all data that could possibly be of interest to the physician in charge should be communicated to him. The same reasoning leads to other evidences : the proper treatment of the child does not require a full knowledge of the gynecological past of the mother. Since this is no more in the utmost interest of the patient, it is actually forbidden and should be made technically unfeasible by appropriate software or hardware protections. The same rules hold true for the information flow through pediatrics and school medicine.

4. INFORMATION SHARING

Some data deserve special consideration, like the neonatal metabolic screening results and the immunization timetable. The diagnosis of an inborn error of metabolism should be considered as a medical emergency and the relevant information should be communicated by phone or by fast mail to the family practitioner. Changing house or changing of family doctor at the moment of communication can be very hazardous and threaten to reduce the benefits of screening to nil. It is suggested that some device should be used to signal the temporary ignorance of the screening results. Since these results are usually not yet available in the hospital, and since communication between screening labs and hospital are likely to be faster than between lab and family, the ID copy of the hospital will signal by a "red flag" that the screening produced pathological results requiring immediate care. As soon as the hospital doctors are confident that a proper treatment has been started, the "red flag" is replaced by the identification of the family doctor, the screening results and the opening of a special subfile for subsequent biochemical controles.

A similar system can be used with the patient's ID in order to avoid interfering with the immunization timetable.

4.1. Ownership and use

Use and ownership are kept separate. The child (and up to the adult age, his legal representations i.e. his parents) is the owner of the ID. But he receives no program that would enable him to read it. He is also supposed to maintain the physical integrity of the ID. Copies will be supplied to compensate for loss or destruction but this will be charged to the owner in order to discourage carelessness.

The user can read all the data he produced himself as well as the data shared by other producers, and with the taking into account the programmed compliance with medical confidentiality regulations. Any copy to a central file (as would occur for scientific processing for example) requires that identifications should first be cleared. Some users are completely outside of the medical secrecy. Hence their share of the data is narrowly limited.

For example, hospital administration, insurances, the Ministry of Health and the town administrations can be allowed to a specific subset of the data without any medical connotation, related to accounting and certifications.

One of the long-term objectives of SCAMPI is to ban hand copies from the processing of medical informations.

4.2. Advantages and inconveniences

- a) Increased productivity of the hospital manpower.
Clericals tasks and, prominently, hand copies, were consuming a disproportionate share of their time availability.
- b) Information persistence : In the present situation, a large percentage of medical information can be considered as lost within a few years. To support this opinion, it is well known that the paediatrician has to trust the mother's memories at the first visit in school medicine in order to learn the birth weight of a child. Later on, those information losses will grow worse and more expensive (E.g. some cases of PKU went three times through the diagnostic procedure and the parents are still not aware of the specific risk at the time of the next pregnancy).
As a rule, the multiplication of diagnostic procedures in patients whose previous results were lost contributes for a certain degree to the costs of our expensive medical system.
- c) Scientific use of information : Most retrospective epidemiologi-

cal studies are plagued by the inaccuracy of collected data.

- d) Threats : Floppy disks cannot be seriously contemplated as long-term information supports. Hence, they should be dropped from the SCAMPI system as soon as possible. A satisfactory support should fulfill at least the following requirements :

- 1° not erasable (READ ONLY)
- 2° not alterable
- 3° cheap
- 4° transportable
- 5° inexpensive retrieval
- 6° obsolescence-protected by being readily transferred on a new storage medium

(N.B. : The "smart card" could probably meet those requirements)

- e) Threats : The impression of confidentiality for non-computerized medical records reflects principally their heterogeneity and systematic loss. Abuses are uncommon because information synthesis is as difficult for abusers as it is for legitimate users. As soon as the ID can be accessed, the potential profit of abuse increases dramatically.
- Summarizing, the medical secrecy of non-computerized records is usually the secrecy surrounding information already lost. Problems of security are likely to be met as soon as the information is available. Potential abusers include business, insurance companies, security agencies.
- f) Conclusion : In order to earn acceptability, SCAMPI needs efficient coding procedures. These procedures should be allowed to evolve at the same pace as the cryptological art since no absolute protection is known.

BIBLIOGRAPHY

- K. Boehm in "Medical Informatics Europe 78" Ed. by DAB Lindberg and PL Reichertz, Springer, Berlin, 1978.
- B.R. Beier et V.M. Brannigan in "Medinfo 83" Ed. by JH Van Bommel, MJ Ball et OW Gertz, North Holland, Amsterdam, 1983.
- L. Horbach in "Cryptography" Ed. by G Goos et J Hartmanis, Springer, Berlin, 1983.

4.3. Medical Secrecy

The very old concept of medical secret is a particular by sensitive aspect of professional secret.

The interests of patient and society are best served if information collected in the course of medical activities is kept secret, and this has been acknowledged for centuries. The medical secret is not only binding for the doctor but also for all persons that happen to be drawn into it (family, colleagues, medical personnel). Secret sharing by medical practioners is a natural complement of the medical secret since there is no other way to allow the patient to benefit from the diverse experience of the medical community. Its limits are consequent to the well-understood benefit of the patient : RX data, are obviously to be shared with the surgeon while psychiatric data are not to be communicated to the dermatologist.

Moreover, the patient has no right to allow disclosure of the medical secret whose provisions can only be lifted by a higher imperative justified by patient and society benefits. To accept the patient's right to lift the secret would obviously make him vulnerable to blackmail.

Since the principles are clear enough and received detailed legal and jurisprudential treatment, their translation into software should still be realized. Such a translation will make sure that no confusion arises between normal secret sharing between colleagues and criminal secret violation.

4.4. Notary

In the present situation, the medical secret is kept by the doctor and follows him in the grave. This is clearly information spillage and does not agree with our general concept of ownership of the data. A computerized protocol of data communication can be implemented that is activated not only by the patient's decision of going to another practioner or by the loss of ID copy but also by the death of the practitioner.

Generalizing, it amounts to using a medical notary rather than the family doctor as the respository of the integral ID.

Such a move involves more than just a change of denomination. The functions of keeper and user of medical information are more clearly

dissociated in the person of a medical notary than they could be in a medical practitioner.

Moreover, the medical notary does not need to be one physical person. It seems specially attractive to consider the case where two physical halfnotaries must cooperate. Each halfnotary owns a half ID and the addresses that allow the merging of the two halves and each halfnotary is aware of the identity of his counterpart.

No program allows to read the complete ID. The halfnotaries can only use programs that allow to merge specific parts of the ID as a preliminary step towards data communication on authenticated request of both patient and doctor.

4.5. The dictator

A dictator is useful to access some data of the ID without preliminary agreement of patient or doctor. The dictator's usefulness is obvious in cases of unconsciousness, emergency, etc. ... Many deaths or severe injuries result from the difficulty of communicating with unconscious persons. Most frequently, this occurs with brain damage, uraemia, metabolic shock ... and these emergencies are routinely neglected because they are misclassified as acute alcoholic intoxication. In large towns of the West, problems like drug, AV block, diabetes, are likely to be met with increasing frequency as a result of medical progress and social involution.

The dictator is able to read some informations (related to drug allergies, blood groups, previous diagnoses of shock). Any dictator's access should be authenticated and identified; the dictator will be held responsible of his actions before the appropriate jurisdictional court.

4.6. Transfers

A specific physically unique localization is unnecessary for the ID. It is sufficient to be able to assemble its halves, possibly by telecommunication from the halfnotaries.

4.7. Benefits and inconveniencies

The legal concepts that call for data encryption can also be invoked for new definitions like the medical notary and dictator. The theo-

retical legal basis of these concepts should be carefully investigated at the same time as the relevant software is developed.