

SO
H30

Draft of a chapter on "HEALTH INFORMATION SYSTEMS" to be part of a book "Epidemiology and Health Care - the application of epidemiological methods to the study of health services and disease prevention in the European Community".

The publication will be prepared by the Education Working Party of the Society for Social Medicine with financial support of the European Communities. Chairman of the Editorial Board: Prof. W.W. Holland (London).

Dr M.J. Hartgerink

november 1976

34/12/76

"EPIDEMIOLOGY AND HEALTH CARE"

Chapter: Health Information Systems - M.J. Hartgerink, M.D.

Definition

The meaning that should be given to the term "health information system" is more defined by common understanding than by very precise definition. As for the word "system" alone most definitions agree on something like "any whole from the standpoint of the methodic connection and arrangement of its constituent members". Even so the word sometimes gets very different meanings dependent on whether the natural coherence of components is stressed or the emphasis is put on the operational principle for the specific purpose at stake.

For the purpose of the subject of this book an attempt by the World Health Organization at defining a health information system seems adequate: "A mechanism for the collection, processing, analysis and transmission of information required for organizing and operating health services, and also for research and training" (WHO, 1973).

Goals

We must recognize that in the practice of medicine recollection of previous experience and collecting of information is as old as medical practice itself. And certainly already in the 19th century a lot of information from different sources in the medical and social field has systematically been brought together much to the advantage of decisions to be taken for the purposes of public health and for the treatment of the individual patient. But both the expansion of medicine as a science and the increasing complexity of the organization of health care delivery have put us before the task of mastering the increasing flow of information by carefully selecting and integrating the relevant facts.

"Information is the essential ingredient in decision making. The need for improved information systems in recent years has been made critical by the steady growth in size and complexity of organizations and data" (Rosove, 1968).

One of the goals for a health information system can be to provide the doctor with the sort of information that will help him to take the best possible decision towards any single patient under his care. The other goal lies in the experience that an effective functioning of public

health administration, of planning and of control demands readily available information and that this information is indispensable for adequate policy-making. The delivery, evaluation and control of the health care system is just not possible without the aid of reliable information about the functioning of the health care system. Equally, the collection, processing and analysis of data becomes an imperative necessity when it is a question of providing guidelines for planning, for rules and regulations pertinent to an effective and comprehensive operation of the health care system and of norms governing the quality of health services. Epidemiology, whether shortly defined as the study of the distribution and determinants of disease in man (MacMahon, 1970; Klark, 1974) or more broadly, including the determinants of the course of diseases (Hartgerink, 1972) is greatly dependent on the proper availability of information. Hence the interest of epidemiologists for health information systems. Certainly epidemiology can be one of the major scientific goals for it. These are in general terms the goals. But it should be stressed that the goals to be achieved by setting up a health information system must be specified clearly and more in detail for each single case before the development is undertaken. If the objective is primarily the type of information that doctors wish to have for their practice the systems design will be very different from the design that would be made if the objective is mainly to provide the managers, administrators and planners with information relevant to them. On certain conditions the requirement of both can be combined in one system. But generally it is very difficult to combine the amount of detail of medical observations and treatment necessary for pathological studies in depth with the much wider overall information which is usually required by administrators and planners. Basically a health information system whether automated or not should only be set up after careful consideration of its goals and the demands of its users (Brauers, 1976; Atsumi & Kaihara, 1975).

Sources

The sources which can provide input into a health information system appear to be manifold. From the epidemiological point of view: Birth Certificates, Medical Records of sickness episodes, Death Certificates, Notification of disease and special disease-registers, Insurance registers and Population registers should be considered as possible sources from which information can be transferred to a health information system. But also facts from morbidity surveys and special investigations or even

the results from medical screening of parts of the population can provide input-material for a more integrated health information system. Health administrators often will be interested in yet other facts like general operational information from the health services and financial implications.

The great diversity of sources and the need for more general availability of facilities for control, surveillance and investigation do pose organizational problems when considering the scope and structure of a health information system. Most of the more strictly medical information can only be obtained from the stations where primary health care is delivered or from hospitals. If the system is to cover an area or even a wider region it is certainly no mean task to bring all the sources together in a co-operative effort to make the system work. An example of how this can be imagined as an array of several dozens of information-sources linked to a "master patient register" is described and showed in diagram by Bodenham & Wellman. But such an elaborate system has as yet nowhere been undertaken and it remains to be seen if the complexity of a design on that scale could in practice be made to work (Bodenham & Wellman, 1972). The basic question is what sort of information and how much of it we want to integrate into the system and for what use it is meant. A well-considered selection of contents and sources is essential for any achievement.

As we will see later on in this chapter it is of great importance to decide if perhaps the facts from some sources should be handled in terms of a partly autonomous subsystem or even should be accepted as being pooled entirely separate with yet the technical possibility to be called upon. Then they can incidentally or at intervals be linked with the main system. In practically every country this will be the case with the data from legally prescribed procedures like notification of birth and death and general population registers. Moreover, if one can rely on the completeness and the accuracy of the sources, it is not necessary that the central part of the system should contain the same amount of detail that is contained in the contributing sources.

Types of information

As to the information-gathering on health care two main categories can be distinguished. One are the facts directly related to the condition of the individual patient. The other comprises data from the organization of health care delivery. It must be clear that epidemiology and operational

research (or management control) will have to draw information from both categories. Some people will claim that a third category of information can be distinguished which then are the services rendered directly to the patient. These of course can have a direct medical significance and at the same time be regarded as organizational items. Figure 1 shows how these categories of information can be seen at the same time as separate and in interrelationship. The two main lines can lead to reporting and archiving for each of them. Quite often it also will prove to be useful to report on the basis of facts from both the main streams of information. For planning purposes the mixed reporting appears to be essential. The facts directly related to the condition of the individual patient together with the identifying indications of the patient usually are called the medical record. A vast amount of literature has been published on the topic of the medical record, most of it concerning the hospital in-patient. This since many years ongoing discussion about the medical record shows in itself that it is not very easy to decide what should be the formalized contents of a medical record. Most doctors and certainly medical specialists usually make notes about the patient-history and the findings during observation and examination and of the outcome of treatment. But how much of it should be kept and fed into a central health information system? In fact there is no general agreement on the ideal medical record for general use (Central Health Service Council, 1965; Hartgerink, 1975).

From a purely medical point of view this uncertainty about the medical record as part of a health information system is not astonishing. There is a great difference in the way doctors in different medical specialisms describe the development of a disease. Can we really expect e.g. that a dermatologist would use the same indications for his findings in a patient as a physician or a neurologist? The way out of this dilemma which is commonly proposed is to choose for a restricted summary of the medical data. Sometimes this is described as the "Minimum basic data set" (U.S. Committee on Vital and Health Statistics, 1972; De Heulme, 1973). With a stronger accent on the medical findings many hospitals already use a standardized "Hospital Discharge Summary" as an important part of the input into their hospital information system (Wagner, 1968; Griesser, 1973; Van Egmond, 1974; Wagner - WHO, 1976).

Information from the primary health care entails problems of its own when it is to be used in a health information system. Some of the preventive work in primary health care can be stated as facts: e.g. vaccinations given, specific diagnostic tests performed. Also when a

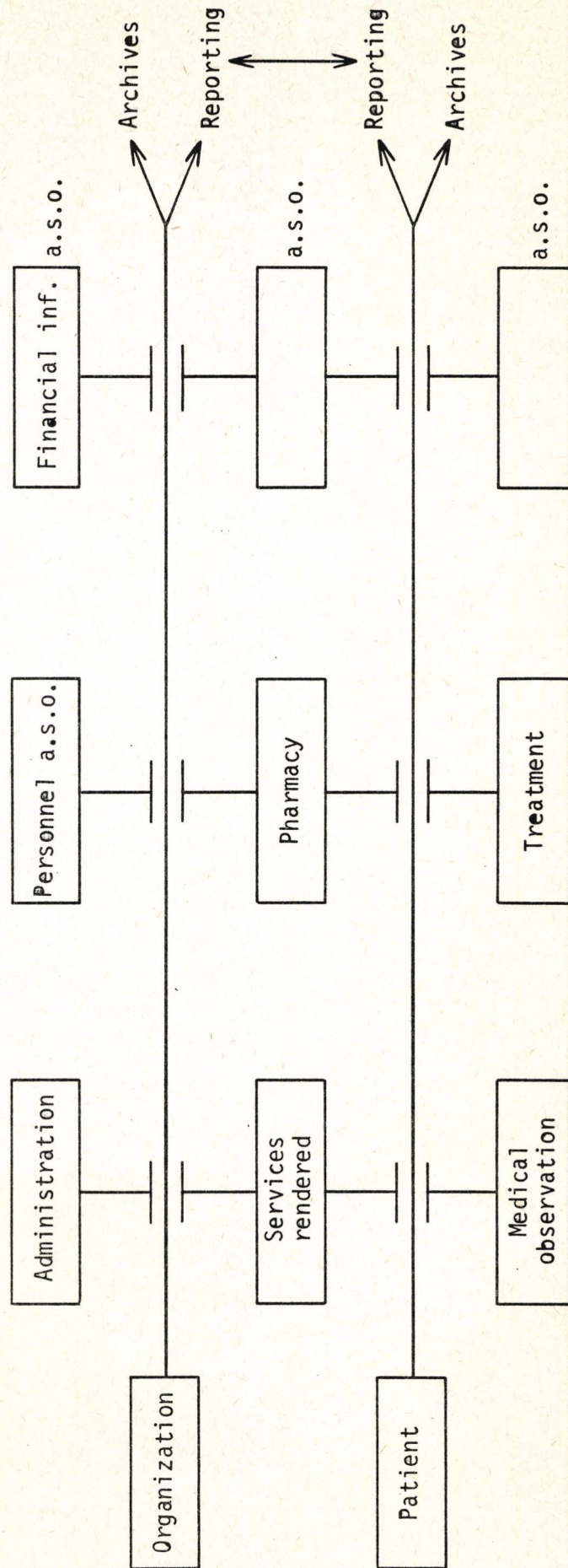


Fig. 1. Main streams in a health information system

diagnosis can be handled as a hard fact the findings in general practice are a valuable contribution to epidemiology. But much of the work and the observations in primary medical care is not easy to be defined in precise terms of diagnosis and concrete acts of treatment. In the contact between patient and general practitioner or public health nurse many general impressions and preliminary findings precede an eventual diagnosis and quite often no diagnosis at all can be reached. It is certainly difficult to feed a formalized report of this sort of medical transactions into an information system. Several attempts have yet been made to do so (Bradshaw - Smith, 1976; Van der Kooy, 1975). From the viewpoint of a more integrated health information system these have not been very successful. On the other hand it is often claimed that pooling of information from the primary medical care may in the future lead to a better understanding of early symptoms and developments of disease. However, for this purpose the best approach will be to deal with this relatively "soft" information in a separate information system.

The use of health information systems

Looking at what has been attempted and achieved three main domains of usage can be distinguished: epidemiology, management and planning. Other chapters of this book deal extensively with epidemiology and the material that is at the basis of this discipline. Obviously many of the facts the epidemiologist draws upon can be collected, stored and compiled in a health information system. There are good examples of special information systems for epidemiological purposes: disease registers, health surveys, information pooling in special investigations, etc. These dedicated systems themselves are usually not regarded as health information systems. But they can be part of a more complex system of recording what are the findings in a population. Perhaps it would be proper then to distinguish between general or integrated health information systems and specific dedicated systems.

Also good recording of sickness episodes and storage of these records is of value for future medical treatment of the individual patient.

Support for management and planning are often the goal for a health information system. The basic idea then is that a mixture of information from the organization of the services and from the medical reports on the patients can provide a better understanding of the real needs, the usage and efficiency of the health care provisions.

A special survey by the World Health Organization in the European Region

recently showed for hospital discharge summaries what the actual use of this type of information has been (Wagner - WHO, 1976).

Uses made of the hospital discharge summary forms

Uses	No. of answers	%
(1) Hospital activity statistics	77	85
(2) Hospital operation and management studies	44	48
(3) Hospital planning	38	42
(4) Administrative purpose	47	52
(5) Epidemiological research	55	60
(6) Patient scheduling	11	12
(7) Forwarding information to physician	32	35
(8) Entering information in data-bank	45	49
(9) Other	17	19

Total number of answers	366	

Number of respondents	91	100,0

Levels of use

The possible use of information on patients and health care delivery as indicated in the preceding paragraphs already suggests that this use will be different on different levels of the organization of the health services. It should be realized that all information is generated on the level where the health care delivery takes place. But part of the information that is to be collected has a meaning only on certain levels of administration, control and planning. On the other hand much of what is relevant on the lowest executive level will only for a small part serve the purpose of the co-ordinating level. Still further restrictive selection of information will be required for the policy making levels. As much consideration as has to be given to selecting of items for the primary input into the system is due to the selecting of items to be brought upwards to other levels in the system (De Paula, 1966; Hartgerink, 1975).

It also can be envisaged that some types of information from the primary input will only be brought up in a specific line for a specific purpose. This "splitting of the system" above the level of the primary input has sometimes been stressed by health economists for their purpose (Lièss, 1976). This can also be desirable for the purpose of epidemiology and other specific uses. Figure 2 represents an example of such a model.

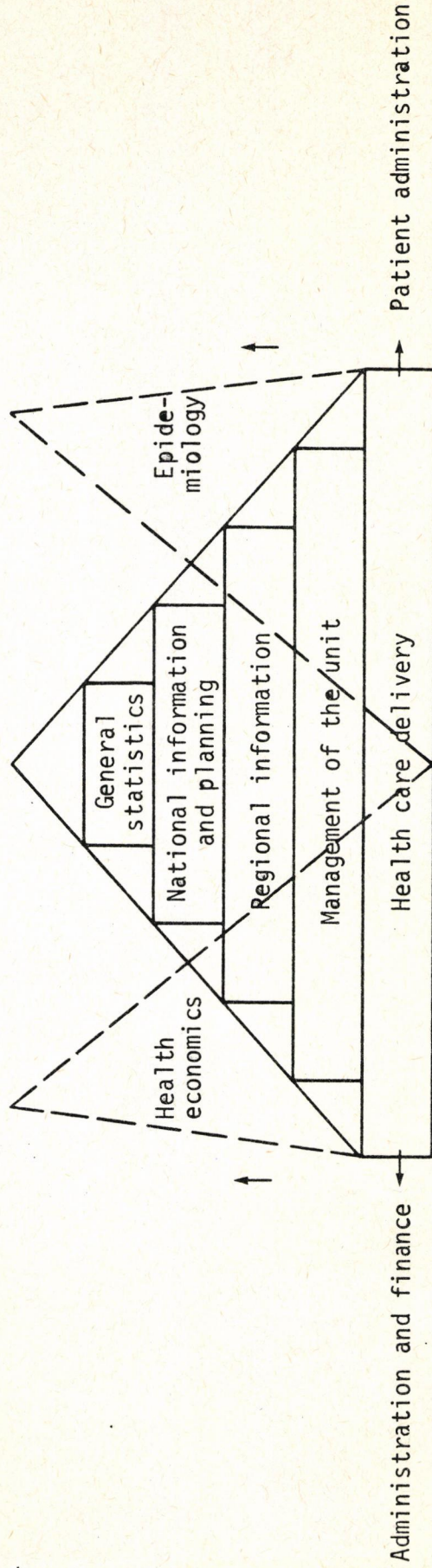


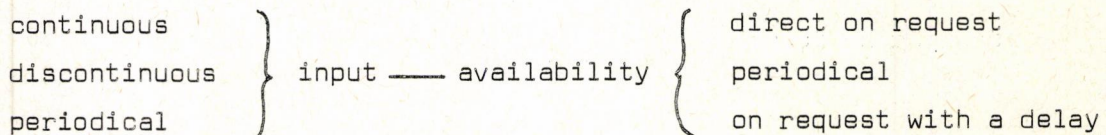
Fig. 2. A model for use of health information on different levels and for different purposes

Continuous and discontinuous information

Perhaps the computer-era has led to emphasize too much the possibility of coping with large continuous streams of information. There is no doubt that under circumstances the immediate input of facts and observations into a health information system can have advantages. The advantage can be the elimination of intermediate information carriers in the form of coded or uncoded written statements by the use of direct input-terminals for computers. Another advantage can be the very quick availability of the statement in the system. But usually very little need exists for this quick input-procedure. With rare exceptions a reasonable input delay is acceptable which then allows for batch-processing.

On the level of health care delivery exists often the wish for direct availability of data on individual patients in the system that will allow for quick and uninterrupted answers to requests for information. This has led to the concept of medical data-banks. These are basically stores of mostly medical information about individual patients that can be consulted at any time. For organizational purposes this will rarely be a necessity.

In any case for setting up a health information system a choice has to be made as to this matter. The following scheme shows the principle possibilities for this choice:



There is, however, a more fundamental aspect of the choice between continuous and discontinuous input in a health information system. Quite often the facts from a continuous flow of information in health care will in the output of the system be presented in relation with facts that represent a situation on a certain date. For example the calculation of the incidence of a disease relates the sum of a series of cases of this disease to the number of the population. In practice epidemiologists content themselves with a population number on a certain date if it may be presumed that during the period that the cases have been notified no significant change in the population has occurred. We have to realize that this is not always justified. Specially when organizational facts from the health service are brought into the system it is essential to verify if the number or the value of these facts (e.g. number of personnel or hospital beds available) have remained relatively stable for the period during which the continuous count of other items has taken place.

Pre-requisite condition

An information system brings together a great many facts of different kind and origin. It is essential that all contributors to the system will use exactly the same words with the same understanding of their meaning. Uniformity of definition is an absolute pre-requisite for a good information system. If number-values are used there must be absolute uniformity in the underlying measurement. It is astonishing how often this principle is disregarded. For instance: one hospital may count the number of days for treatment of a patient including the first and the last day, another hospital may subtract the dates of discharge and admission and by that will count one day less. Of course the total count of bed-usage and mean-values of duration of stay in hospital will then be very different.

Uniformity of definitions and uniformity of measurement are the cornerstones of information systems and statistics. Several techniques have been introduced to facilitate and to safeguard the fulfilment of these conditions. Agreement on classifications for diagnoses and for medical procedures has been furthered by the World Health Organization. This has laid the basis for some reasonably reliable international health statistics. But there remain several subjects, specially of organizational nature, that can not yet be presented in international comparisons.

A further aid for making information homologous and for formalizing it is the introduction of coding systems. This is a method of great value in automated data processing. But there exists as yet very little uniformity of coding systems. When setting up a health information system it is essential that this should be settled beforehand.

A great deal of the work of an epidemiologist is to safeguard the quality of the primary data. When these data are to be brought into a health information system it is just as important that the classifications, codes and further handling-procedures will assure that the quality of the information will meet the highest possible standards of uniformity and reliability.

Integration of facts from different sources

It is somehow remarkable that in the field of health care during the last twenty years so much attention has been given to the record linkage as a new technique. In essence it is just a better and more formalized way of bringing together parts of the life-history of a patient. As such it is exactly the same as "taking the patients history" which has been

practised in medicine ever since the days of Hippocrates. In medical literature the term "record linkage" has for the first time been used by Dunn in 1946. He then was head of the National Office of Vital Statistics in the USA and brought forward the question if not we could do much more with the statistics of birth and death when linkage of facts from different sources could be practised: "It would greatly enhance the significance of such statistics if they could be linked to other facts about the same individuals, such as, what sort of jobs do they hold, how many children do they have, what sort of illness did they suffer from, what kind of social environment do they live in?" (Dunn, 1946). In Europe it has among several others been specially Acheson who has contributed to the general acceptance of this viewpoint in health care and the necessary techniques. He defined medical record linkage as "..... the process of bringing together selected data of biological interest for a population commencing with the conception and ending in death, into a series of personal cumulative files, the files being so organized that they can also be assembled in family groups" (Acheson, 1967). For medical care and epidemiology this technique has for several problems opened new perspectives. As much for the study of the development of disease in individual cases as with regard to family-relationships and the relationship between diseases and social factors (Hartgerink, 1972).

In discussing health information systems we implicitly accept the idea that different sources will contribute to the system (cf. p.2). If the individual patient remains recognizable by name or code in bringing together the information from different sources we in fact use record linkage.

Also when no integration of personal files is contemplated the health information system will usually combine data from different sources for the compilation of statistics. If these sources are not already independently organized a good solution can be to handle a specific source as a sub-system with both an independent aim and a contributory task to the main system (Härö, 1976).

Privacy and responsibility

The development of integrate health information systems has confronted the individual with the problem of his privacy. What and how much do we allow other people and the public administration to know about us?

In the complexity of modern society it is unavoidable that for purposes of the administration of taxes, social assistance, health care, etc. the individual will reveal certain conditions of his personal life and status. He will to a certain extent have to accept that decisions are taken on the basis of his personal information. This request for information need not be extended as an encroachment on the rights and independence of the individual but on the contrary as an instrument for legal security and equality before the law. Thus the individual can not claim an absolute right to privacy.

As to medical care in the past every patient has accepted that his doctor keeps a record on his health. Since the medical profession and the health administration claim that more adequate medical care can be delivered if more of the facts about the treatment and previous treatments of a patient be known to them, many individuals have started to ask how much about their very private life then should be known by the claim of professionalists and by the administration. The progress of socio-medical care and of the science of medicine are considered to be too general an argument to be valid without restrictions for pooling identifiable personal medical information in a health information system.

It must be realized that the introduction of computers for information systems has influenced the discussion about privacy in two ways. On one hand it is undeniable that the written patient-file of a doctor can much more easily be read and be misused by unqualified persons than a computer-memory. On the other hand a computer-based system concentrates so many personal files that breaking into the system can hit a great many people's interests. And more specially a computer-based system can so easily combine information from different sources that the possible inroad on privacy is apparent.

Several techniques can be chosen to safeguard privacy in health information systems. The principal possibilities are:

1. Omitting the name of the patient and of the unit where the health care delivery has taken place. These can be replaced by identification numbers which are only known outside the system itself. There are many different approaches to the use of personal numbers instead of names. These can be completely unique or can be composed of a combination of personal characteristics like age, sex and date of birth. The degree of selectivity for the composition number has to be calculated in order to avoid misrouting of future information about the same individual. Sometimes a mixture of these two possibilities is to be preferred. For every

solution with numbers instead of names it has to be taken into account that patients rarely will know their personal "health care number" by heart. If they are to carry their number with them on a written document the secrecy of the number will soon be broken.

2. A safeguard against abuse of personal records through introducing secret call-numbers for qualified users in the communication with the information system. In practice these secret call-numbers soon become known to others than the only one entitled to it (be it only the secretary).

3. Much more a matter of principle is that personal records of different kind should not be brought together in one information system. E.g. no combination of judicial records with medical records. In several countries legislation excludes these combinations and has introduced supervisory boards to enforce and control the measures that have to be taken to safeguard the privacy of the individual.

But whatever the official rulings and the legally implemented safeguards are it remains of great importance that also at the level of the primary data collection all those who are engaged in the health care delivery and the reporting on it should stick to a formal discipline in this matter. For this it can be useful to draft a "confidentiality matrix". Such a matrix will give very precise indications as to what sort of information may be seen by whom. An arbitrary example is shown in fig. 3. It is essential that the matrix fits the specific situation of the unit for which it is meant.

A health information system is certainly a very powerful tool for the advancement of medical knowledge, for health surveillance and for the management of health services. But the implementation increases the responsibility of the medical profession for the welfare and the privacy of the individual patient.

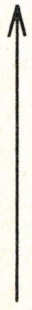

	Information levels 			
<u>Users</u> 	Medical record	Summary record	Anonymus record	Statistical presentation
Patient	X	Through own doctor	X	X
Doctor in charge	Unrestricted	ibid.	ibid.	ibid.
Department of treatment	X	Unrestricted	ibid.	ibid.
Research personnel	Without patient identity	ibid.	ibid.	ibid.
National Health Authorities	X	X	Unrestricted	ibid
Other institutes	Under special arrangement	Under special arrangement	Unrestricted	ibid.

Fig. 3. Example of a confidentiality matrix for a health information system

REFERENCES

- ACHESON, E.D. Medical record linkage. London, 1967
- ATSUMI, K., & S. KAIHARA. Planning a national medical information system. In: N.T.J. Bailey & M. Thompson (eds). Systems aspects of health planning. Amsterdam, 1975. p. 111 f.f.
- BODENHAM, K., & F. WELLMAN. Foundations for health service management. London, 1972
- BRADSHAW-SMITH, J.H. A computer record-keeping system for general practice. British medical Journal 1 (1976) 1395-1397
- BRAUERS, K.M. Systems analysis, planning and decision models. Amsterdam, 1976
- CENTRAL HEALTH SERVICE COUNCIL. The standardization of hospital medical records. London, 1965
- DUNN, H.L. Record linkage. American Journal of public Health 36 (1946) 1412
- EGMOND, J. van, et al. Belgian inter-university project on computerization of the medical record. Medinfo 1974. p. 45-49
- GRIESSER, G. Ansätze zu einem Gesundheits Informationssystem in Schleswig-Holstein. In: N. Hollberg (ed.). Computer: Aufgaben im Gesundheitswesen. Berlin, 1973. p. 88-102
- HÄRÖ, A.S. Strategies for the development of health related statistics. In: M. Pflanz & E. Schach. Cross national sociomedical research, concepts, methods, practice. Stuttgart, 1976
- HARTGERINK, M.J. Record linkage, verbinding van intermitterende ziekte-geschiedenissen. In: Computer en medische zorg. Leiden, 1972. p. 118-134
- HARTGERINK, M.J. Epidemiologisch onderzoek van bevolkingsgroepen. Tijdschrift voor sociale Geneeskunde 50 (1972) 935
- HARTGERINK, M.J. Traitement automatique des données cliniques. Nouvelles Techniques 4 (1975) 109-111
- HEAULME, M. de, et al. Archivage et interrogation de dossiers évolutifs. Revue Information méd. 4 (1973) 201-208
- KARK, S.L. Epidemiology and community medicine. New York, 1975. p. 1 f.f.
- KOOP, S. van der. Registratie van medische gegevens door de huisarts. Huisarts en Wetenschap 18 (1975) 261-263
- LIËSS, M. Economical aspects of central record systems; what the economist can expect from a centralized hospital data system. Edinburgh, 1976. Workshop on hospital statistics of the European Economic Communities

MacMAHON, B., & Th.F. PUGH. Epidemiology, principles and methods. Boston, 1970

PAULA, de. In: Report at British Joint Computer Conference, Eastbourne 1966. Cited by B. Barber in: The place of the computer in the delivery of health care. Univac, 1975

ROSOVE, P.E. Developing computer-based information systems. New York, 1967

U.S. NATIONAL COMMITTEE ON VITAL AND HEALTH STATISTICS. Uniform hospital abstract, minimum basic data set. Rockville, Md., 1972. (Vital and Health Statistics, Series 4, number 14)

WAGNER, G., M. IMMICK & C. KÖHLER. Der Krankenblattkopf der Heidelberger Kliniken. Methods of Information in Medicine 7 (1968) 17-25

WAGNER, G. Uses of the hospital discharge summary forms in the European region. Copenhagen, WHO, 1976

WORLD HEALTH ORGANIZATION. Health information systems; report on a conference. Copenhagen, 1973