Cardiological database in a community hospital – 10 years of experience.

Marek Demczuk[1], Jerzy Janecki[2], Anna Barbara Wardzińska[3], Teresa Marczykowska[4], Agnieszka Skomerska[5], Jakub Wróblewski[6].

1) Cardiology Department, Primate Cardinal Stefan Wyszyński Hospital, 98-200 Sieradz, ul. Armii Krajowej 7, demczukm@medprakt.pl fax. +(48) 43 8275286;
2) Institute of Biocybernetics and Biomedical Engineering, Warszawa, ul. Ks. Trojdena 4, jejan@ibib.waw.pl fax.+(48) 22 6599143;
3) ABACO s. c., Warszawa, ul. 3 Maja 2, E-mail abaco@pobox.com ;
4) Institute of Biocybernetics and Biomedical Engineering, Warszawa, ul. Ks. Trojdena 4, E-mail pawel@marcel.com.pl ;
5) Cardiology Department, Primate Cardinal Stefan Wyszyński Hospital, 98-200 Sieradz, ul. Armii Krajowej 7, askomer@medprakt.pl fax. +(48) 43 8275286;
6) Institute of Mathematics, Warsaw University, Banacha 2, 02-097 Warsaw. jakubw@mimuw.edu.pl, http://alfa.mimuw.edu.pl/~jakubw/

A multi-ward community hospital providing care in a rural area and localized in a relatively small town creates some specific possibilities in management of medical data. After the appropriate amount of time every one inhabitant of the area who has some health's problems must have his/her file in such a hospital. Thus, routine observational data collected in such a hospital may be a very rich source of a valuable and unique information, useful both while adequate territory-oriented medical policy is discussed as well as more general evidence-based medical scientific aims are considered[1].

However, to have such data exploitable the adequate computerization of the collection process is necessary[2,3]. The solutions of the problem which exist in some developed countries cannot be simply transferred to our country both due to financial limitations and due to differences in medical care systems[4,5,6,7]. Thus, original solutions are necessary.

Primate Cardinal Stefan Wyszyński Hospital in Sieradz (about 45.000 inhabitants) has been erected in 1990 as a community hospital providing care over the population of some 120.000. Additionally the hospital was served as a local center providing specialized care for the former Sieradz Province inhabited by some 450.000. In the Cardiology Department in close cooperation with the Institute of Biocybernetics from the very beginning we collected the basic medical data both in obligatory paper files, and in computer files as well.

The aim of this paper is to discuss, on the basis of our ten-year experience, some specific problems we found and to present our solutions.

In the computer management of routinely generated data we were oriented to achieve three at least partly discordant but equally for us important goals: 1) to improve the generation of some obligatory documents and/or some necessary administrative procedures; 2) to collect the most important part of raw data in the form exploitable for future scientific usage, while the strict and precise definitions of scientific goals were not existing at the moment of data collection; 3) to achieve both above aims without adequate financial support and in the human environment completely lacking of any preliminary knowledge in computer usage.

Due to financial limitations we restricted the computerization to four segments: 1) admission/discharge procedure for the whole hospital; 2) admission procedures in our department; 3) echocardiographic laboratory; 4) discharge procedures in our department.

1. The admission/discharge procedures in our hospital are supported with the software package originally created by some of us (AW, JJ). The unique government personal identification numbers (PESEL) is serving as the system's identification number. On Dec. 31, 1999, after 10 years of exploitation, the personal database of the package was filled with data of 81.324 patients admitted to our hospital in 1990-ties. The patients' names, PESEL numbers, addresses are available, as well as addresses of family members. The medical database of the system was filled with 173.895 records related to hospitalizations. The unique hospital medical record number provides the identification of the appropriate paper files, if necessary. The dates of admission and discharge are available, together with

the codes (six fields) of final diagnosis (9-th ICE Edition up to 1998, and 10-th Edition later on). The between-ward movement may be easily tracked, if necessary. The dates of death – if that happened in our hospital – are available, in the separate database, together with the code number of the cause. Many additional functions, necessary for the management of the hospital, are not discussed in this paper.

2. The admission procedure in our department is supported with the same above described software package. However, some modifications were performed by some of us (MD, AW). Thus, our cardiology system is additionally fitted with the medical database specifically designed for patients admitted to our Intensive Care Unit. Moreover, we have ten fields for diagnostic codes and from the beginning we are strictly using the 4-character version of ICD (3-character version is used elsewhere). Our medical bases have additional field for the ward-specific medical record number while the hospital medical record number is present, too. Thus, we have in our system all identifiers necessary for the identification of our patients' medical documentation (paper and digital as well) in our department, in other departments of our hospital, and in government databases, if necessary. On Dec. 31, 1999 our personal database was filled with the data of 7.610 patients admitted to our department on 12.410 separate occasions. The medical ICU base started to operate from 1993 and is filled with 3.852 records.
3. The echocardiographic laboratory in our department is supported with the software package originally designed by some of us (MD, TM, JJ). The construction of this package is similar but much less complicated than the previously described one. There are two bases – the personal one and the medical one. Because a relatively large proportion of patients have their echocardiographic examination on the out-patient basis, their PESEL number may not be available at the moment of examination. Thus, the PESEL number, the hospital medical record number and the ward-specific record number are collected in the personal database, but the system is generating its own patient's identification number. The data collected during the single one examination are identified through the unique examination's identification number. Previously we collected the values of weight, height, and blood pressure, together with 26 distinct measurements taken from 2D-guided echo M-mode, but only 12 of them obligatory. Those were the basic measurements of a geometry of hearts' structures and the basic measurements of a timing of some cardiac events. Six logical fields were present, too, to mark some specific echo views as "correct" or "incorrect" ones. Several multi-character memo fields are available for the written notes or additional descriptions, or conclusions as well. The medical database evolved in 1995 when the new echocardiographic equipment started to operate. From that moment 24 echo-Doppler measurements are collected additionally, as well as four logical fields. On the basis of stored data the system is able to perform automatically many algebraic calculations with the use of equations known from the literature. The strictly defined part of stored data, as well as some most popular and world-wide used results of calculations (less than 10% of all available in the system) are used to generate the Echo Examination Report which is automatically printed at request. In our system on Dec. 31, 1999 there were present the data from 11.753 echocardiographic examinations performed in 8.813 patients
4. The administrative part of the discharge procedure in our department is served by the previously described (see point 2) system. However, we have some further segment of the routine discharge procedure partly computerized. It is mandatory in our country for every hospitalization to create a Discharge Summary document. This document must consist of some mandatory sections: the patients' and hospitalizations' identification data, the diagnosis, the analytical laboratory data, the specific examinations and tests' results (such as X-ray, echo, stress test, e.t.c.), the other specialists consultations' results (if were performed), the performed treatment section, the written summary section, and the proposed further treatment section. Such a document is an integral and mandatory part of a paper medical record and usually it is written manually by a physician in paper files and later on retyped by secretary staff. We are using the text editor to print the document. All edited documents are stored. Some formalization of the mode of editing of the Discharge Summary is taking place. Our secretary staff have to their disposal the "empty digital form" of the document. Every one section in this "empty form" is edited in the same mode. Every one piece of information is following in some arbitrary defined order. Every one name of an analytical laboratory examination is available in the "empty form", as well as the adequate abbreviation of the units' denomination. The open language written sections are written by physicians in a slightly formalized way. The secretary staff is obliged to fill the empty spaces with the adequate data values, to erase the unused parts of the "empty form", to adjust the document appropriately, to print it, and to store it in the

digital archive under the name constructed on the basis of the unique ward-specific medical record number. This way all raw data existing in the medical record are stored in the digital form – and without any additional amount of working time. Our Discharge Summary document is usually adequate and accurate source of information what seems to be the basis when the routine usefulness of documentation from physicians' point of view is considered[7]. In our routine everyday work, while dealing with the previously hospitalized patient, we very rarely need to check some details in the original paper medical record. Every one hospitalization in our department has its representation in our Discharge Summary archive. Moreover, all the content of the archive is transferable to the database format.

At the moment the above described segments of computer system are working separately. However, every one record existing in any part of the system is clearly identified, and this is the basis to the successful merging. Moreover, we have in our records the key element, i.e. the PESEL number, which is necessary to verify, with the use of government data bases, the date of out-patient death. Thus, we have the dates of birth, events, co-morbid events (other departments including), and death, we have code numbers for diagnoses of events and co-morbid events, we have the laboratory data, we have the results of diagnostic examinations and tests (non-invasive and invasive ones, cardiological or other), we have the names of medicaments used, we have the dates and results of treatment procedures performed, we have the summarized descriptions which may be necessary to categorize the patient from some specific point of view, and we have the names of medicaments prescribed for further out-patient treatment.

Thus, we have in our disposal all the data necessary to address many scientific evidence-oriented questions related to the results of distinct forms of treatment. Some epidemiological analyses of these data may be of value, too. We present here some examples of research analysis we performed with the use of our system of computer data collection.

The data collected before April 26, 2000, were analyzed. All ten diagnostic fields of patient's records of the data base were searched for the diagnosis of the ischemic heart disease. 4.214 records of patients admitted to our department met the above criteria. In 435 cases diabetes mellitus was diagnosed, too. The data on treatment applied to these patients were taken from the Discharge Summary archive, after transformation of the appropriate sections of the archive to the data base format. The presence/absence of date of the out-department death was checked in the hospital archive as well as in the government PESEL archive. Thus, only the data from echo sector were not used in this analysis. Fig. 1. presents the results, in terms of all

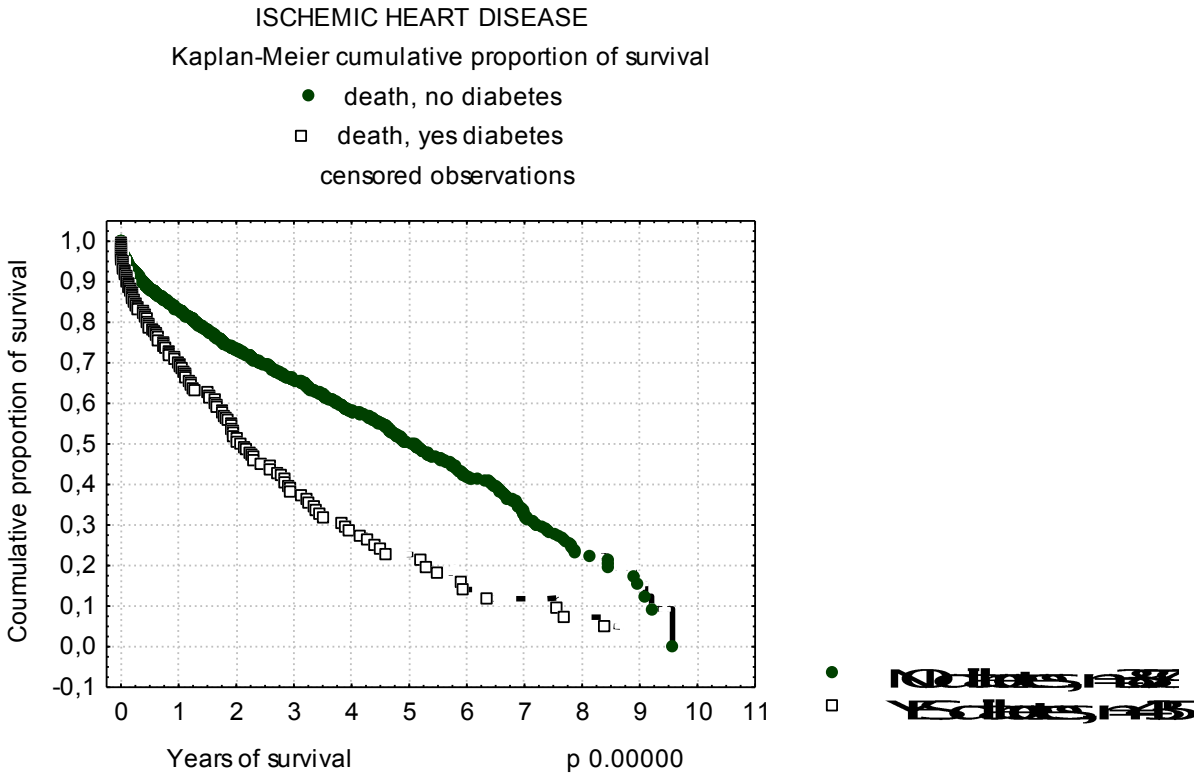

Fig.1. All causes mortality in ischemic heart disease – a comparison between diabetic and non-diabetic patients.

causes mortality, of the treatment of ischemic heart disease in comparison to ischemic heart disease complicated by diabetes mellitus. It should be stressed that the life survival curves presented here are raw, i.e. unadjusted to any confounding factor. However, with the use of our system we are able to consider many detailed data as confounding factors (i.e. concomitant diseases, concomitant therapies, specific test and/or analytical results, user-defined compact confounding factors, e.t.c). Thus, we are able to address very specific and strictly defined scientific questions with the use of our data

Fig.2. is presenting the results of research performed mainly with the use of the data from echocardiographic laboratory. However, the data from all parts of the system were necessary, as well as the confirmation of the dates of death from the government PESEL system, to complete and/or appropriately verify the true dates of events. In 1993 the data taken from 1.897 patients were searched to select the patients who had echocardiographic examination and no identifiable significant cardiovascular risk at the Rx day. The composite event has been defined as death, myocardial infarction, unstable angina or stroke. After three years on the basis of the presence/absence and the date of that event three separate groups of patients were found: “high risk”, “low risk” and “no risk”. The complex analysis[6] of 80 variables taken from the echocardiographic data base revealed that the groups may be successfully discriminated on the basis of that echo data. The validation with the use of the data taken from another 64 patients who died after examination revealed that the existence of risk was possible to detect in 91% of that patients. These results were achieved while the available numbers of records were much too small to give clinically valid results. However, that is a matter of time only. We present here Fig. 2. to prove how deep scientific analyses are possible with the use of our very simple system of data collection.

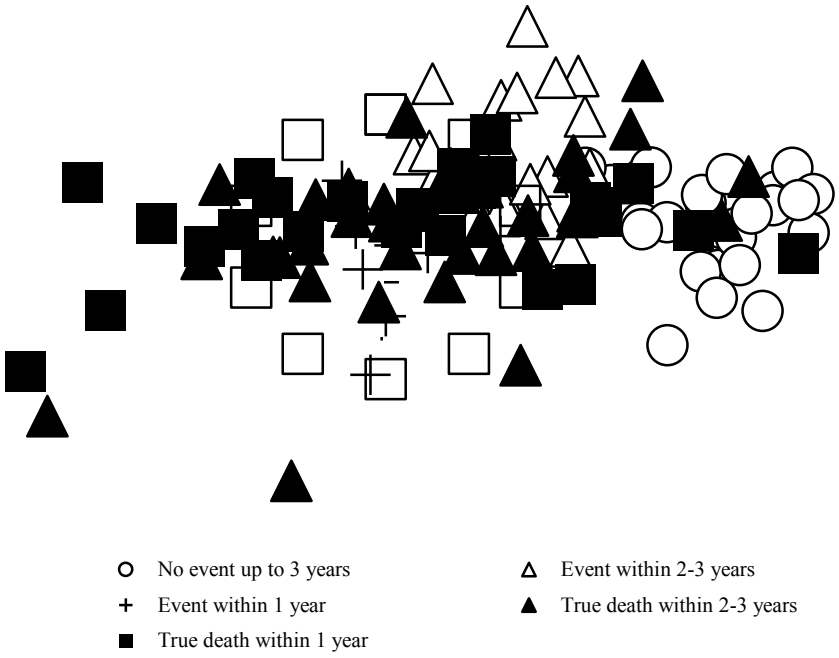

Fig. 2. The identification of risk on the basis of simple echo measurements.

Ten years ago, when we started to build this system, the financial support from medical administration practically did not exist, except of the very beginning when six computers (5 PC and one AT) were bought. There was no money for the additional staff. There was no staff members (except one of us – MD) who saw the computer in their previous professional life. Thus, the actual state is the result of many compromises forced by the evolving situation.

While dealing with the problem described in the point 4) - we could create the base for the data collection, but we had no software necessary to automatically generate the Discharge Summary document. Thus, such a base could not be used in the routine work of our department. Thus, some additional staff to operate such a base should be necessary, and there was no such a staff. On the other hand the Discharge Summary document ought to be generated. Thus – our solution is a kind of workable compromise.

While dealing with the problem described in the point 3) - we needed to generate the Echo Examination Report. It might be simply written manually, and later on rewritten to the Discharge Summary document, like other documents. But we need to learn, how to achieve two distinct goals: to create the necessary document and to create the adequate for scientific purposes database – all the above without any additional time for the operational procedures. Our solution works much better than we assumed. Due to scientifically oriented part of the system the contents of the Report provides the very deep and adequate insight into the structural and hemodynamic situation of the examined heart. Thus, the examination may be performed by any physician, and in the majority of occasions must not be personally overviewed by the one who is responsible for the patient's care. The differences in the successive examinations of the same patient are clearly visible, and the visual inspection of the stored video tapes provides the possibility to clarify the situation. Moreover, although theoretically possible, the same amount of information is practically unavailable without the computer system. The impact of this relatively small part of the computer system on the organization of physicians' everyday work is so tremendous that without this segment of the computer system our department would have to be reorganized.

While dealing with the problem described in the point 2) - in our department this part of the computer system is operated by a single one secretary officer, although all members of our administrative staff, as well as a few nurses, are able to operate the system. In every one hospital ward in our country some secretary officer must spend every day some amount of time to write down the personal data of the newly admitted patients to the paper files. Our staff is filling paper files, too, because that is mandatory. However, after a few years of experience we cannot detect the additional amount of time our staff needs to operate the

computer system. The only exception is when the staff of hospital admission room (where a great fluctuation of personnel exists) was made some mistake. The correction in our computer files takes a few minutes. Some corrections in the paper files may be much more time-consuming.

After 10 years of exploitation we still have our system separated from that hospital one. Our system is much better operated and the key data are much better controlled than in a much larger hospital system. If the everyday monitoring of the quality of the data is not adequately performed, such a simple question like the number of myocardial infarctions treated in a particular year cannot be addressed successfully without the manual searching of the paper files. On the other hand – if a system is able to serve automatically the true answers to the questions of such a type, any scientific question may be addressed to these data. We have learned that each one working day the data generated in the personal module of the system during the previous day must be carefully checked by the secretary staff, the same must be done every week for the data from this week, some key elements must be checked every month, and finally some elements of the database should be checked yearly. Interestingly, this is not the additional amount of working time. Every one ward must report their results on some timely basis. The adequate quality control of the data needs in the final effect less time yearly than the manual counting from the paper files necessary to create a single report. The rest may be done by some ready to use piece of software.

However, even after 10 years of operation we still have problems, mainly in the Discharge Summary sector of the system. Due to previous fluctuations in our personnel our quality control procedures failed in this sector. Thus, we were forced to check (AS), file by file, the whole content of this archive. Despite of this many typing mistakes are still present. In effect while the transformation of the data to the base format is taking place many additional work is necessary (JW, JJ, MD). Such a transformation is a must while we are preparing to use data from Discharge Summary (including these written in formalized natural language) in further analyses. To utilize all advantages of modern symbolic knowledge discovery methods (e.g. based on rough set theory[9,10]) we should collect in relational databases as much information about patients as possible – even if they look to be not concerned with analysts goal in direct way

In conclusion - the creation of the hospital computer system restricted to some key elements seems to be relatively inexpensive. The collection of data for scientific purposes seem to be possible, but should be performed automatically while the system should be designed mainly to help in the edition of the mandatory documents. This seems to be the only way to generate the accurate level of interest in the quality control. The carefully designed improvements in the contents of the routinely generated documents may have a very positive impact on the organization of routine everyday work.

---

[1] Cotham P, Cunningham B, Gilmer C. Outcomes analysis with case management. J Cardiovasc Manag 1994; 5: 23-4, 26-33.

[2] Thiesemann R.: Dokumentationsverfahren in der Geriatrie--Zur Unscharfe alltagsnah erhobener Klinikdaten. [Documentation procedures in geriatric – on unreliability of routinely gathered clinical data]. Z Gerontol Geriatr 1999; 1: 29-32.

[3] Tache A, Chatellier G, Menard J. Evaluation de la qualite du remplissage des dossiers d'une consultation d'hypertension arterielle. [Evaluation of the quality of filling out medical records in a hypertension consultation]. Arch Mal Coeur Vaiss 1997; 90: 1189-93.

[4] Taddei A, Dalmiani S, Cecchetti G, Macerata A, Carpeggiani C, Chelozzi L, Marchesi C.: Java-based medical record system for cardiology. Stud Health Technol Inform 1999; 68: 791-4.

[5] Dayhoff RE, Kuzmak PM, Kirin G, Frank S.: Providing a complete online multimedia patient record. Proc AMIA Symp 1999; 241-5.

[6] Fabbretti R, Dorsaz PA, Doriot PA, Rutishauser W.: Applying the object paradigm to a centralized database for a cardiology division. Int J Biomed Comput 1996; 42: 129-34.

[7] Stam H, van Ginneken AM.: Computer-based patient record with a cardiologic extension. Medinfo 1995; 8 Pt 2: 1666.

[8] Demczuk M, Janecki J, Jóźwiak-Limanowska L, Bednarczyk G, Kanicka-Lisiewicz A, Podsiadły T: Prediction of lethal or life-threatening cardio-circulatory events in patients who apparently are not at risk: a preliminary retrospective echocardiographic study. J Cardiol 1996; 28: 143-154.

[9] Pawlak Z. *Rough sets: Theoretical aspects of reasoning about data*. Kluwer, Dordrecht 1991

[10] Polkowski L., Skowron A. (eds.). *Rough Sets in Knowledge Discovery*. Physica-Verlag, Heidelberg 1998.