Towards Unambiguous Representation of Patient Data

AM van Ginneken, J van der Lei, PW Moorman
Dept. of Medical Informatics, Erasmus University
P.O. Box 1738, 3000DR Rotterdam The Netherlands

ABSTRACT

In the early eighties, the goal set for the development of computer-based patient records was the creation of patient records that were analogous to the paper record. In the Netherlands, where the number of physicians using computer-based patient records is steadily increasing, this strategy has been proven successful. Although these "paper-like" computer-based patient records were suitable for patient care, they were much less suited for other purposes. Experiments showed that the use of data for other purposes than those for which they were recorded, can only be performed reliably when these data permit unambiguous interpretation.

Due to a physician's limited time there is a constant tension between benefit and effort. Therefore, we should not expect the physician to provide the large amount of additional information, required for unambiguous interpretation of his record. Many of the inferences made by physicians are based on general knowledge and do not require specific, patient related information. We have focused our research on the potential of using knowledge about concepts in the patient record, to infer information, that is implicit in the patient data. The paper discusses considerations with respect to possible strategies to elicit a maximum of information with a minimum of effort from the physician.

INTRODUCTION

The development of computer-based patient records has been an active area of research in the past years. Despite these efforts, wide spread use of computer-based patient records in daily practice has not yet occurred. In The Netherlands, however, we observe a rapid introduction of computer-based patient records in primary care: more then 50% of the Dutch general practitioners have purchased an information system, and the computer-based patient records provided by these systems are replacing conventional paper-based records [1]. Professional organizations played an important role in promoting information systems in primary care by defining software requirements for these systems [2]. In some regions of The Netherlands, paperless medical offices are integrated in regional electronic networks that connect dozens of primary-care physicians with the hospitals, pharmacies, and laboratories in their area.

Our research group has been conducting research into computer-based patient records during the past 10 years. The rapid introduction of computer-based patient records in primary care, has provided us with an opportunity to study actual use of computer-based patient records, and to conduct experiments using data from those records. Based on these experiments, we have set a research agenda for our work in computer-based patient records. In this paper, we discuss our current research focus in the area of modelling the computer-based patient record. This paper is based on our experience in The Netherlands; in view of the significant difference between countries, the reader will have to determine the degree to which our experience is relevant for other settings.

REPLACING OR AUGMENTING THE PAPER-BASED PATIENT RECORD

In the early eighties, we developed a primary-care information-system, called Elias, for the general practitioner [3]. In The Netherlands, Elias pioneered the notion of a paperless practice. The principal design decision of that system was to replace the paper-based patient record rather then to augment a paper-based patient record. Augmenting a paper-based patient record leads to a hybrid record; the complexity introduced by the combination of paper record and computer-based record may outweigh the benefits gained from automating part of the patient record. The consequence of replacing rather then augmenting the paper-based patient record was that the general practitioner had to do record keeping with the system herself, thus forcing her to acquire new skills such as typing. We had, at that time, no definite evidence that replacing paper-based patient records with computer-based records in primary care could be achieved. Elias allowed us to test the feasibility of replacing paper-based records with computer-based records on a large scale.

Subsequent developments have shown that general practitioners are able to use computer-based patient records. At present, several commercially-available information systems provide the general practitioner with computer-based patient records; 25% of the Dutch general practitioners maintain such computer-based records [1]. We use the term "computer-based patient record" for a patient record, that is entered and maintained by general practitioners on a computer, and that fully replaces the paper-based patient record. The Elias system supports SOAP (Subjective-Objective-Assessment-Plan) coding and a problem-oriented medical record. Although all data within each SOAP category can be entered as free text, the system supports coding of the data [4]. Medication can be coded for the purpose of drug-interaction and drug-doses monitoring.

A recent study by the Institute of Medicine argues that most technological barriers that have in the past impeded development of computer-based patient records have disappeared or are disappearing [5]. Our experience supports, at least for Dutch primary care, the validity of that claim.

69

COMMUNICATION

As the number of general practitioners that use information systems increased, our next experiment centered on electronic communication. In a medium sized Dutch city (population of ca. 110.000), approximately 37 general practitioners use the Elias system. These systems were integrated into an electronic network with hospitals and laboratories in the same city. Admission/discharge -, laboratory reports, and messages from colleagues are received by electronic mail, and the computer-based patient records are updated automatically. We measured the flow of data over the network, and we surveyed the general practitioners about the use of electronic communication. We discovered that general practitioners were able to integrate this technology into their practice [6]. For example, a survey of the participating practitioners showed that electronic mail was rated positive and became accepted in daily practice.

The results of this experiment show that general practitioners are, in addition to using computer-based patient records, also able to integrate electronic communication into their style of practice. The acceptance of both computer-based patient records and electronic communication creates new research opportunities: it enables research directed at computer-based patient records that are shared by physicians working in different environments.

IMPROVING QUALITY OF CARE

In general, improving the quality of care requires (a) a means of assessing care as it is currently delivered, (b) the identification of areas in which care can be improved, (c) the identification of recommended behavior (e.g., a consensus meeting to establish guidelines), and (d) a method for informing the physician of preferred alternatives. Previous research, however, has shown that efforts directed at changing the behavior of physicians have met with varying (even disappointing) degrees of success. For example, Lomas at al. report that, even when practitioners view guidelines favorably, they are slow to change behavior; the development and dissemination of consensus recommendations seems to have little apparent effect on medical practice [7,8]. Other researchers have argued that, if the patient record were automated, computer-based audit of those records could constitute an important mechanism for improving quality of care.

In yet another experiment, we had physicians and a critiquing system review computer-based records of patients treated for hypertension [9,10]. In that study, we could demonstrate that both human reviewers and the critiquing system were able to monitor the hypertension treatment. An important finding of that study, however, was that a significant part of the data in the patient records were not coded (that is, available only in free text), and the data could not always be interpreted unambiguously. The study confronted us with the consequences of data that are not described in sufficiently formal notation.

LESSONS LEARNED

For the Dutch situation, we observe that, in primary care, computer-based patient records can replace paper-based records: Computer-based patient records are replacing existing paper-based records, and use of computers both for maintaining patient records and for communication is becoming an integral part of the physician's style of practice. Thus, in The Netherlands, the available technical infrastructure and the attitude of general practitioners create an environment in which new research topics such as shared patients records (that is, a single patient record shared by physicians working at different sites) becomes feasible. One important factor, however, is the limited degree of formalization on the computer-based patient record: data in such records (like data in paper-based records) are subject to different interpretations by different physicians.

RESEARCH AGENDA

Developments, such as shared records, computer-based audit of records, and quality assessment, indicate that the use of data in patient records is expanding beyond patient care only [5,11-14]. Using data for different purposes than for which they were recorded carries the risk of erroneous interpretation of these data, unless those data permit unambiguous interpretation. We believe, therefore, that it is important to focus our research on the formal representation of patient data.

Unambiguous interpretation of data requires unambiguous representation of these data. In this paper, we articulate considerations for the design of a new generation of the computer-based patient record. This is research in progress; we do not provide solutions. We present our present research issues and we will discuss the problems and trade-offs we face.

We have chosen to focus on the formalization of the patient history and physical examination, which are still the least formalized parts (predominantly free text) of the currently available Dutch computer-based patient records. Although the general strategy of history taking and physical examination is well described in books, the physician has freedom with respect to the completeness and detail in which the information is acquired from the patient and documented in the record. As a result, patient records of different physicians show much variation in style and content. This variation is no problem in itself. Difficulty in interpretation of a patient record arises from the fact that the data in these records reflect only partially the information acquired from the patient and the thought processes of the physician who created the record. Based on her own notes in the record, a physician is able to recall a more complete picture of the patient than reflected by the notes only [15]. After the process of data reduction from the physician's brain to the patient record, information, crucial for unambiguous interpretation of the record by other care providers and researchers, may be missing. This is even more pronounced when

data is to be interpreted by decision-support systems.

An important reason for ambiguity is lack of contextual information, which may be caused by missing data or by missing relations between data. When a record contains a description of an ulcer without mentioning where it is located, a physician may be able to infer the location from the description, but the location is not explicitly known. However, it is not always possible to make correct inferences from patient data. We were confronted with the limited ability to interpret patient data when we developed Hypercritic, a critiquing system that audits treatment based on computer-based patient data [9,16]. For example, when a diagnosis is recorded in the patient record and the physician prescribes a drug that has as one of its indications that specific diagnosis, then Hypercritic assumes that the drug was given for that diagnosis. After discussing Hypercritic's remarks with the physicians, however, a significant number of such assumptions proved to be erroneous. The reason for these errors was that the patient records involved, lacked information that enabled unambiguous interpretation of their contents.

We are well aware of the tension between unambiguity and efficiency. Even if it were possible to keep a record exhaustive in the sense that every possible item is explicitly described, than the effort would outweigh the benefit of such extensive record keeping. Therefore, the strategy to reduce ambiguity in the patient record should not be based, for example, on forcing physicians to fill in large numbers of screens with predefined items. Even optional screens for further specification of signs and symptoms are likely to be skipped when physicians are often confronted with them.

It is our belief that part of the information, that is not explicitly represented in a patient record, can be inferred from that record, using knowledge about the patient record in general: its concepts and how they are related. When a record contains two notes concerning an ulcer, a physician will conclude from the location in these notes, wether these pertain to one or two ulcers. Such a conclusion can also be inferred without human intervention when the knowledge, needed to arrive at that conclusion, has been made explicit. Hence, the use of knowledge about the patient record in general may help us to reduce ambiguity while asking relatively less information from the physician.

We will discuss the types of problems that we encountered in our attempt to reduce ambiguity when representing patient data. Some of these problems apply to patient records in general, while others are related to the (automated) interpretation of a computer-based patient record.

CONSIDERATIONS CONCERNING AMBIGUITY IN PATIENT DATA

Absent Data

Since it is impractical to describe every possible detail about the patient, a physician records information related to the complaint of the patient and the corresponding physical examination, whereas other topics are briefly dealt with or not mentioned at all. As mentioned previously, data reduction takes place when data are recorded. A physician, who reads a MR of a colleague will not be able to reconstruct the thought processes of her colleague and will develop her own hypotheses [16,17]. As a result, questions may arise with respect to topics that are not mentioned in the record. How should absent data be interpreted: "it is not known because it was not asked or observed", or "it is probably not abnormal, otherwise my colleague would have recorded it"? To solve such a question may require repetition of questions to the patient, physical examination, or even lab tests. Several studies account of missing data in patient records; studies that illustrate the trade-off between completeness and benefit [5,18]. One strategy to encourage data-entry is to bring items to the attention of the physician. It has been shown that data collection and information content of patient records can be improved by using well-structured forms or computer applications for data-entry [19,20]. Another strategy, to be discussed in the next section, may be to facilitate the entry of data, explicitly representing the absence of abnormalities. The main question is how we can reduce missing data, taken into account that it is not practical for a physician to be exhaustive.

The Meaning of the Statement "No Abnormality"

Many items in patient records are described by the simple statement "no abnormality". The statement constitutes a summary of a set of findings that have not been made explicit. Our clinicians brought up that the interpretation of such a statement poses questions, closely related to those in the previous paragraph. The question is no longer if the physician dealt with the item, but how he dealt with the item. Take for example the statement "heart: no abnormality". The meaning of this statement depends on what investigations a physician is used to include in the physical examination of the heart. In the absence of abnormalities, one physician may suffice with listening for cardiac murmurs, while another physician may carefully determine the heart size by palpation and listen to the heart tones as well. In the absence of abnormalities, a physician may simply make the statement: "heart: no abnormality". Inter-observer agreement regarding "normality" is not complete, although it is usually greater than agreement regarding "abnormality" [21]. Hence, it is important to make explicit what a physician means when making the statement "no abnormality". As argued in the previous section, we cannot expect a physician to make a statement about every possible item. To what extent can the meaning of the statement "no abnormality" be made explicit?

We are presently investigating wether or not it is feasible to define a default interpretation for the statement "no abnormality" for a large number of items. In daily practice, the physician would either make the statement "no abnormality" or only specify the findings that differ from those

specified in the default description. Since each physician has his or hers own style of history taking and physical examination, such a default would have to be tailored to the individual physician. The question, therefore, is whether there are strategies, such as a physician-specific default for the statement "no abnormality", that stimulate physicians to explicitly record findings that cannot otherwise be inferred.

Descriptions Involving Combinations

Many items in the patient record are characterized by a number of aspects. For example, a cardiac murmur has a loudness, pitch, relation to the heartcycle, shape (e.g. crescendo), and location. The item becomes specific for the patient when these aspects are filled in with the findings in that patient. Since physical examination represents the condition of the patient as it is observed at a specific moment in time, only one description is applicable for each aspect of a finding. In contrast with a momentary observation, the patient's history covers a certain time span and so do the descriptions of complaints. Hence, a symptom may have different expressions over time. In other words: one description may apply to a symptom in period A, while another description applies to the same complaint in period B. Let the symptom dyspnea be described by the aspects: season of occurrence, frequency, and severity. A patient may suffer from frequent attacks of severe dyspnea in spring, whereas in winter attacks of dyspnea are mild and rare. Here, the combination of aspect descriptions has medical meaning: they point into the direction of an allergic disease. The information conveyed in the combinations "spring - frequent - severe" and "fall - rare - slight" would be lost when the data would be stored as: "spring and fall - rare to frequent - slight to severe". It is important that different expressions of the same symptom can be explicitly represented in the computer-based patient record. In other words, it must be possible to make explicit which descriptive statements belong together, denoting a specific expression of a symptom. At present, we are exploring how to explicitly elicit this type of information from the physician in a natural way.

Multiple Occurrence of Concepts

Some concepts in a patient record constitute body parts, symptoms, or findings that are always singular. Examples of singular body parts are the heart or the liver. Examples of singular symptoms and signs are cough and bowel sounds, respectively. Other concepts may have multiple occurrences: paired body parts and findings such as skin lesions and enlarged lymph nodes. When a physician describes a patient's cough and she returns to that topic later to add some more information, it is obvious for the physician that she is adding information involving that same cough. However, when a physician describes a skin lesion and returns to the topic "skin lesion" again, it is not evident, unless made explicit, wether she wants to describe an additional skin lesion or add information to a previously described skin lesion. In a paper medical record such information is often conveyed in the spatial arrangement of the notes and can be inferred by the physician. However, when data in a computer-based patient record are to be subjected to research or to interpretation by other programs such as critiquing or expert systems, the presence of more than one skin lesion has to be made explicit and each skin lesion must be explicitly linked with the proper description. Hence, knowledge about the possible multiple occurrence of concepts plays an important role in minimizing ambiguity in a computer-based patient record.

However, there is more to multiple occurrence of concepts than the aspects just mentioned. For example, is cough always a singular complaint? Assume, we have a patient who presents with a cough, that has been dry and itchy since several years, but that became productive during the last two weeks. We can speak of one cough that changed in time. However, if the patient is a smoker with a frequent, dry, itchy cough and he contracts a bronchitis, we may prefer to speak about a second cough superimposed on the first one. When we bring into mind Weed's problem oriented medical record [22], we may want to link one cough with the patient's habit of smoking and the other with his bronchitis, or we may prefer to link the combined cough temporarily to both the smoking habit and the bronchitis. The question is whether or not it is possible to define which criteria should be used to determine when to treat a concept as being singular.

The Problem of Multiple Views

Many terms in medicine play a role in more than one context: "Frequency of urination" may be important in the history taking of the urinary tract, circulatory tract, or endocrine disorders. Hence, data specified at different locations in the patient record, make up the information about urination. "Urination" is an example of a singular concept. In general, every description pertaining to a singular concept always involves that same concept. This is evident for a physician. Hence, when dealing with such a concept in a computer-based patient record, the physician expects all information with respect to that topic to be presented and accessible, irrespective of the steps that the physician took to arrive at that topic. In other words, the semantic coherence of descriptions, pertaining to a singular concept must be maintained, while access to the information is always available from any appropriate medical point of view. Rector also recognizes the importance of supporting "perspectives" that allow items to be viewed in different ways [23]. We are currently investigating the potential of using knowledge about concepts being singular to infer when descriptions pertain to one item.

SUMMARY

In the early eighties, the goal set for the development of computer-based patient records was the creation of patient records that were analogous to the paper record, which was considered a requirement to make computer-based patient

record keeping acceptable for physicians. In the Netherlands, this strategy has been proven successful. In the years following the introduction of the computer-based patient record, expectations with respect to the use of these records evolved and expanded rapidly. The conviction gained ground that the computer-based patient record did, indeed, open up new possibilities for research by efficient retrieval of information, for sharing of patient data, for automated audit, for the integrated use of diagnostic expert systems, and for quality assessment. Although these "paper-like" computer-based patient records were suitable for patient care, they were much less suited for other purposes. Experiments showed that the use of data for other purposes than those for which they were recorded, can only be performed reliably when these data permit unambiguous interpretation.

We believe that research is needed to find formal notations for patient data that allow unambiguous interpretation of that data. Unambiguous interpretation of patient data requires these data to be complete and explicit. Due to the limited time of physicians there is a constant tension between benefit and effort, that has led us to conclude that we cannot expect the physician to provide the required additional information. Many of the inferences made by physicians are based on general knowledge about medical concepts and do not require specific, patient related information. Hence, we have focused our research on the potential of using knowledge about concepts in the patient record, to infer parts of the information required for unambiguous interpretation. In the Netherlands, with the relatively large and still growing acceptance of computer-based patient records in daily practice, the circumstances for experiments are fortunate. We are exploring the limitations of automated inference of information based on patient data as part of possible strategies to elicit a maximum of information with a minimum of effort from the physician.

References

[1] Landelijke Huisartsen Vereniging. Huisarts en Automatisering. Utrecht, september 1991.

[2] Werkgroep Coordinatie Informatie Automatisering. Referentiemodel Huisarts Informatie Systemen, versie 1985. Nederlands Huisartsen Genootschap, 1985. (Working Group on Coordination of Information Automation. Requirements Information Sytems for General Practitioners, version 1985. Dutch College of General Practitioners, Utrecht, The Netherlands, 1985. In Dutch.)

[3] Westerhof HP, Boon WM, Cromme PVM, et al. Elias: Support of the Dutch general practitioner. In: Reichertz PL, Engelbrecht R, Picollo U, eds. Present Status of Computer Support in Ambulatory Care. New York: Springer-Verlag, 1987:1-10.

[4] Lamberts H, Wood M, eds. International Classification of Primary Care. Oxford, Oxford University Press, 1987.

[5] Institute of Medicine. Committee on improving the Patient Record. The Computer-Based Patient Record: An Essential Technology for Health Care. Washington DC: Natinal Academy Press, 1991.

[6] Branger PJ, Duisterhout JS. Electronic Data Interchange in Medical Care: An Evaluation Study. In: Clayton PD, ed. Proc. 15th SCAMC. New York: McGraw-Hill, 1991:58-62.

[7] Lomas J, Anderson GM, Domnick-Pierre K, et al. Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. New Engl J Med 1989;321:1311-1315.

[8] Lomas J. Words without action? The production, dissemination and impact of consensus recommendations. Am Rev Publ Health 1991;12:41-65.

[9] Van der Lei J, Musen MA, Van der Does E, et al. Comparison of computer-aided and human review of general practitioner's management of hypertension. The Lancet 1991;338:1504-8.

[10] Van der Lei J, Musen MA. A model for critiquing based on automated medical records. Comp Biomed Res 1991;24:344-78.

[11] Mc Donald CJ, Tierney WM. Computer-stored medical records: Their future role in Medical Practice. JAMA 1988:259:3433-40.

[12] Shortliffe EH, Tang PC. Patient records and computers. Ann Intern Med 1991;115:979-81.

[13] Mc Donald CJ, Barnett GO. Medical-record systems. In: Shortliffe EH, Perreault LE, eds. Medical Informatics: Computer Applications in Health Care. Reading MA: Addison-Wesley, 1990:181-218.

[14] Van der Lei J. Editorial: Use and Abuse of Computer-Stored Medical Records. Meth Inform Med 1991;30:79-80.

[15] Mc Donald CJ. Action-Oriented Decisions in Ambulatory Medicine. Chicago: Year Book Medical Publishers, 1981.

[16] Van der Lei J, Van der Heijden P, Boon WM. Critiquing Expert Critiques: Issues for the Development of Computer-Based Monitoring in Primary Care. MEDINFO 89. Amsterdam: North-Holland Publ Comp, 1989:106-10.

[17] Quaak MJ, Westerman RF, Van Bemmel JH. Comparisons between written and computerised patient histories. BMJ 1987;295:184-90.

[18] Reiser J. The Clinical Record in Medicine. Part 2: Reforming Content and Purpose. Ann Intern Med 1991;114:980-85.

[19] Whiting-O'Keefe Q, Simborg DW, Epstein WV, Warger A. A Computerized Summary Medical Record System Can Provide More Information Than the Standard Medical Record. JAMA 1985;254:1185-92.

[20] Kent DL, Shortliffe EH, Carlson RW, Bischoff MB, Jacobs CD. Improvements in Data Collection Through Physician Use of a Computer-Based Chemotherapy Treatment Consultant. J Clin Oncol 1985;3:1409-17.

[21] Koran LM. The Reliability of Clinical Methods, Data and Judgments. N Engl J Med 1975;293:642-46 and 695-701.

[22] Weed LL. Medical Records, Medical Education, and Patient: The Problem-Oriented Record as a Basic Tool. Cleveland OH: Case Western Reserve Press, 1971.

[23] Rector AL. Helping with a humanly impossible task: integrating knowledge based systems into clinical care. In: Hanmu J, Seppo L, eds. Proc. 2nd Scandinavian Conference on Artificial Intelligence, SCAI '89. Tampere, Finland, 1989:560-72.