A.L. Rector

Medical Informatics Group,
Department of Computer Science,
University of Manchester,
Manchester, United Kingdom

Introduction

Medical records continue to be at
the centre of developments in Medical
Informatics. Standards bodies struggle
with definitions and interchange stan­
dards; controlled vocabularies and ter­
momologies continue to be sources of
both controversy and research issues;
how best for clinicians to interact with
medical record systems remains a cen­
tral issue; commercial vendors are in­
creasingly offering
"clinical
worksta­
tions", but few yet support anything
like the richness of a complete medical
record.

Underlying all this activity are two
questions:

1. To what extent it is practical -or
even possible- to capture, represent
and present electronically the infor­
mation now contained in manual
records;
2. What is the value -in improved
health care or cost savings- of so
doing and what are the potential
disadvantages and hazards.

If we focus more narrowly on the
technical problems, we might analyse
the issues to be faced into a six-stage
sequence:
1. Capturing the information;
2. Representing the information;
3. Sharing the information;
4. Presenting the information;
5. Evaluating the process of manag­
ing information;
6. Evaluating the outcome on patient
care.

To this linear progression through
stages we must add four additional is­
ues which are nearly orthogonal to it:
1. Fitting into the social and organiza­
tional context of care;
2. Being sensitive to patients’ reac­
tions and concerns;
3. Exploiting the possibilities of the
technology and coping with its limi­
tations;
4. Integrating into the broader techno­
logical framework of health care
information systems and standards.

Any given study -any given research
program- can deal with only a limited
number of these issues. Studies of
medical records per se tend to concen­
trate on the core issues of capturing
and representing information and
evaluating the quality of the informa­
tion and the acceptability of the pro­
cess.

Structured Data Entry and
the Computer in the
Examination Room

The papers by Moorman [1], Ham­
er [2], Kalra [3], and Solomon [4] all
deal with data capture and its evalu­
ation. Between the four studies, six
major criteria were used for evalua­
tion: completeness, accuracy/consen­
sus, availability, speed, user accept­
ability, and patient acceptability.

Solomon’s paper is unique in dis­
cussing the issues of patient accept­
ability which is still a worry to many
doctors. Despite the study’s admitted
limitations, it is gratifying to find fur­
ther data suggesting that patients are
not disturbed by the use of computers
during the consultation. Solomon con­
siders a number of different areas in
which use of a computer might be
expected to disrupt the physician-pa­
tient interaction and finds no support
for any of them.

The other three papers deal with
structured data-entry systems for di­
rect use by clinicians. (Studies of the
use of natural language for data cap­
ture, e.g., [5, 6], are not included in this
section.) Moorman’s study is an evalu­
ation of a system whose technical de­
scription was published last year [7].

The other two papers combine tech­
nical description and evaluation. Each is
of a slightly different task - Kalra’s
paper formally concerns an “integrated
assessment system” rather than “medi­
cal records” - illustrating the difficul­
ties of drawing the boundaries be­
tween which studies constitute “Medi­
cal Records” and which decision sup­
port or some other topic.

Hammer’s and Kalra’s studies con­
centrate on speed, acceptability and
completeness, while Moorman’s con­
centrates more on coverage. Interest­
ingly, speed of use is now so widely
accepted as a key aspect of user ac­
cceptability that it is often reported sepa­
rately from other factors. Both Ham­
ner and Kalra found the systems faster
than the alternatives and well accepted.
Hammer, in particular, emphasises the
potential incidental benefits of the sys­
tem: improved access to the literature
and to definitions for coding using
DSM-IV. It is also worth pointing out
that the decision-support systems cited were passive rather than active; the ability to access definitions and summaries of past experience rather than to obtain direct advice.

Kalra’s study is also noteworthy for its use of a portable pen-based system. Many of us have long thought that pen-based computing would have important consequences for clinical use. Even without touch screens, the easy availability of portable computers has changed the practical considerations for introducing electronic patient records to situations where physicians are mobile and go to patients rather than remaining in a consulting room where patients come to them. However, touch-screen hardware has often proved disappointing. It is to be hoped that the absence of any discussion of hardware issues in Kalra’s paper means that the hardware is finally proving satisfactory.

Studies of the Coverage of Compositional Representations

Both Kalra’s and Hammer’s studies were performed in the context of a well-defined target for data collection, the international classification of psychiatric disorders DSM-IV for Kalra, and a series of measures of impairment and morbidity for stroke for Hammer. Both might be regarded as special situations in which the line between “medical record” in the usual sense and “diagnostic” or “psychometric” instrument is blurred.

Moorman’s paper, by contrast, is an attempt to study more usual routine medical records, albeit in the restricted context of reporting the results of endoscopic studies. It concentrates on completeness and coverage, and takes as its starting point the natural-language reports written independently of the computer. Furthermore, it deals with a compositional formalism rather than the atomic coding schemes used in the other studies in this group and, hence, it has greater ambitions in what it seeks to represent. In this respect the Moorman study is closer to our own study using the GALEN and PEN&PAD formalisms [8] (see below).

Several points from Moorman’s study are worth noting. Unsurprisingly, there is greater consensus among the physicians when reporting using the structured reports than when reporting using free text, as indicated by the greater number of endoscopists describing each feature in the structured reports. This is useful confirmation of our intuition that, given a fixed structure which constrains and cues the physician, physicians record information more consistently than when writing in free text guided only by a largely unwritten framework. Note also that in the free-text reports, “unclear” feature descriptions accounted for nearly 10% of the total. Our own experience suggests this is relatively low and may reflect the fact that endoscopy is already a highly constrained domain. Overall, the evidence is accumulating that structured data entry improves consistency and consensus in data capturing, although worries remain about its expressiveness. Worries also remain that physicians may be forced to distort information by entering it in inappropriate categories.

The other side of the issue is to determine what information is lost in structured data entry; what information could not be recorded. Moorman’s paper gives some indications. Most information was recorded successfully. Of the rest, roughly half required only simple additions to the underlying model.

The fact that only minor additions were needed to accommodate new information leaves open the question which a brief study cannot answer, namely how would the rate at which simple additions were required decrease over a long period of use? Would the rate of new additions ever fall near zero? Experience in the natural language community suggests that the requirement for new lexical items declines rapidly initially and then remains relatively constant. Whether or not this is true for structured representations such as Moorman’s or GALEN’s remains to be seen.

Moorman raised two quite different considerations. Firstly, there is the issue of what should be considered “inherent” in a concept, e.g., since erosions are, by definition, superficial is the phrase “superficial erosion” therefore redundant? (This issue is also a central focus of GALEN’s concerns with normalisation of concept representations and bridging levels of granularity [9].) These issues require the ability to manipulate the logical structure of definitions and descriptions. The second issue is that of alternative coordinate systems and units as illustrated by the request of some endoscopists to use “hours” on a clock face rather than the anatomical descriptors “left”, “right”, “front”, and “back”. Such problems require not just the ability to cope with descriptions but also the ability to perform various calculations and transformations.

Our own study [8] was part of a workshop exercise organised by the CANON group [10] which gave rise to a number of other studies and commentaries [11-14], some of which appeared in last year’s Yearbook. It dates from an early stage in the development of GALEN. Like Moorman’s study it concentrates on coverage of a compositional formalism by comparison with naturally occurring free-text records. By contrast to Moorman’s study, it attempted to construct the complete theoretical representation rather than constraining itself to what could be entered by a particular structured data-entry system.

In so doing, it raises two groups of
further issues. Firstly, there are issues of the expressiveness of the formalism itself. These include the representation of negative findings such as "absence" or "without", the handling of temporal relations and references to previous (or subsequent) events within the record, the representation of uncertainty, and how (or whether) to distinguish between observed manifestation and the inferred condition of the patient, e.g., between the "opacity" on a radiographic film and the "atelectasis" in the patient's lungs. Approaches to each of these issues is presented, but none can be considered definitive.

Furthermore, each issue dealt with adds to the complexity of the representation. Often the additions add global complexity even though they arise relatively rarely. For example, to deal with the possibility of negative findings using the form "absence which isStateOf...", an extra level must be included in all findings which is usually the seemingly redundant wrapping "presence which isStateOf...". (Note that while an implementation might choose only to store the "absences" and assume the "presences", the formalism must have both if it is to be consistent and avoid anomalies such as retrieving an "absent ulcer" as a kind of "ulcer disease".

The second set of issues are those of implementation and appropriateness. Despite the emergence of object-oriented databases, most medical records continue to be implemented on relational databases which require predefined fixed structures. Descriptive formalisms such as Moorman's or GALEN's fit such systems poorly. Either they must be constrained to limit the depth and length of expressions, or special mechanisms must be developed. The emerging CEN standards on medical records represent one attempt to provide a modest restriction while allowing expressivity which is believed adequate for most purposes; GEHR [15] represents another related such effort. GALEN's approach of a terminology server [16] represents an alternative approach of encapsulating descriptive complexity so as to hide it from relational systems.

The broader issue of how much complexity it is worth implementing can only be answered over more time with experience than is covered in these papers. Those of us advocating more complex systems must show that we can produce applications with them which would not have been possible without.

### Presenting Information

The final paper in this Section, by Goldschmidt [17], reports a novel means of presenting the information in medical records, specifically the laboratory findings and functional test results. Goldschmidt's technique attempts to present information in a way which makes effective use of clinicians' visual pattern recognition. The information is presented in a circular array in which different patterns of anomalies give rise to strikingly different shapes. Like the first of the Moorman papers last year, it presents the technique but does not provide an evaluation of its practical use. Of particular interest in this case will be information on the training and experience required to learn to recognise the patterns generated. We await such further evaluations with interest.

In the mean time, the most important aspect of the paper may be that it breaks new ground with a radically different means of presenting information, which could only be possible with computers. To date, most presentations are extensions of how we present information in manual records. There is no reason to believe that these techniques, despite their familiarity, will be the most effective possible. The use of computers presents opportunities for radical new presentations using shape, colour, three dimensions both in abstract forms such as presented here and in more representational pictorial, diagrammatic or iconic forms. Some will turn out to be "just pretty pictures". Others may give physicians a much more powerful way of apprehending patient information quickly.

### Conclusion

Data capture, representation and presentation are each moving forward. It is becoming clear that physicians will use structured data entry if it can be made convenient, quick and available, at least in certain situations. It also seems likely that if physicians do use more structured methods, they will record data more consistently and with less ambiguity than otherwise. The price is that some things may not be able to be said or may require supplementary free text or dictation.

What is also clear is that complete representation of the clinical notes with their full complexity of time, uncertainty, evidence, causality, etc., is a difficult task which will take time and tax the capabilities of existing data-management systems. A balance of immediate application, long-term development, and careful construction of migration pathways from existing to new systems will be required before the potential of electronic health care records is achieved.

### References


Yearbook of Medical Informatics 1996

Acknowledgement: I am grateful to be asked to summarise this year's papers, but reticent about the fact that one of our own studies is included. Readers should be aware that I can claim greater knowledge but must admit to less objectivity with respect to that work than would an independent reviewer.

Address of the author:
Alan L. Rector,
Medical Informatics Group,
Department of Computer Science,
University of Manchester,
Manchester M13 9PL, United Kingdom.
email: rector@cs.man.ac.uk
URL: http://www.cs.man.ac.uk/mig/