Bernt I. Olsen, Niels W. Lund, Johan G. Bellika, Eirik Årsand,
Per Hasvold, Gunnar Ellingsen, Alexander Horsch,
and Gunnar Hartvigsen

Documents in Medicine: from Paper Documents to Quality-Healthcare?

1   Introduction

The US, as well as most Western nations, is experiencing a transformation from paper-based health documents to digital documents. In this study, we will explore how and why these changes are happening and aim at analyzing what the transformation from paper to electronic documents in healthcare might bring. We are also concerned about how this transformation affects our understanding of the notion of "documents", and in particular "medical documents".

The current replacement of paper-based information sources in health care has several causes. The development of new treatment methods and technological innovations as well as national health policy initiatives are some of them. These generate enormous amounts of information. For instance, the database Medline[1] grows by about 30.000 new publications each year, requiring specialized systems to organize all this information. Large hospitals also face this challenge. For instance, Haux (1998) argues that University Hospital of Heidelberg, with its 1700 beds, creates about 400,000 new medical records per year. These records contain 6.3 million pages and require 1.7 km of storage. Physicians create over 250,000 reports and 20,000 procedure reports each year. The service departments create around a million medical results.

These problems underscore that the traditional paper documents are insufficient in the modern healthcare system. Thus, future electronic documents will have to be dynamic and flexible, fully capable of supporting intensified documentation processes.

---

[1]   Medline is a bibliographical database with references to journal papers in medicine, nursing and odontology.

In this study, we will present some of the situations in which medical documents are produced. In particular, we will elaborate on how digital documents make a difference compared to paper based documents. To get a feel for what these documents are and will be in healthcare and how they actually come about, we will proceed along the following questions and lines:

- What is a document in healthcare?
- What makes medical information and medical information systems special?
- We will look at some specific cases/genres of documents:
    - Medical images
    - Collaborately created documents
    - The role of disease surveillance
    - Medical sensors

2   What is a medical document?

Interestingly, the terms "medical document" and also "clinical document" are frequently mentioned within the medical literature, but searching for some explanation on what actually lies behind the terms yields little information or discussion on what is meant by them. "Clinical" means direct medical treatment of patients, as in what is observed by a doctor: does the patient have a temperature, is he/she pale, does he/she have a fast pulse, etc. A clinical document would therefore describe clinical findings about the patient. "Medical" is, by the *Oxford English Dictionary's* definition: "of, relating to, or designating the science or practice of medicine in general, or its practitioners". From these two definitions one can argue that clinical documents are documents created during care, while medical documents are knowledge-containing documents used to implement care. It would seem that the "medical community" is not engaged in the discussion of what is meant by identifying something as a document, although the term is frequently used. Furthermore it seems that the meaning behind the words "document" and the process of "documentation" seems every bit as much blurred and ambiguous as it is in contexts other than medicine. The word "document", in its medical context, does not appear problematic in any particular way and apparently there exists a very good consensus about the meaning of it. In order to get an impression of what is considered as a document in healthcare let us first take a look at how the term is used. Please note that the following examples are quite randomly selected from the medical (informatics) literature.

The American Society for Testing and Materials' (ASTM) committee E31.25 was established "to develop standard electronic document representation of paper-based healthcare documents and forms", using XML for the purpose. In this con-

text it is interesting to see their mention of "documents", exemplified in the following citation:

> In order to develop an electronic medical record, an electronic representation of the paper documents must be determined. However, standard description of the types of paper documents in medical records does not exist. A set of the types of documents in the medical record needs to be derived so that electronic representations of the documents can be developed (Sokolowski and Dudeck 1999, 148).

"Documentation", which generally concerns the process of creating a document, is in medicine very much associated with evidence of medical care (or healthcare), or the "legally" influenced report on how care was provided in a specific case (one patient) and is often collocated with the word "malpractice". The following is a quote from an article in the *Journal of Family Practice Management* authored by Teichman:

> Excellence in medical documentation reflects and creates excellence in medical care. At its best, the medical record forms a clear and complete plan that legibly communicates pertinent information, credits competent care and forms a tight defense against allegations of malpractice by aligning patient and provider expectations (2000, 1).

Teichman further states that physicians typically document patient care with a goal of effectively communicating with themselves. This is especially true for primary care like family practices. In other situations of patient care such as hospital settings, home care and nursing homes, this might also be the case, at least for example with patients with chronic illnesses, but within secondary and tertiary care, the documentation contains more communication between healthcare professionals. We will here make clear the distinction between the notion of document and the notion of documentation. In the case of the notion of document, it means the discrete limited number of data appropriate for a certain case, for example a consultation at the GP's office, which is the main focus. In the case of documentation, it is the recording of the whole process, each step in a treatment/care of a patient which is the focus.

We note the use of the term "record" in the previous quotes. In the medical setting "record" and "document" is related, "record" is typically used in the "Patient Record", the "Medical Record" or "Health Record". In other words, it would seem that the collection of documents concerning one patient's care or health is typically regarded as a record within healthcare. In his book on information retrieval (IR) within healthcare, Hersh points out in his definition of an IR-system that:

> An IR system consists of content, computer hardware to store and access that content, and computer software to process user input in order to retrieve it. Collections of content go by a variety of terms, including database, collection, or – in modern Web parlance – site. In conventional database terminology, the items in a database are called records. In IR, however, records are also called documents, and an IR database may be called a document database (2003, 4).

This is an interesting remark, even though we are not specifically interested in IR in this context, as it brings some clarification to the relationship between "record" and "document" – especially since the author of the book is talking about IR within healthcare. Records are, in database terminology, defined as related data values or items physically stored together. In a relational database the equivalent might be a row (a tuple) in a table. This is interesting because it says that a document is a (small) set of related data and values, that create some meaning when grouped together. When a medical or health record is a collection of "documents" we could interpret this as "record" being the broader term, while "document" is the more narrow and specialized unit of information. This is how we could interpret the meaning behind the two terms in a medical setting.

Following this trail of thought, the medical documents are seen as parts of medical records and are "instantiated" either in an ad hoc manner, i.e. like narrative text, or according to some predefined structure, a template or the like. As for what goes into clinical documentation, quotes from the guidelines for *Good Medical Practice* (2006) by the General Medical Council provides a fair understanding:

> In providing care you must:
> […] keep clear, accurate and legible records, reporting the relevant clinical findings, the decisions made, the information given to patients, and any drugs prescribed or other investigation or treatment (8);
> […] Sharing information with other healthcare professionals is important for safe and effective patient care (25).

This quote points to two of the key elements of medical documentation: to record and document findings and actions and to inform colleagues or others who are involved in the shared care of the patient. From this we can also distinguish between the types of documents that store a documentation of objective observations (such as measured temperature, weight, height, etc.) and more subjective data, such as possible diagnosis and assessments of the medical condition of the patient, and the types of documents that are used for sharing care (such as referral notes and laboratory reports).

An important aspect of medical documentation is that it is created, maintained, and protected by the medical professional, but contains information about the patient. The patient has wide rights regarding insight into what is contained in the medical documentation, but cannot change or influence what is in the documentation directly. Because of the sensitive nature of medical information, a great responsibility lies on those who are responsible for the integrity and accessibility of the medical information. Access to medical documents is controlled by legislation that specifies who can access these documents and under what conditions. Such legislation usually also places a responsibility on the healthcare professional to not abuse the information and to maintain the privacy of the patient.

The use of medical documentation is not limited to the diagnosis and treatment of individual patients, but also plays an important role in monitoring practice (both in auditing one's own practice and reaching consensus on the best practice) and assessing population-wide trends (such as the outbreak of an epidemic, see section 4).

Health Level 7, an American non-profit organization, has in its Clinical Document Architecture (CDA) made an effort to standardize electronic clinical documents for exchange between systems and organizations using an XML structure. Having the "document" as part of the label, it is difficult to have a discussion about medical or "clinical" documents (or documentation) without taking a closer look at it. The CDA specifies the structure and semantics of a clinical document. A CDA is defined as follows:

> A CDA document is a defined and complete information object that can include text, images, sounds, and other multimedia content. It can be transferred within a Message, and can exist independently, outside the transferring message (Dolin et al. 2006, 1).

The CDA is mainly developed for interoperability purposes, making institutions able to exchange documents such as discharge summaries or progress notes. The CDA is "richly expressive" and contains text, images, sounds or any other multimedia content and is thus descriptive in terms of what we can expect from future health information systems. The CDA's mission is both to make information structured and interoperable and to make it interpretable by machines, so that they can be acted upon – by machines (like issuing alarms, reminders, etc), hence the adoption of XML for this purpose. The CDA is, like all standards from HL7, based on a framework called Reference Information Model, RIM. As the name suggests, this is a model of the grammar for the HL7 "language", i.e. it gives the basic building blocks of the language and the permitted relationships between them. The CDA is now being adopted quite widely within research as well as

within clinical information systems and is a promising alternative when documents are being digitalized within healthcare.

The most interesting characteristic of medical information is the complexity inherent in such information. Since the 1970s, the field of Medical Informatics has had the creation of a worldwide exchangeable and interoperable medical- or health record system as a research goal, but it still remains distant. The issue of complexity, it is argued, is the main cause of the lack of such an electronic health record. The complexity of patient information is due to the fact that medicine in general mirrors the most complex organism that we know of, namely the human body. Furthermore, systems handling this information must meet requirements from a social and societal point of view (such as legal requirements etc.). Beale, in his comment in the IMIA Yearbook of 2005, uses the eight levels of complexity in living systems proposed by Miller (1995) to describe the task of making a information system, or an EHR, to support the task of giving health care, arguing that the complexity of life is relevant for clinical information systems at all eight levels – ranging from the micro-system level of the cell to the macro-perspective of the "super-societal" (global) level. Beale here argues strongly that the health information system differs from other systems in that the "complexity of life" it examines exists at all eight levels, and cannot easily be simplified by abstractions, but that the human patient needs to be captured in his/her entirety; any small detail might be important for care-giving and treatment purposes. This stands in contrast to other information systems that HISs are commonly compared to, like banking systems or ticketing/air travel reservation systems. In these systems, it is only a fraction of human activity that needs be modeled for a successful system. In addition to the complexity of the human body and the social and societal issues, the continually evolving medical knowledge base (e.g. Medline) creates further complexity, describing both the results of centuries of medical scientific achievement – an already overwhelming base of input to clinical decision making - along with an accelerating rate of new knowledge, crucial for success in research-intensive areas like oncology. Beale lists some interesting requirements for the health record:

- information and efficient user interface reflecting multiple levels of hierarchical biological and social organization;
- mobile patients;
- longevity of information (e.g. 100 years);
- multi-lingual;
- data shared and authored by multiple users simultaneously;
- integrated with knowledge bases such as terminology and clinical guidelines;
- wide geographical availability of a given record to multiple caregivers and applications;

- consent-based, potentially fine-grained privacy rules on information use (with exceptions for emergency access);
- multiple sources of constant change to requirements including medical technology, clinical procedures and guidelines, genomic/proteomic medicine;
- reliable medico-legal support for all users (Beale 2005, 302).

In addition to these requirements, one might also add the already mentioned potential size of the record, as some of the compounds might require terabytes of storage. However, as Beale points out, in all of the above requirements there might be at least one application or system more exigent than the EHR. The problem is that medical information needs encompasses multiple areas of ICT at once. As health care is provided, the assisting information systems need to perform at several of Miller's eight levels of complexity all at the same time, for different users with different needs. An example of the difficulty caused by this is the problems faced in the design of user interfaces for medical applications. Electronic information systems, in contrast to older (paper-based) systems have the potential to inform caregivers in almost any situation of nearly every aspect of the medical decision process. The risk of overloading the user with information, putting too much into one interface, is imminent and tempting for system designers. The problem at this point is: how does one know what is relevant at any point in time? How does the GP know what is relevant when referring patients to secondary care? Only the specialist, the recipient of the information, would know what to extract. Many applications fail at this point to attract users in clinical settings as the requirements for both user-friendliness and the felt usefulness of the application are extremely high. Health care workers are generally people with very little time to waste with systems that fail to inform them efficiently and are the first to stop using them (or refrain from starting to use them). Having the definition of HL7's CDA freshly in memory, the description of a "complete information object […] that can exist independently from a transferring message" perhaps preceded by the word "minimum" might seem compelling in this context. The first apparent task is of course to build systems to make all the information available, the next to find appropriate methods of selecting the information most relevant to any particular situation.

3 Images in medicine

As we have now taken a look at the general "nature" of medical documents, we will move on to take a closer look at some specific concepts and cases within modern healthcare and the potential within technology and the digitization of medical documents, and the issues that they raise.

In medicine, images play a central role in the concept of a document. Their counterparts in everyday life are photos, illustrations, sketches, drawings and other examples from art or mass-media. However, images in medicine have very distinct properties, as they most often serve a very specific use in treatment, research or within medical education.

In medicine, an image, or a series of images, a 3D volume, a video, or any other visually perceivable entity, together with a minimum set of meta-data, is considered to be a medical document. That is: anything from a photo or a video to a 3D-model together with describing "data" is considered to be a document. The minimum set of meta-data depends on the context in which the document is used (or in other words: the purpose for which the document is created). In the context of patient care, the minimum set of meta-data comprises all information that is mandatory to fulfill the requirements of health service provision. For this purpose, especially information about the patient, the acquisition device, and the producer of the images is included. A special case are teleconsultations, where images or videos are captured and stored for documentation of the consultation. The project ENDOTEL led by Munich University of Technology (Horsch et al. 2003) has used the CDA document standard for this purpose.

In the context of medical education, the minimum set of meta-data comprises information needed to make the images a learning unit. This includes information such as author of the learning unit, and case data (e.g. age and sex of the patient, medical problem, therapy regimes).

Within the context of medical research, the minimum set of meta-data consists of information required by the trial protocol. Patient-identifying information is in this context either anonymized or pseudonymized. A special case are reference datasets for the validation of computer-based image analysis tools, such as computer-aided diagnosis tools: here, an image or imaging study together with meta-data needed for assessment of quality features (e.g. accuracy, reproducibility, robustness) can be considered a document (Horsch et al. 2005).[2]

Some examples may help to understand this definition. DICOM (Digital Imaging and Communications in Medicine) (DICOM, 2006) studies are documents in the context of patient care. A DICOM file representing a DICOM study on a computer includes a large header part containing detailed information about the patient, the health service provider that has performed the study, the imaging device, and the image acquisition parameters. ODITEB (Open Distributed Textbook, an eLearning application for tumor diagnosis with x-ray, CT, MRI, and endoscopy)

---

2   The bare pixel data of a digital image is not a document in medicine, due to the fact that it lacks necessary information that makes it a self-contained information unit.

cases are documents in the context of medical education (Horsch et al. 2000). Such an eLearning case comprises imaging studies together with non-identifying (anonymous) information about the patient and the case such as anamnesis, or an expert guided tour through the case. Mammo-iCAD (a current research project on computer-aided diagnosis for breast cancer detection of Definiens Inc., Munich, in collaboration with University of Erlangen and Munich University of Technology) cases are documents in the context of medical research. Such a case contains the breast images (mammograms) and a subset of patient data from the standard documentation used in mammography. In all these examples there are core meta-data which are essential for a document, including author(s), creation and modification date and time, and the property of being a self-contained information unit.

Documents for education and research are usually derived from documents which have been created in a patient care setting. For instance, in order to create an ODITEB eLearning case, DICOM studies are anonymized and non-identifying information from the patient record is taken and worked upon, providing a learning case for medical students.

In practice, medical images occur as parts of documents. So, looking at practice-relevant information units in medicine, one can always perceive the bare image without any meta-data as a part of a medical document. From this perspective, a DICOM image is a document with a structure defined by the DICOM standard, which contains an image or a series of images as a part of the document. Printed documents such as reports from a hospital to the primary care physician often contain images for illustration purposes (not for diagnosis). Information systems for pathology, microbiology, ophthalmology, etc. store visible light images and meta-data in a database, e.g. in an EPR system or a clinical trial system.

4 Multiple authored documents in healthcare

The medical domain is collaboration intensive by nature; many different participants are required in the act of caring for a patient. In hospitals, different kinds of clinicians bring their specialty area forth in the effort to give care to the patient, while in other settings such as home care and nursing homes several people share responsibilities in caring for multiple people. Within primary care, the most typical encounter in public health care (which accounts for about 90% of patient cases) cooperation also plays an important role. Documentation is a way of ensuring that someone else can easily take over care for the patient (i.e. secondary care, or another GP). Other than in the case of family practice kind of primary care, documents are rarely created in a medical context for only one person to see or use, and hence documents are both means for cooperation and artifacts that encourage cooperation, as we shall see. The role of cooperation within healthcare is

strengthened by a current and worldwide emphasis on "patient-centric care" and "quality of care". Concepts as teamwork, the sharing of care for patients between clinicians, are regarded as crucial success factors in achieving increased quality of care and an improved health for the patient (Dichter 2003; IOM 2001, i.e. 27; Olender 2005). Shared care refers to clinicians who jointly treat the same patient, while continuity of care is the 'complete care' provided during all transitions, such as home to hospital, between institutions, and so on.

"Computer supported cooperative work" (CSCW) is a field of research that aims to understand the nature and characteristics of cooperative work so as to build adequate systems to support and enhance this cooperation. The nature of a collaborative activity is generally classified along the dimensions of time and space, that is – synchronous versus non-synchronous and same place (i.e. the people are in the same room) versus geographically distributed collaboration.

	Same time (synchronously)	**Different time (asynchronously)**
Same Place	*A meeting*	*A hospital ward*
Different Place	*Telephone call*	*Referred patient*

Table 1

All four modes of cooperation (see Table 1) are present when health care is provided in all kinds of medical issues and are traditionally supported in different ways. Documentation in a case of cooperative activity could be either conclusions made by the team or a report of a meeting. Cooperative activities spanning time and space have traditionally been mostly 'supported' by paper, either as paper journals (or parts of them) or forms and schemata, either transported between the cooperating parties or stored in a common place. We acknowledge that paper has a lot of qualities that make it well suited to support cooperative activities, as for instance medical records are small enough to be held in the hand while the doctor communicates with the patient, a nurse, or a specialist on the phone. The paper can readily be moved both within and between institutions, and has inherent (physical) limitations regarding unwanted distribution of sensitive data. Paper can easily be annotated for the next viewer or for later use. All of these qualities and many more are easily forgotten when documents are 'reproduced' or replaced by an electronic counterpart. But paper has also severe limitations when supporting collaborating parties or persons.

The introduction of information systems into healthcare has profound implications for the potential support and enhancement/augmentation of cooperation for the abovementioned quality of care issue and making the health of the patient at the centre of attention. The transition from paper to electronic documentation makes

long-distance exchange of information and communication manageable. Information technology can provide for tele-consultations, bringing expertise and specialists to work in areas where it was nearly impossible before. Secondly, in the future, we expect to have health record systems that allow for both cooperative documentation of the care given to patients, allowing for authoring and editing by all types of clinicians and healthcare workers, suiting their need, from wherever they find themselves caring for a patient. Cooperation via IT-infrastructure brings forth new possibilities and requirements for documentation of care. Systems will allow for constructive and easy co-authoring of documents while at the same time also allowing for communication and coordination of activities.

Electronic documents are especially useful in terms of availability. Electronic versions of documents can easily be distributed to several locations at the same time. People can share both information, thoughts and ideas and work on the same document at just about any given time (only limited by the availability of the people themselves). The ease of distribution/availability of documents means that more people can be included into the care regime, which means there are better grounds for interdisciplinary cooperation, which in turn can lead to better care. Even the patients themselves can be encouraged, through availability of data about themselves and their health condition, to actively contribute and cooperate with healthcare personnel. This can be done through web interfaces containing parts of the patient record that the patient can access from home. Examples of such contributions are sensory equipment at the patient's home, or sensors to be worn by the patient, as discussed in chapter 6.

Research within CSCW has thrown light on several aspects of cooperative activities that have profound implications for systems that are built for health care purposes. Agreeing that cooperation and collaboration are important for the quality of given health care and the effectiveness of health systems means agreeing that all of these aspects must be taken into account in designing systems for the healthcare domain. Documents have a special role when it comes to cooperative settings in caring for patients. They can, for example, be seen as "collaboration mechanisms", helping in both coordinating and structuring work. Both paper and electronic documents can act as such mechanisms; thus for example in the co-located situation there are documents that are collaboration mechanisms, such as when physicians discuss an MRI image, lab reports, or any other such document. There are similarities between the co-located and the distributed collaboration modes that relate to the role and importance of collaborating mechanisms. In health care, documents (including forms, records, images, sounds, medical data, etc.) play an important role in collaboration as healthcare services are bound by legal requirements to document all actions and assessments related to the diagnosis and treatment of a patient. Digital documents have opened up new ways of collaboration and organization of work flow. They further enable completely new ways of per-

forming care audits, both by allowing for more and better quality of documentation, and in allowing mechanisms for developing and controlling routines and protocols. The availability of electronic documents, along with communications tools, make them ideal for informing the cooperating people of each others' actions, creating 'awareness' within the cooperating party. Information technology and – systems can provide for quite outstanding support for cooperation in healthcare, but care must be taken when designing systems in order to use this potential to the highest possible degree. We are only starting to see systems designed with such issues in mind, and little empirical data exists on their use. The scientific body of knowledge on the subject of information systems supporting cooperation is perhaps still not matured to the level that the potential for 'augmented cooperation' can be fulfilled. There is still a lot to be learned about systems for collaboration and the documents that are to come from them.

5   Medical Documents for Population; 'aggregate' documents

At the other end of the scale from the patient-specific collaborately authored documents are population-specific issues. Today, disease surveillance is an important part of most countries' protection against epidemics and their global counterparts, pandemics. Most countries therefore have legislation that permits extraction of data from medical information systems for disease surveillance purposes. Disease surveillance today is dependent on specific 'documentation practices' and will likely remain dependent on a very wide range of information sources in the future.

An early example of the role of a document, and the documentation process, is the Broad street map, made by John Snow, a physician in the Soho area in London during the reign of Queen Victoria. From 1831 and onwards, London suffered from recurring cholera outbreaks that caused many deaths. Snow became interested in the cause and transmission of the cholera disease and used a map of the Soho area in London to plot the location of deaths caused by cholera. He used this map to show that there was a correlation between the water supply and deaths caused by the cholera disease. In his publication, "On the mode of transmission of cholera" (Snow 1965) from 1855, he argued that cholera was transmitted through contaminated water, against the more commonly held belief that cholera was transmitted through miasmata (i.e., bad air) (Frerichs 2001; Buechner et al. 2004). By looking at the map, everyone could see that the cases were scattered around the location of the Broad street pump. Snow managed to stop an outbreak, giving undisputable evidence for his theory, by convincing the authorities to remove the pump handle from the Broad street pump. John Snow is today regarded as a pioneer in epidemiology and is by some called "the father of modern epidemiology".

Today both we and the diseases that we are the carriers of travel faster and further than ever before. The SARS outbreak in 2003 showed how quickly a local outbreak of an unknown disease in Hong Kong spread to a lot of countries. The systems that are able to discover disease outbreaks have therefore become more important than before. Our ability to discover disease outbreaks is today limited by our access to digitally documented evidence of confirmed cases or digitally recorded documentation of actions that are caused by people affected by a disease.

Traditionally, it has been the medical laboratories that have been the data providers for the disease surveillance systems. The medical laboratories provide evidence for the presence of a disease from the samples that are provided by physicians. By counting the number of confirmed cases of a predefined set of diseases, typically in centralized institutions, we achieve some preparedness against the threat from contagious diseases. The record of previous cases provides data about the normal distribution of cases in the population. If the number of confirmed cases of a disease suddenly increases, we can discover the situation and potentially initiate a public health response to stop the outbreak.

A new threat which the disease surveillance system must face is bioterrorism. Bioterrorist actions may cause a large number of people to become ill simultaneously by spreading a disease agent through air, water or food. The fear of bioterrorist attacks has initiated research in syndromic surveillance systems (Henning 2004). The term "syndromic" refers to the set of early symptoms that characterize a disease. The goal of syndromic surveillance systems is to discover disease outbreaks as early as possible, hopefully early enough to initiate public health actions that can limit the effects of such actions. New data sources as school and work absenteeism, over the counter sales in pharmacies, and many other sources, have therefore become important. The most important source is the electronic patient record systems used by the health personnel that meet the patient. However, the text for documenting the chief complaints presented by the patient may cause problems related to misspellings, local slang and lack of a common terminology. If the routines for documenting a case are followed, and the documentation is performed immediately, this is a good data source for surveillance systems. However, the lacking utilization of electronic tools for documentation is limiting our ability to discover bioterrorist actions or naturally occurring disease outbreaks. To have an impact, syndromic surveillance systems may need to discover bioterrorist actions within the incubation period of the disease (Kaufmann et al. 1997, i.e. 88). A future source for syndromic and disease surveillance systems is therefore data provided by humans that volunteer data from sensors for disease surveillance purpose. A very relevant group are the diabetics that regularly measure their blood glucose level (Årsand et al. 2005). We will have a look at diabetes disease in relation to how sensory output is a part of medical documents later in the text.

Laboratory-based disease surveillance has limitations. It may take a long time to confirm the presence of a disease, we may not test for the correct disease in the laboratory, and it may take a long time before the data becomes available for the electronic disease surveillance systems, if the disease is unknown, how can we count occurrences of the disease electronically that lacks a classification? If the disease is new and unknown, as in the SARS case, we may not have a test that we can perform. We may therefore need to collect data about the symptoms of the patients, recorded by health personnel, and the category of tests that are performed to provide early indications of disease outbreaks. In the case of an unknown disease, we may lack a clear definition of the symptoms that we can use for screening, and the definition of the disease may change, based on updated knowledge which invalidates earlier definitions of the disease. Management of disease definition information is therefore also a challenge in being prepared for every possible outbreak of a communicable disease.

Another development that threatens our disease surveillance systems is the increasing availability and use of disease specific polymer chain reaction (PCR) tests by general practitioners. Unless the results of these tests are collected, our disease surveillance systems may become undermined from lack of case data. In the future we therefore need to develop disease surveillance systems that also collect case data from primary care physician's electronic health record systems.

Disease surveillance is a subject that requires thorough and accurate documentation, both in health care and, in the future, from other places. All sources of information are important in such a context, including other documents and documentation practices which, unlike today, are not specifically or explicitly made for the purpose of detecting epidemic outbreaks. They are differently sized 'streams' or 'rivers' of documents that altogether will contribute to the creation of powerful documents that detect epidemics or pandemics, hopefully early enough for us to take action. They will be powerful documents in that, like the map created by Dr. Snow, they will initiate actions against outbreaks or reducing the spread of diseases. However, finding signs of epidemic outbreaks is in many cases not trivial. It involves, in this context, searching through potentially large and distributed data-sets or documents. This approach is often labeled "data mining" within computer science. We will next briefly discuss this technique, which is regarded as very important in the future of health care.

As we have already discussed, patient databases and knowledge bases are growing rapidly within medicine and healthcare. With paper based information systems, this means growing mountains of paper documents that require a lot of storage space (patient records) and books and journals printed, delivered, indexed and filed for the use of medical decision making. With digital systems, it means growing patient databases and growing knowledge bases stored either in local systems

where the data/knowledge is produced, or in centralized settings on servers. The transformation to digital documents offers some possibilities that are not available using paper documents. With paper, the information stored is not easily available, as documents have to be tediously collected in order to be available for evaluation in medical decision making. Furthermore, the information has to be processed by humans in order to be useful. The affordances of electronic documentation in regards of availability of the information and the potential for machine processing is given by the technique known as data mining. In essence, it means browsing through large amounts of data, looking for relevant (new) information and knowledge. This means that one can have agents or processes running that look for i.e. specific patters (one technique) either on a specific case (patient) level or in larger groups of people.

The increasing complexity of the medical field calls for tools that can keep up with this development. Having patient records available online enables automatic search routines and pattern matching. For example, the content of a person's record can regularly be used as a search filter in order to locate relevant medical information in medical databases and/or on the Internet.

For difficult cases where we have symptom descriptions, but not have been able to form a diagnosis, automatic pattern matching against other patients can help in the identification of the correct diagnosis. More general algorithms can be used to find similarities in medical data. This will be similarities between patient's records, between patient's records and medical databases, etc. Based on such a data mining system we can uncover relevant relationships in medical data, both expected and unexpected. For a document, this means that it will live its own life. The extension of the document occurs automatically.

The digitalization of documentation processes within healthcare has profound implications for the amount of available data in the future. Another factor that will contribute to this volume of information is the growing number of (digitalized) medical sensors. These range from hospital equipment that delivers data directly to patient records to equipment worn by the patient or body-sensors. It is likely that we will see this data present in future health records, along with patient contributions in various forms. We will now have a look at a specific and very relevant disease that illustrates this: the case of diabetes.

6   Sensory input - automatic documentation

Many chronic diseases need to be closely followed up in order to achieve better health prognoses, and medical sensors constitute good tools for this. Wireless communication technologies have exhibited huge progress the last years, with

growth of facilities such as Bluetooth, ZigBee, WiFi, UMTS, and others. An increasing number of medical sensors support at least one of these communication standards, among them blood-pressure measuring systems, asthma peak aspiratory flow monitors and blood glucose measurement systems. Likewise, technological development has lead to miniaturized electronics that enable medical sensors either to be worn on the body with minimal inconvenience, or even to be implanted in the body. Based on one of the most common chronic diseases, diabetes, we will here briefly discuss both the state of the art and future perspectives of this technological progress, related to the change in medical documents and medical documentation.

The number of cases of diabetes worldwide in the year 2000 among adults (+20 years of age) was estimated to be 171 million (Wild et al. 2004). Most of the instances of the disease are defined as Type 2 diabetes (~92 %), where an unhealthy lifestyle is the main contributor to the imbalance in the essential insulin production and/or utilization system. Type 1 diabetes contributes to nearly all the remaining instances, where the cause is unknown and the individual's insulin production has stopped. For both types of the disease, the blood glucose level increases to unhealthy levels due to the defect in the insulin system. Thus, frequent use of blood glucose sensors is essential for this patient group as an aid in achieving healthy values and reducing chances of complications such as progressive development of retinopathy with potential blindness, nephropathy that may lead to renal failure and/or neuropathy with risk of foot ulcers, amputations, sexual dysfunction and substantial increased risk of cardiovascular diseases. For Norway alone, such complications are calculated to cost the country 3.5-4 billions NOK per year (Norges Diabetesforbund 2000). In addition to the artificial supply of insulin, physical activity and proper nutrition are actions, and thus parameters, that both patients and care givers need to relate to for achieving healthy blood glucose values. So are parameters such as weight, cholesterol level, micro-albumin, blood pressure, and so on.

People with diabetes are in regular contact with health care personnel like general practitioners (GPs), hospital nurses, diabetes specialists, foot therapists and eye specialists. Usually, these contacts result in the sampling of one or more of the parameters mentioned above. This data is stored in one or more electronic health record (EHR) systems. In Norway, a lot of data is still stored in unsynchronized systems. Data entered into EHR systems is stored in an underlying database, and the different medical documents are created real-time based on predefined report templates. Today, all the documents are both designed and generated by professional health care actors, i.e. none by the patients themselves.

Even though many of the blood glucose meters store the last measurements and have export possibilities to a PC, few of the hospitals or GPs have procedures for

capturing these data in their EHR. Instead, the patients are asked to remember their last measurements at their annual, bi-annual or quarterly visits. Thus, the most important parameter for avoiding diabetes complications is still not utilized in medical documents, neither for use by health care personnel nor for the patients or relatives. The same is true for the two other most important parameters for people with diabetes; nutrition and physical activity data. The latter two are hardly ever captured at any health institutions. Prescriptions of medicines and measurement equipment are still not generated or transferred automatically to pharmacies, but issued as a paper document from the doctor to the patient (often handwritten).

Figure 1: Automatic blood glucose data transfer to EHR and as an SMS to relatives. (Photo: Jarl-Stian Olsen/Hilde Pettersen NST)

A likely future scenario is that all important patient-sampled sensor data is stored, arranged, displayed and used by both the patient and the health care personnel to achieve an optimal health management. Prescriptions are automatically generated based on the actual use and effect of the medicine and measurement equipment, and transferred electronically to both pharmacies and patients. One of the projects at the Norwegian centre for telemedicine (NST) has designed and tested a concept that exemplifies part of this scenario for the blood glucose parameter. The prototype combines the short-range communication standard Bluetooth and GSM for long-range communication, in automatically transferring the medical sensor data both directly to an EHR system (Årsand et al. 2004) and to a relative's mobile phone (Gammon et al. 2005), see Figure 1.

If sensor data such as blood glucose were part of the EHR database, it provides an opportunity to generate novel medical documents to aid the patients, both in the treatment of illnesses and in prevention of getting them. Using data analysis prior to automatic document generation gives both the patients and the health care personnel answers to questions like: "At which time of the day are the parameters at a less healthy level?", "Which weekday(s) does the patient need to focus most on achieving healthy values?", "How many percent of the values are within the healthy range?", "How well regulated is this person compared to the average within her/his profile?", and so on. NST is also working with capturing and integrating nutrition and physical activity data (Eirik Årsand and Gunnar Hartvigsen, 2005) to provide a better clinical overview for this patient group. The idea is that the concept of direct capture of patient data into medical documents will shortly manifest itself in more patient groups, and even be used as a preventive measure among people at risk for developing certain diseases. Such an extensive use of patient-generated data needs subsequent routines for document generation, distribution and follow-ups.
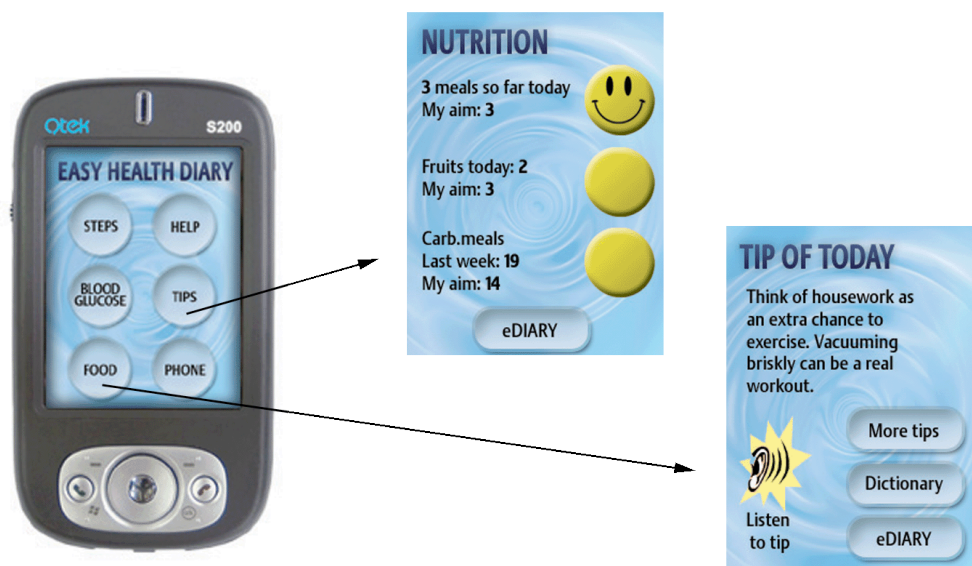

Figure 2: Example on Digital patient diary for people with diabetes; main menu and two example functions (Design: Årsand/Varmedal/Olsen, NST).

Future medical documents will, instead of mainly being either paper-based or electronic documents on hospital data terminals, to a large degree be constituted of electronically displayed information on the patients' stationary terminals (PCs, TVs) and mobile terminals (mobile phones, PDAs). The growth of digital patient

diaries like the one displayed in Figure 2 is part of this scenario. So is the introduction of automatic computer-generated handling and feedback on the patient-sampled data.

7 Closing remarks

The notion of a "document" has profound relevance within healthcare. It is used for just about anything, from an X-ray image, a discharge report or a lab report to a prescription. Multimedia content is in some cases regarded a document. Aggregated and processed data is in the case of an epidemic a powerful document that can affect entire populations. The spectrum of objects called documents is, in other words, very wide. As we have touched upon in this essay, complexity plays an important role when speaking about documents in healthcare, both at the molecular level and the societal level, which documents within healthcare are operating on at all levels. Regarding documents as "complete information objects", as HL7's CDA does, provides us with a useful insight into the very crux of 'documentation' in this context. As the field of medicine and healthcare in general is highly complex, involving many people in the process of care-giving for patients, or also research and education, it is quite a challenge to produce such complete information objects at any given time. As documents are to be useful for such a broad range of people and uses with severe requirements for availability, security, informability and nearly every other aspect that could test the boundaries of information and communication technology, it is no wonder that standardization processes – trying to deal with the very issue of complexity – is having such a hard time making general models of the systems that deal with giving care to human beings.

There is a contradiction between the need for generalizing to build information systems and the specialization, where every patient, every human being has a need for a personalized treatment and follow-up, both in terms of medical treatment and social/cultural preferences. The tools available so far are classification and standardization, structuring and modeling information systems. Some issues are solved by such an approach, and others arise from it. Medical work, especially critical care, is exceptional by nature and collaboration-intensive. The most valuable resource in providing quality care is well-informed and competent medical staff. The paper-regime dominating healthcare from its infancy until now is rapidly being replaced by its electronic counterpart. Digital documents offer new opportunities, as information can be shared easily and be a part of electronic communication services on the Internet.

8 References

Beale, Thomas, 2005. The Health Record – why is it so hard? In *IMIA Yearbook of Medical Informatics 2005*, eds. R. Haux and C. Kulikowski. Stuttgart: Schattauer GmbH. Verlag für Medizin und Naturwissenschaften.

Buechner, Jay S., Constantine, Herbert, and Annie Gjelsvik, 2004. John Snow and the Broad Street pump: 150 years of epidemiology. *Medicine and Health Rhode Island* 87, 314-315.

Dichter, Jeffrey R., 2003. Teamwork and hospital medicine: A vision for the future. *Critical Care Nurse* 23, 8-11.

DICOM, 2006. Digital Imaging and Communications in Medicine (DICOM). Rosslyn, Virginia, USA, National Electrical Manufacturers Association (NEMA).

Dolin, Robert H., Alschuler, Liora, Boyer, Sandy, Beebe, Calvin, Behlen, Fred M., Biron, Paul V., and Amnon Shabo, 2006. HL7 Clinical Document Architecture, Release 2. *J Am Med Inform Assoc* 13, 30-39.

Frerichs, Ralph R., 2001. History, maps and the internet: UCLA's John Snow site. *SoC BULLETIN* 34(2), 3-7.

Gammon, Deedes, Årsand, Eirik, Walseth, Ole Anders, Andersson, Niklas, Jenssen, Martin, and Ted Taylor, 2005. Parent-Child Interaction Using a Mobile and Wireless System for Blood Glucose Monitoring. *Journal of Medical Internet Research* 7(5):e57.

Haux, Reinhold, 1998. Health and medical informatics education: perspectives for the next decade. *Int J Med Inform* 50, 7-19.

Henning, Kelly J., 2004. What is syndromic surveillance? *MMWR Morb Mortal Wkly Rep* 53 Suppl, 5-11.

Hersh, William R., 2003. *Information Retrieval: A Health and Biomedical Perspective*. New York: Springer-Verlag, Inc.

Horsch, Alexander, Balbach, Thomas, Melnitzki, Susanne, and Jürgen Knauth, 2000. Learning Tumor Diagnostics and Medical Image Processing via the WWW – the Case-based Radiological Textbook ODITEB.

Horsch, Alexander, Blank, Ricarda, and Doris Eigenmann, 2005. The EFMI reference image database initiative: concept, state and related work.

Horsch, Alexander, Itälä, Timo, Mikola, Tuire, Leonhardt, Petra, Tobman, Matvei, Nätscher, Catharina, and Helmut Sußmann, 2003. CDA-based Integration of a Teleconsultation Service into a Regional Electronic Patient Record System. *Medical Informatics Europe MIE.*

IOM (2001) Crossing the quality chasm: a new health system for the 21st century. Washington, DC, Institute of Medicine (US) Committee on Quality of Health Care in America. National Academy Press.

Good Medical Practice, 2006. Online: http://www.gmc-uk.org/guidance/good_medical_practice/GMC_GMP_V41.pdf (accessed 29 November 2006).

Kaufmann, Arnold F., Meltzer, Martin I., and George P. Schmid, 1997. The economic impact of a bioterrorist attack: are prevention and postattack intervention programs justifiable? *Emerg Infect Dis* 3, 83-94.

Miller, James G., 1995. *Living Systems.* Niwot; Colorado: University Press of Colorado.

Norges diabetesforbund, 2000. Noen fakta om diabetes – Program og handlingsplan 2001-2003. Oslo.

Olender, Lynda M.A.R.N., 2005. Hospitalists: A Chief Nursing Officer's Perspective. *Journal of Nursing Administration* 35, 497-501.

Snow, John, 1965. On the mode of communication of cholera. *Snow on Cholera* (Reprint). 2 ed. New York, New York: Hafner Publishing Co.

Sokolowski, Rachel, and Joachim Dudeck, 1999. XML and its impact on content and structure in electronic health care documents. *AMIA Symposium.* Washington DC.

Teichman, Peter G., 2000. Documentation tips for reducing malpractice risk. *Family practice management* 7, 29-33.

Wild, Sarah, Roglic, Gojka, Green, Anders, Sicree, Richard, and Hilary King, 2004. Global Prevalence of Diabetes – Estimates for the year 2000 and projections for 2030. *Diabetes Care* 27, 1047-1053.

Årsand, Eirik, and Gunnar Hartvigsen, 2005. A Wearable eHealth System for People with Type 2 Diabetes. In *Scandinavian Conference on Health Informatics SHI 2005,* eds. O. Hejlesen and C. Nøhr. Aalborg, Denmark: Virtual Centre for Health Informatics, Aalborg University.

Årsand, Eirik, Walseth, Ole Anders, Andersson, Niklas, Fernando, R., Granberg, Ove, Bellika, Johan Gustav, and Gunnar Hartvigsen, 2005. Using blood glucose data as an indicator for epidemic disease outbreaks. *Stud Health Technol Inform* 116, 217-22.

Årsand, Eirik, Walseth, Ole Anders, and Eva Skipenes, 2004. Blood glucose data into Electronic Health Care Records for diabetes management. In *HelsIT 2004,* ed. Ø. Nytrø. Trondheim, Norway: Norwegian Research Centre for Electronic Patient Records (NSEP).