1 Introduction

The traditional paper-based patient record used in a clinical setting generally contains the notes of clinicians and other care providers. These notes are often supplemented with data from other sources: laboratory test results and reports describing the results of other tests that have been performed, such as X-rays, pathology, ultrasound, lung function, and endoscopy (see Chapters 12 and 13). With the exception of electrocardiograms, some images, or drawings, the majority of information in the paper-based record involves data that can be expressed in characters and digits (alphanumeric data). In most European countries, the nursing record is usually kept as a separate document (see Chapter 14). Most nontextual information, especially images, can be viewed only upon request, and it may even be necessary for the clinician to go to a special location to view the materials. Hence, the set of patient data is generally not yet available as a whole at the place and time it is needed.

This chapter first provides a brief overview of the history of the patient record, starting with Hippocrates, to be followed by modern views on the structure of the patient record, the development and use of computer-based patient records (CPRs), the entry of data into the CPR, coding and standardization, the representation of time in CPRs, and the clinical use of the CPR. Chapter 29 provides more details on structuring patient records in computers.

2 History of the Patient Record

The patient record is an account of a patient’s health and disease after he or she has sought medical help. Usually, the notes in the record are made by the nurse* or the physician. The record contains findings, considerations, test results and treatment information related to the disease process.

* In some countries, both physicians and nurses have equal access to medical records, which are then most often called patient records. In other countries, there are separate medical records and nursing records.

In the fifth century B.C., medical reporting was highly influenced by Hippocrates. He advocated that the medical record serve two goals:

1. it should accurately reflect the course of disease, and
2. it should indicate the possible causes of disease.

With the medical insight of those times, the records contained descriptions of events that preceded disease rather than real causal clarifications. Panel 7.1 shows how Hippocrates described the course of a disease (see Fig. 7.1 for the text in Greek). The example shows
Panel 7.1

Hippocrates Describes a Disease

The description starts with the patient’s history prior to his request for medical help:

“Apollonius was ailing for a long time without being confined to bed. He had a swollen abdomen, and a continual pain in the region of the liver had been present for a long time; moreover, he became during this period jaundiced and flatulent: his complexion was whitish.”

Hippocrates proceeds with the reason for seeking medical help:

After dining one day and drinking to excess, Apollonius “at first grew rather hot and took to his bed. Having drunk copiously of milk, boiled and raw, both goat’s and sheep’s, and adopting a thoroughly bad regimen, he suffered much therefrom.”

Reports on the progress of the illness follow from that time onward. They are not provided daily, but are provided only at times when important changes in the symptoms occur.

There were exacerbations of the fever; the bowels passed practically nothing of the food taken, the urine was thin and scanty. No sleep. Grievous distention, much thirst, delirious mutterings. ……

About the fourteenth day from his taking to bed, after a rigor, he grew hot; wildly delirious, shouting, distress, much rambling, followed by calm; the coma came on at this time. … About the twenty-fourth day comfortable; in other respects the same, but he had lucid intervals. ……

About the thirtieth day acute fever; copious thin stools; wandering, cold extremities, speechlessness. Thirty-fourth day: Death.

'Απολλώνιος ὁ Ῥηθοστάθην ὑπέφερετο χρόνον πολὺν. ἦν δὲ μεγαλὸ-σπλαγχνος, καὶ peri ἦπαρ συνήθης ὀδύνη χρόνον πολὺν παρείπτετο, καὶ δὴ τὸ τέλε καὶ ἱκτερώδης ἔγεντο, φυσώδης, χροιῆς τῆς ὑπολεύκου. φαγὼν δὲ καὶ πίων ἀκαρότερων βέβιον ἑθερμάνθη σμικρὰ τὸ πρῶτον, κατεκλίθη. γάλαξι δὲ χρησάμενος ἐφοίσε καὶ ὕμοισι πολλοίσι, αἰχέοισι καὶ μηλείσια, καὶ διαστῇ κακῇ πάντων, βλάβαι μεγάλαι. οἶ τε γὰρ πυρετοὶ παρεξωθήσον, κούλη τε τῶν προσενεχθέντων οὐδὲν διέδωκεν ἄξιον λόγου, οὐρὰ τε λεπτὰ καὶ ὄλης διημέ. ὕπνοι οὐκ ἐνήσαν. ἐμφύσια κακῶν, πολὺ δίους, κωμά-τάδης, ὑποχονδρικοῦ δεξιοῦ ἐπαρμα σὺν ὀδύνη, ἀκρεὼ πάντωθεν ὑπό-ψυχρα, σμικρὰ παρέλεγε, λήθη πάντων δ ὦ λέγω, παρεφέρετο. peri δὲ τεσσαρεσκαιδεκάτην, ἀφ᾽ ἢς κατεκλίθη, ὑγίωσας ἐπεθερμάνθη. ἑξεμάν. βοῆ, ταραχή, λόγιοι πολλοί, καὶ πάλιν ἰδρυσις, καὶ τὸ κύμα τηνικάυτα προσῆλθε. μετὰ δὲ ταῦτα κούλη ταραχώδης πολλοίσι χολώδεσιν, ἀκρή-τοισιν, ὕμοισιν. οὐρὰ μέλανα, σμικρὰ, λεπτὰ. πολλὴ δυσφορία, τὰ τῶν διαιρημάτων ποικίλως. ἢ γὰρ μέλανα καὶ σμικρὰ καὶ ιῶδεά ἢ λιπαρά καὶ ὕμια καὶ δακρύλεα. κατὰ δὲ χρόνους ἐδόκει καὶ γελακτώδεα διδόναι. peri δὲ εἰκοσθῆν τετάρτην διὰ παρηγορήσεις. τὰ μὲν ἄλλα ἐπὶ τῶν οὕτων, σμικρὰ δὲ κατενόηνεν. ἐξ οὐ δὲ κατεκλίθη, οὐδενὸς ἐμνήσθη. πάλιν δὲ ταχὺ παρ-ἐνεί, ὄρμητο πάντα ἐπὶ τὸ χείρον. peri δὲ τριηκόσιον πυρετοὺς ὄξυς, διαιρημάτα πολλὰ λεπτὰ, παράληπρος, ἀκρεὼ ψυχρά, ἄφωνος. τριηκόστῃ τετάρτῃ ἔθανε.

Figure 7.1
Description of a disease by Hippocrates 2,600 years ago. The patient history is that of Apollonius.
that he recorded his observations in a purely chronological order. We call such a record a time-oriented medical record. The descriptions mainly reflect the story as it is phrased by the patient and the patient’s relatives. In Hippocratic medicine, it was very important to estimate the prognostic value of findings. Well-documented disease histories play an important part in achieving that goal. It is the physician’s and nurse’s most important task to relieve suffering, but these providers must know their limits and refrain from pointless interference. Hippocrates’ vision is still the basis for the oath or promise that all physicians must take before they can start to practice their profession.

Until the early 19th century, physicians based their observations on what they could hear, feel, and see. In 1816, Laennec invented the stethoscope. This instrument contributed considerably to available diagnostic techniques. When more diagnostic instruments became available, such as the ophthalmoscope and the laryngoscope, a terminology was developed to express the new findings with these instruments. The advent of this new technology caused the emphasis of the patient record to expand the scope from the story told by the patient or the patient’s family to the findings of the physician and the nurse. Shortly after 1880, the American surgeon William Mayo formed the first group practice, which became the now well-known Mayo Clinic in Rochester, Minnesota. In the early Mayo Clinic, every physician kept medical notes in a personal leather-bound ledger. The ledger contained a chronological account of all patient encounters. As a result, the notes pertaining to a single patient could be pages apart, depending on the time intervals between visits. The scattered notes made it complicated to obtain a good overview of the complete disease history of a patient. In addition, part of the patient information could be present in the ledgers of other physicians. In 1907, the Mayo Clinic adopted one separate file for each patient. This innovation was the origin of the patient-centered medical record. The fact that all notes were kept in a single file did, however, not mean that there were criteria which the content of those records had to meet. In 1920, the Mayo Clinic management agreed upon a minimal set of data that all physicians were compelled to record. This set of data became more or less the framework for the present-day medical record. Despite this initiative toward standardization of patient records, their written contents were often a mixture of complaints, test results, considerations, therapy plans, and findings. Such unordered notes did not provide clear insight, especially in the case of patients who were treated for more than one complaint or disease. Weed tackled the challenge to improve the organization of the patient record, and in the 1960s he intro-
duced the **problem-oriented medical record**. In this problem-oriented medical record, each patient was assigned one or more problems. Notes were recorded per problem according to the **SOAP** structure, which stands for subjective (S; the complaints as phrased by the patient), objective (O; the findings of physicians and nurses), assessment (A; the test results and conclusions, such as a diagnosis), and plan (P: the medical plan, e.g., treatment or policy). Besides further improvement in the standardization and ordering of the patient record, the main purpose of the problem-oriented SOAP structure is to give a better reflection of the care provider’s line of reasoning. It is immediately clear to which problem the findings and the treatment plan pertain. Although Weed’s problem-oriented record was readily accepted on a rational basis, it proved to require much discipline to adhere to the method in practice. Data associated with more than one problem need to be recorded several times. Panels 7.2, 7.3, and 7.4 provide three versions of the same notes in time-oriented, source-oriented, and problem-oriented formats, respectively.

### 3 The Present-Day Medical Record

Most modern patient records are not purely time oriented, because strict chronological ordering makes **trend analysis** difficult. Laboratory test results may be separated by visit notes, X-ray reports, and other kinds of information. In such a record, one cannot quickly obtain insight into the course of, for example, the
**Problem-Oriented Medical Record**

**Problem 1: Acute bronchitis**
**Feb 21, 1996**
S: Shortness of breath, cough, and fever.
O: Pulse 95/min, Temp: 39.3°C.
Rhonchi. ESR 25 mm.
Chest X-ray: no atelecitasis, slight sign of cardiac decompensation.
A: Acute bronchitis.
P: Amoxicillin caps. 500 mg twice daily.

**Mar 4, 1996**
S: No more cough, slight shortness of breath.
O: Pulse 82/min. Slight rhonchi.
A: Sign of bronchitis minimal.

**Problem 2: Shortness of breath**
**Feb 21, 1996**
S: Shortness of breath.
O: Rhonchi, RR 150/90.
Chest X-ray: no atelecitasis, slight sign of cardiac decompensation.
A: Minor sign of decompensation.

**Mar 4, 1996**
S: Slight shortness of breath.
O: RR: 160/95, pulse 82/min.
A: No decompensation.

**Problem 3: Dark feces**
**Feb 21, 1996**
S: Dark feces.
Present medication Aspirin 64 mg per day.
O: Abdomen not tender, no blood on the glove at rectal examination Hb 7.8.
A: Intestinal bleeding possibly due to Aspirin.
P: Reduce Aspirin to 32 mg per day.

**Mar 4, 1996**
S: Normal feces.
O: Occult blood feces.
A: No more sign of intestinal bleeding.
P: Keep Aspirin at 32 mg per day.

---

hemoglobin level. To facilitate trend analysis, current records are generally source oriented. The contents of the record are ordered according to the method by which they were obtained; notes of visits, X-ray reports, blood tests, and other data become separate sections in the patient record. Within each section, those data have a chronological order. Problem-oriented recording following Weed’s SOAP code affects only the clinical notes.

An important question is how well the current paper-based record is suited for its purpose. As one may rightfully expect, the patient record is used first and foremost to support patient care. However, developments in health care have made this task more complex, and there is also a greater demand for patient data for purposes other than patient care. Well-recognized ways of using the patient record today include the following:
Data from Patients

- Supporting patient care:
  - a source for evaluation and decision making, and
  - a source of information that is shared among care providers.
- A legal report of medical actions.
- Supporting research:
  - clinical research,
  - epidemiological studies,
  - assessing quality of care, and
  - postmarketing surveillance of drugs.
- Educating clinicians.
- Healthcare management and services:
  - providing support for billing and reimbursement,
  - a basis for pre-authorization by payers,
  - providing support for organizational issues, and
  - providing support for cost management.

Although medical notes are usually recorded on paper, paper-based notes have disadvantages that mainly stem from medical progress. The enormous growth of medical knowledge has led to an increasing number of clinical specialties. Specialization leads to multidisciplinary care, so that more than one care provider is involved in a patient's treatment. In such a setting, one physical record per patient causes too many logistical problems. Therefore, there are often as many records for a patient as there are specialties involved in his or her treatment. Patient data then become scattered among a variety of sources. When clinicians want to form a complete picture about a patient's health, they may need to consult records that are kept by their colleagues. Paper files can only be in one location at a time, and sometimes they cannot be found at all. Handwriting may be poor and illegible, data may be missing, or notes may be too ambiguous to allow proper interpretation.

The rapid advances in medical technology and information make it difficult even for specialists to have state-of-the-art knowledge at their fingertips. Yet, patients may expect treatment according to the best available insight. A fundamental limitation of paper-based records is that they can only contribute passively to the decision making of the clinician (see Chapters 15-17). The record cannot actively draw the care provider's attention to abnormal laboratory values, contraindications for drugs, or allergies of the patient, for example, to iodine and penicillin.

Beside limitations that directly involve patient care, the paper-based record also has disadvantages that are related to research purposes and healthcare planning. To support these goals, patient data need to be unambiguous, well structured, and easily, but not unlawfully, accessible. Respecting the privacy of the patient is a topic of continuous concern (see Chapter 33).

It is obvious that retrospective research on the basis of large numbers of paper-based records is extremely laborious and that many data would prove to be missing or useless. This is one important reason why most studies are conducted prospectively.

In summary, paper as a storage medium for patient data has the following disadvantages:

- The record can be only at one place at a time. It may not be available or it may even be missing.
- The contents are in free text; hence they are:
  - variable in order,
  - possibly illegible,
  - possibly incomplete, and
  - possibly ambiguous.
- For scientific analysis, the contents need to be transcribed, with potential errors.
- Paper-based notes cannot give rise to active reminders, warnings, or advice.

4 The Computer-based Patient Record

The increasing demand for well-structured and accessible patient data, in combination with developments in computer science, sparked a great inter-
heavy demands on data collection. The following sections offer an introductory overview of the CPR. Chapter 29 provides a more elaborate description of the reasons for structuring the CPR.

For more than 25 years people have tried to develop the CPR. The first developments were in a hospital setting and focused on those parts of the patient record that were relatively easy to structure, such as those containing diagnoses, laboratory test results, and medication data. Narratives proved to be far more difficult to collect in a structured format. Typical examples of narratives are notes on the history of the patient and the physical examination. Not only do clinicians vary widely in the phrasing of their findings but they also appear to be reluctant to enter data directly into a computer, because they felt that data entry on a terminal would be time-consuming and unfriendly to the waiting patient.

Several elaborate systems developed in the 1970s continue to remain in use. Examples are COSTAR, TMR, RMIS, STOR, and ELIAS. For the collection of notes from physicians or nurses, these systems use so-called encounter forms. An example of a form is the one used in the Regenstrief System (RMIS),

---

Image of a form from the Regenstrief Medical Record System. The upper left is a list of diagnoses. Below this list, some structured data, such as vital signs, can be entered. The progress notes are handwritten and will be entered into the system by trained personnel. Each specialty has its own dedicated forms.

Figure 7.2

---

Sample patient record page.
shown in Fig. 7.2. On these forms, the system has printed part of the patient data, such as diagnoses and problems that the patient is known to have, medication prescribed at the previous visit, and test results that have become available. In the case of a new patient, only basic administrative information appears on the form. Most of the encounter forms have a number of fixed items, which the care provider is expected to fill in. Examples of such fixed items are weight, blood pressure, pulse rate, possible new diagnoses, medication, and medical decisions. The physician or the nurse can add notes pertaining to history and physical examination in writing, if they are considered to be relevant. A variety of different forms are usually available to accommodate preferences at the level of a clinical specialty or department. After office hours, the contents of the forms are entered into the computer by clerical personnel. Clinicians can consult the patient record on the computer at any time and generally do so mainly outside office hours.

Transcription of freehand dictations by clerical staff has the disadvantage that the data are not immediately available and may contain errors as a result of misinterpretations. The next sections address the use of CPRs in primary and specialty care.

4.1 Primary Care

Over the past decade, general practitioners (GPs) in The Netherlands and the United Kingdom have made considerable progress with respect to the use of the CPR. In 1997, more than 90% of the Dutch GPs were using an information system and more than 50% of them had replaced their paper-based charts by CPRs. GPs are far ahead of the specialists, who rarely use a computer for record keeping, and if they do, it usually involves the recording of data in the context of research. Besides electronic record keeping, primary care information systems also support the administrative and financial aspects of running a practice. Usually, the CPR is an electronic version of the paper-based patient record with options for problem-oriented record keeping. The information system is usually able to print referral letters and prescriptions, and it often provides the option to code diagnoses and findings according to the International Classification of Primary Care (ICPC) or the ICD-9-CM (see Chapter 6). Although the latter requires initiative from the care provider, there is a strong incentive on the part of the clinicians to do this themselves when coding is required for reimbursement.

4.2 Specialty Care

One would expect specialists to have a far greater need for computerization of the patient record than GPs. GPs treat patients on a long-term basis and they receive reports from all other patient care providers, whereas the specialist is often confronted with fragmented patient data. Yet, there are several explanations for the fact that specialists are not as eager to adopt the CPR as GPs. In some countries, GPs either run a practice by themselves or participate in a small group practice. In such a setting, fewer other people are involved in decisions concerning practice management. Specialists, however, work in a complex environment, where management staff must consider the influence of information technology on a variety of future users and on available resources and logistics. This interdependence between decisions by specialists and organizational aspects is very strong in a managed-care environment, such as in the USA. Although GPs may be confronted with a broad spectrum of pathologies, their notes are usually less extensive and less detailed than those of a specialist. Hence, the interactive use of a CPR may be more time-consuming for a specialist than for a GP. Finally, there may be different specialties in a given clinic, each with its own requirements for the contents of the patient record. It is unlikely that one CPR could satisfy the majority of specialists. System developers must tailor the CPR for a specialist in such a way that the record can accommodate a variety of domains, while the record's contents can be merged with those of other providers to form a complete record of the patient's medical history.
5 Data Entry

Present CPRs usually support time-oriented, source-oriented, and problem-oriented views of the patient data, although the latter requires that clinicians explicitly specify the patient's problems and the relationships between problems and findings. It is widely recognized that patient descriptions created via structured data entry are essential to obtain reliable patient data that are suitable not only for patient care but also for example for decision-support and research. Therefore, the entry and presentation of data in the CPR are primary topics in CPR research. Although a variety of combinations may occur, there are two main strategies for the collection of structured data:

- natural language processing, and
- direct entry of data in a structured fashion.

5.1 Natural Language Processing

Natural language processing (NLP) has the advantage that it can be applied to existing free text. The text must be obtained by dictaphone or a speech recognition system, but the most current CPRs allow physicians and nurses to remain fully free in the amount of detail that they provide and their choice of words. Thesauri of medical terms (see Chapter 6) and knowledge of language structure (syntax, synonyms, etc.) can assist in parsing sentences in free text (or natural language). The most elementary form of NLP produces an index of the terms used. Such indices are used to retrieve texts in which one or more specified terms are present. However, when the search criterion is, for instance, cough, the result will contain cases in which cough has been confirmed as well as denied. Therefore, medically correct coding requires semantic knowledge about medical terms, their synonyms, and how the terms may be combined into meaningful expressions.

Examples of such knowledge are:

- stomach: is an organ,
- cough: is a complaint,
- AIDS: is a disease,
- dyspnea: is synonymous with shortness of breath, and
- pain: can be described by location, severity, progression, radiation, precipitating factors, etc.

These descriptors, in turn, can be defined by other terms. This knowledge, which has a semantic basis, in combination with syntactic knowledge, can be applied to the interpretation of sentences in a medically meaningful fashion. Even if terms such as "cough" and "sputum" are separated by several other words, there will be no question that they belong together since sputum is one of the descriptors of cough. Take the following sentence in a patient history: “Maximum walking distance without pain is 200 m.” NLP methods can deduce that “leg claudication” is involved, provided that “maximum walking distance” is only known as a descriptor for “leg claudication.” Note that the knowledge would be incomplete if “walking distance” would be the only descriptor for “leg claudication,” because pain elsewhere may also limit walking distance.

Yet, numerous problems remain. Take, for example, the following phrase: “Walking distance was 200 m due to chest pain, but not to pain in the leg.” The concept “maximum walking distance” may occur in several contexts, among which are angina pectoris and leg claudication. Important in the context of this sentence is the recognition of the negation regarding maximum walking distance due to leg claudication. Furthermore, a good synonym thesaurus is needed to associate “chest pain” with “anigina pectoris”, and “pain in the leg” with “leg claudication.” Because of the possible danger of incorrect interpretation by NLP, the results preferably need to be checked, but this additional step again requires human effort. Unfortunately, narratives often contain ambiguities that can easily be resolved by a human but not by a computer algorithm.

Important advances in our ability to obtain useful data from narrative reports have been achieved in the
Figure 7.3
Example of a screen of the Pen&Pad system that helps the user to select the symptom or complaint to be described. The list on the left presents complaints of a general nature. The symptom list on the right has a specific focus, depending on a location selected by the user, in this case the chest.

Figure 7.4
The screen of the Pen&Pad system offers predefined options for description after a symptom has been selected. Here, "cough" can be further specified by direct descriptors and associated symptoms. More detail can be added via the buttons "more" and "additional." The user can see the result of his data entry on the right side of the screen in textual format.
domain of Radiology by the Columbia-Presbyterian Medical Center in New York. Their NLP software performs as well as coding by radiologists and internists under certain circumstances. The conclusion is that NLP, at best, will be able to extract as many data as a human colleague can but we have reached that point in only a few very limited domains. The advantage that, in principle, NLP does not restrict the clinician in phrasing his or her findings coexists with a disadvantage: NLP cannot stimulate care providers to be more complete and more explicit in their descriptions.

5.2 Structured Data Entry

The other main strategy for obtaining data in a more complete and less ambiguous format is to enter the data directly in a structured way. Therefore, this section focuses on structured data entry (SDE) directly by the care provider. SDE differs significantly from the present routine of patient data entry, but it has proven to result in more reliable and more complete data. It is important that care providers receive as much support as possible when they use SDE. One cannot expect them to use a bulky manual that specifies which terms may be used in which context and in which combinations. This type of information must be incorporated into the user interface in such a way that the user needs only to choose from available options. An example is the Pen&Pad interface for SDE, as shown in Figs. 7.3 and 7.4. The particular strategy that best supports SDE strongly depends on the variety of findings with which the care provider may be confronted; it depends on the size of the clinical domain.

5.3 Forms for Data Entry

In the context of a given medical specialty or of a specific study, only a limited number of items needs to be specified. It may be efficient to combine these items in one or two forms. The data entry task is then reduced to filling out these forms. The advantage of such forms is that users quickly familiarize themselves with the position of the various items and barely need to navigate through the system.

Even when a CPR user is highly specialized or is collecting data in the context of a specific study, users may still need to record patient findings beyond their primary domain of knowledge and interest. These incidental findings must also be recorded in the patient record. The wider the scope of the specialty, it is less possible to predict what will be found. A domain such as internal medicine is so vast that nearly all imaginable findings must be available to the physician. Hence, the CPR must also support SDE for large medical domains. A form-based approach is not suitable for that purpose. The number of potential findings is enormous, whereas a small, unpredictable number of findings may occur in a particular patient. A form that accommodates all potential findings would resemble a soccer field of boxes but with checkmarks in only a few of those. An all-inclusive approach would amount to a large number of forms to accommodate all potential findings. When users must browse through many forms, they may lose insight of the larger picture and conclude that SDE is highly inefficient. They will then probably decide to use the free-text option to present the bulk of their findings.

5.4 Dynamic SDE

When the domain is large and when the findings to be entered are unpredictable, forms ideally should be dynamic, which means that their options should automatically be tailored to the topic of interest. There are several techniques to achieve dynamic forms: using interactive screens, menus, icons, or combinations of these.

Menu-driven user interfaces are well known. In a menu-driven interface, the user chooses an item from a list, which may produce a new list from which another item can be chosen. This procedure repeats until the user indicates that he is finished. The menu offered next depends on the choice of the user.

When the use of sequential menus is the only way that data can be entered, the CPR user may face the problem of not knowing how to navigate through the...
Panel 7.5

Speech Recognition for Reporting of Medical Findings

An increasing number of clinicians are using a speech recognition system to streamline the reporting process. For instance, radiologists, pathologists, and other clinicians use it to report diagnostic imaging findings. Currently, most clinicians use dictation, a slow and expensive method that, due to delays in typing pools, may require days to produce a typed report. Because typing errors may be introduced, the documents must be reviewed by the clinician and mistakes must be corrected.

With a speech recognition system linked to, for example, a radiology information system, the physician can dictate, edit, and instantly create electronic reports. These reports are immediately available to other clinicians and can be integrated with electronic patient records. This leads, in principle, to a considerable saving of time, offering a better service and reducing costs.

At present, several companies provide speech recognition systems for a limited number of languages and medical professions, but each system has its own features. A speech recognition system can usually be installed on a personal computer, equipped with a microphone, with typically 16 to 32 megabytes of random access memory, required to run the program. The system records the speech signal; digitizes and processes the signal (see Chapters 8 and 25); compares the analyzed speech patterns with a collection of possible words, deciding which of these words is most likely to have been articulated (see Chapter 27); and finally, generates the written text. Information from three sources is used to recognize a word: a phoneme inventory, a pronunciation lexicon, and a language model, which contains syntactic and semantic information.

At present, 90 to 95% of words can be recognized correctly. Errors occur when words are wrongly classified and another word is recognized instead, words are not detected (false negative), or words that were not spoken are inserted (false positive). As a result, a certain number of errors must still be corrected. In this way the output of the system becomes more reliable. However, a number of limitations on today's speech recognition systems determine user acceptance:

- menus to reach a certain item. It is time-consuming and inefficient to have to browse through a number of menus to find the desired item. This problem can be solved by offering the option of shortcuts through the use of keywords. Although the combination of menu navigation and keywords enables “to-the-point” data entry, view of the entire record is poor: one can specify some topic in detail but must ascend the menu tree to describe the next topic.

- 5.5 Interfaces for SDE

A variety of methods have been applied to improve the efficiency of SDE. Some improvement to observing the entire record can be achieved by displaying several levels in the menu tree at once. Other applications use a combination of images and text. Icons can be used to symbolize an option for data entry (e.g., a stethoscope...
* Speaking rate
A distinction can be made between discrete speech recognition and continuous speech recognition. Most existing systems can only recognize words when the clinician uses discrete speech, that is, by ensuring that each word has a clearly articulated beginning and end. This requires a slight pause between the words. Continuous speech recognition does not influence the natural speaking behavior of the user. Recent developments in speech recognition enable the system to identify the start and end of a word automatically.

* Speaker dependency
Speaker-dependent speech recognition systems require a training period during which the user "teaches" the computer his or her specific voice profile (see also Chapter 27). A voice profile contains information on individual pronunciation of various phonetic sounds. The recognition accuracy benefits from training the system.

* Domain dependency
A speech recognition system uses a vocabulary containing the words that the system needs for accurate recognition. The size of the vocabulary should be limited to obtain an acceptable speed in the speech recognition process. This is why a specific system can only recognize words within a specific domain. The vocabulary for radiology contains about 24,000 words.

* The need for a knowledge base
The repetitive nature of certain medical reports can be expressed in a "knowledge base" (see Chapters 15 to 17) by using "triggers." These triggers are used to produce an entire line, a paragraph, or even a full page of text by speaking one or only a few trigger words. In this case, the clinician must learn how to anticipate on what the computer is expecting to realize the most rapid report generation. Knowledge bases must be customized for each user. In the future, speech recognition is expected to achieve a high accuracy in dictating medical reports, with an error percentage close to zero and requiring little or no special training. The clinician is then able to use free-form dictation by speaking naturally and continuously, without the need for a knowledge base. This makes speech recognition usable not only for certain standardized reports but also for all medical reports that a clinician must generate. Speech recognition should become speaker independent and work under any condition (e.g., in the presence of background noise). For the time being, our ears and brain are much better than any computer in understanding masses of analog speech information.

for physical examination). Images provide efficient data entry devices for the user to indicate the location of findings, such as when the user localizes an obstruction in a picture of a common angiogram. No matter how the dynamic interface has been realized, it is always based on knowledge about potential medical descriptions. This knowledge has much in common with the domain knowledge necessary for medically meaningful NLP. Speech recognition can be useful in combination with SDE. Current applications for speech recognition are most reliable and effective when the domain involved is restricted, such as radiology reports (see also Panel 17). In combination with SDE, speech recognition can be applied to the options displayed on the screen, which constitute a very limited domain. To make a CPR efficient and attractive for physicians and nurses, it is essential to make optimal use of available interface techniques.
5.6 User-Adaptive SDE

Many clinicians describe their abnormal findings at a
level of detail that they consider relevant, but they often
express normal findings with phrases such as:
"heart/lungs WNL" (within normal limits) or "abdomen
normal." Such expressions denote a set of findings, but
these may vary for each care provider. Most users develop
their own style for history taking and physical exami-
nation. In daily practice, clinicians follow a certain
pattern of actions, and they only go deeper into some-
thing when they encounter abnormalities during this first
glimpse. One physician may use the expression "heart
normal" to indicate a regular heartbeat and the absence
of murmurs, whereas another physician may use the
same expression to indicate the absence of ventricle
enlargement as well.

For the collection of unambiguous data, users should be
couraged to use summarizing terms, but the presenta-
tion of an explicit description of all normal findings is
extremely time-consuming and will quickly cause
annoyance. The fairly constant nature of routine history
taking and physical examination can be used to make an
SDE application more efficient by expanding it with a
feature that allows it to "learn" the meaning of "normal
findings" for each physician or nurse, that is, to be user
adapted. The application will ask the user to specify the
meaning of a "normal statement" upon first use, and it
will feed this specification back upon subsequent use of
that same statement. In this way, efficiency and comple-
teness of data input can be reconciled.

If the SDE in an application is supported in a more intelli-
gent way, care providers will be more willing to accept it.
For such intelligent support it is crucial that the program
be based on semantic knowledge about potential descrip-
tions of findings. The interface should be transparent and
intuitive, and it should clearly distinguish between
descriptive options and actually entered patient data.

5 Coding and Standardization

5.1 Exchangeability of
Patient Data and
Coded Data

Large-scale projects that involve multiple
institutions, even at an international level, not only
need structured and unambiguous data but also
need semantic exchangeability of data. Coding
items, such as the International Classification of
Diseases (ICD), the Systematized Nomenclature
of Diseases (SNOMED), the Read code, and the
ICPC are under continuous development in order to
apply an internationally accepted vocabulary for
medical domain. The challenge is to obtain not
only structured data but also a mapping of entered
data onto these coding systems.

6.2 Nontextual Data

Many diagnostic techniques produce images and signals,
such as X-ray equipment, CT scanning, or magnetic res-
oneance imaging, endoscopy, anatomy, or pathology;
Doppler ultrasound and echocardiography, electrocardi-
graphy, electroencephalography, electromyography,
etc. So far, we have discussed the electronic equivalent of
the paper-based record, which mostly contains reports of
tests that also produce images and signals. All data col-
lected for a certain patient are part of the patient's record.
At present it is time-consuming for a clinician to obtain
nontextual data. As a result, some clinicians only use the
final reports for their decision making without referring to
the source data. The availability of large electronic stor-
age capacity at decreasing cost, as well as improvements
in data transfer and exchange, brings a multimedia
patient record within reach. Once different data formats
and physical distances can be overcome, care providers
will have the complete patient record at their disposal,
including signals and images.
7 Representation of Time

Time plays a very important part in health care. The patient’s course of disease unfolds over time, the physician’s insight may evolve over time, and protocol-based care involves actions with specific intervals in between. Therefore, time stamping is essential. The patient record is a chronological account of observations, interpretations and interventions (see Chapter 1). The physician relies on time-related data for decision making, such as repeating a test or renewing a prescription. Time stamps are also essential for detecting trends, for example, in an intensive care unit or when following the condition of chronically ill patients. For instance, when the number of white blood cells seems to decrease, knowledge of the amount of time between two consecutive measurements is essential.

Time can be expressed in relative terms ("2 days after") or in absolute terms ("June 5th, 10:30 a.m."). Relative time is used in medical knowledge that must be applied to a particular situation and in, for example, descriptions of the course of disease. Absolute time is often associated with facts, such as the date of a visit or the date of a bone fracture. Relative temporal expressions become practical only when they are applied to a concrete situation. Relative time can be expressed as absolute time as soon as the relationship with real time has been established. If a patient with measles shows the first signs of eruption on May 3, then it is expected that they will disappear between May 8 and 10. A clinical protocol often prescribes intervals for monitoring laboratory test results and the duration of certain medications. All these rules are expressed in relative time. Once the starting date of the protocol is known, it is possible to derive which steps need to be taken at what time.

Absolute time is not always preferable to relative time. On the contrary, it is often more difficult for a physician to interpret a series of calendar-oriented events than it is to apprehend the duration of various phases in a disease process. The key issue is that the relationship between relative and absolute time must be known to offer the representation that is most suitable for the situation at hand.

For the purpose of decision making, data are interpreted within the context of medical knowledge. Interpretation of data and decision making can be difficult when time indications are inaccurate. This problem often occurs when one must rely on one’s own memory. The order of a series of events can be very important, but it cannot always be derived from the available data in a reliable way. Computers can be used to perform temporal inferencing, provided that temporal data are recorded in an unambiguous fashion. Computer applications can monitor critical medical parameters and support adherence to a protocol. To take advantage of such applications, time stamps in the CPR must be recorded in a standardized format. Temporal indications are also essential for the following two reasons:

1. The patient record must be a reliable reflection of reality, and because it is not legal to edit the data in a patient’s record at a later date, there must be an option to record evolving insight.
2. Medical actions must always be interpreted in the proper context: a physician may not be held responsible for improper medical actions taken on the basis of insight that was available at a later time but that was not available at the time that those actions were taken.

A classic example of data where several time stamps are required are the laboratory tests:

1. the moment that the sample is taken,
2. the moment that the sample is tested, and
3. the moment that the test results are available to the physician.

A test result provides information about the sample at the time that it was taken. When too much time elapses between sampling and testing, the result may become invalid. It is also important that abnormal test results come to the physician’s attention as soon
Data from Patients

as possible. In short, there must be a standardized way to record the following for each piece of information in the patient record:

- when it was observed and by whom,
- when it was entered and by whom,
- the time that it begins to apply.

8 Clinical Use of the CPR

Despite all of the developments regarding the CPR during the last few decades, it is still only used on a small scale in most settings. The strong focus on the shortcomings of the paper-based record have pushed its strong aspects to the background. The paper-based record has five strong advantages and the CPR has seven principal strengths (Table 7.1). Apparently, most present CPR applications do not yet outweigh the advantages of the paper-based record. Familiarity with the current routine of using paper-based patient records plays an important part. Developers have understood that it is not sufficient to eliminate the limitations of the paper record; its strengths must also come to expression in its electronic equivalent.

Introduction of the CPR in specialized care requires applications for use in inpatient as well as outpatient settings. Inpatient use of the CPR requires that the nursing record also become available in electronic format (see Chapter 14). An important problem of a logistic and financial nature involves data entry and display at the patient's bedside. Bedside computerized equipment is usually available only in intensive care units. Another option would be the use of portable equipment, preferably with a wireless connection to the host system.

---

**Table 7.1**

Advantages of Paper-Based and Computer-Based Patient Records.

---

**Advantages of paper records**

1. They can easily be carried around,
2. Much freedom in reporting style,
3. Easy data browsing,
4. Requires no special training, and
5. Never 'down' as computers sometimes are.

**Advantages of CPRs**

1. Simultaneous access from multiple locations,
2. Legibility,
3. Variety of views on data,
4. Support of structured data entry (SDE)
5. Decision support,
6. Support of other data analysis,
7. Electronic data exchange and sharing care support
8.1 Other Uses of the CPR

Nonacademic medical centers will not usually be attracted to the CPR for its potential for use in research. However, it is to be expected that hospital management, insurance companies, and government institutions will pose a greater demand on patient data (e.g., for the assessment of quality of care). A CPR is almost indispensable for the efficient delivery of data to other parties. Eventually, data for direct patient care and other purposes will have to be collected in one patient record. User acceptance requires a gradual migration from the present situation to a new one. Therefore, the collection of structured data must be encouraged, but not enforced. Users will have to be strongly involved in CPR development. They must recognize their desires and preferences in the final product. Consequently, the development of a CPR must be a concerted action of designers, implementers, and end users. Many user requests involve routine tasks such as billing, correspondence, and printouts of prescriptions. The CPR can only be used to its full potential when it has been used long enough to contain important data about the majority of a practice’s patients.

Clinicians will have to make an investment before they can take advantage of powerful and flexible overviews, decision support, electronic communication, and shared records. The greatest challenge is the tension between effort and benefit. It should be kept in mind that users will invest in the quality of their patient records only if it is rewarding.

8.2 Multimedia Patient Records

In recent years increasing interest has been shown in making the CPR more complete by adding images and signals: the so-called multimedia patient record (MPR; see also Chapter 20). In practice, the development and installation of one integrated solution did not appear feasible. The reason is not only the size and complexity of such a project and the required scope of expertise, but also the fact that healthcare institutions are reluctant to do away with their current equipment and invest in new systems. An MPR should be achieved as much as possible by using existing systems and software and by the gradual introduction of this new technology. This requires the integration of existing systems with new applications to supply the functionality of the MPR. Expertise in networking, communication, and integration is needed. Yet, it is apparent that currently used software varies greatly in type and age and that the task is complicated by the large number of old applications that do not meet present standards for data transfer and communication. (Integration and communication are discussed in Chapter 5.) At this stage, however, the aspects of user acceptance and financial and logistic problems outweigh the technical problems, and NLP and SDE also face many obstacles.

Key References


See the Web site for further literature references.