4

Health information science

Don Walker

The communication of information is fundamental to healthcare. In pre-history, knowledge and skills were passed from healer to apprentice by talking and example. Early cuneiform writing on clay tablets record some medical techniques and remedies. In more recent times, medical books have documented an accumulating human knowledge. Currently, journals, magazines, books, proceedings and pamphlets deliver a daily avalanche of written healthcare information, while audio and video tapes, radio, television, CD-ROM, video-discs and films produce a mountain of audio-visual material.

There is a well established communication between healthcare providers. Between the nurse and the doctor, the doctor and the pharmacist, the radiologist and the pathologist, and so on. A healthcare system must also be managed and funded. This requires a flow of information from the health care recipient and the provider to various organisations for fund and resource allocation, and for management.

All the above information is small compared to that which passes between the recipient of healthcare and the provider. This intercourse contains the daily application of healthcare. In it, and too often hidden, are the clues to new diseases, complications and interactions. Some clues are recorded. They form actual or potential food for researchers.

Healthcare is about people and mostly one-to-one relationships between recipients and a providers. However, the domain is steeped in information and its management. In this, technology will have an ever increasing role to play.

There is no point in applying information technology unless advantages result which outweigh any associated disadvantages. Let us briefly consider some of these and the building blocks required for advances to occur.

The communication of medical information and its management

Communication with literature

The overwhelming volume of printed healthcare information has, from the turn of the century, been indexed in libraries. In recent years, these indices have been created and stored on computers.

The key concepts within the text are identified and named according to a standardised list of terms. One *controlled vocabulary* used for this purpose by the National Library of Medicine (NLM) in Washington DC is called *MeSH* (Medical Subject Headings). It evolved for computer use from *Index Medicus,* the original list of terms used to index a catalogue of the publications in the US army medical library. It was compiled and published by Surgeon General Dr John Shaw Billings, in 1879. Some medical and other health related literature is indexed using a different vocabulary, for example the Library of Congress Subject Headings (*LCSH*),and the Cumulative Index to Nursing and Allied Health Literature *(CINAHL).* These and others are discussed in greater detail in another chapter.

The volume of new written material has long been far beyond any one person's capacity to read, let alone learn and apply. The challenge for information technology is to present <u>relevant</u> information to the healthcare worker <u>when</u> it is needed. For example, if a combination of disease and medications have potential side effects, then the provider should be warned of these at the moment of prescribing. For this to occur, an up to date database of drug interactions must be provided to the computing industry. The electronic medical record must also record a person's details in terminology equivalent to that used by the database.

In the future, the literature may be searched (in a semi-automated process) according to the contents of a person's electronic medical record. This would save the provider compiling searches, and would thus save time and effort. The major obstacle to be overcome before this is a reality, is the difference in terminology used by the healthcare provider and the library indexing system. The *UMLS* (Unified Medical Language System, a research project sponsored by the NLM) is attempting to provide a solution.

All people will inevitably have greater access to current literature and databases. Searching will be easily done from the home computer or local library. Some more aware patients will result. An informed recipient of care is a two edged sword - more able to self-care, but more demanding of new investigations, therapies and procedures. To avoid increasing costs, coordinated education programs may be needed. Thus requirements for improved literature communication are:

- Increased awareness of the existence of available resources
- Wider use of computers
- Unification of healthcare terminologies
- Improved searching programs
- Educational programs which improve user interpretation of the literature

Communication between providers

Healthcare is very much a team game. Communication between players is constant - between general practitioners, physicians, surgeons, pathologists, radiologists, nurses, pharmacists, therapists, lawyers and many more. The essence of any communication should be recorded. The background history of the person concerned, and a summary appropriate to the provider is usually required. The management of a person seeking care can evolve into a complex round of repetitive forms and letters. With each communication, the risk of inaccurate information increases. If there is a coordinator such as a general practitioner, or a case manager from a hospital, or other health care agency who compiles and collates all documents into a person's record, then this record can become a confusing collection of difficult to find documents. Its usefulness declines as its thickness increases.

Electronic communication between members of the team such that each has a view of, and some access to a person's medical or health record, would save time, increase accuracy, and result in better management. In general terms, as the number of pieces of paper are reduced, so the efficiency increases. The pharmacists prescription should be an electronic message sent directly by the doctor to the pharmacist. Likewise the laboratory and radiology requests and reports should not involve paper. For this to materialise:

- Providers must use computers.
- The terminology used by each provider must be consistent and universally understood.
- All involved must agree as to what will be transmitted.
- An adequate communications network must exist .
- Each network node must adhere to standardised message protocols.
- Adequate security, confidentiality and integrity safeguards must exist.
- The provision of up-to-date information resources in a standardised form which are computer understandable.

Communication with organisations

Current communication with funding and resource management organisations involves a complex paper trail with an occasional interspersed computer. Medicare slips, discharge details, workers compensation forms are posted by the millions. Wishes, advice, warnings, demands, rules and regulations issue forth from administrative and professional organisations or associations to healthcare providers, on pieces of paper, brochures and books. Much of this traffic could be replaced if a healthcare network of provider computers existed, and each computer understood the concepts or facts used by the other. Wishes, advice, warnings, demands, rules and regulations could be composed in a computer understandable way, such that when appropriate, the provider would automatically be advised, prompted or warned. Electronic fund transfer is already well established (eg. EFTPOS). It will inevitably be used by the healthcare industry. For these things to eventuate the same conditions listed previously as required for electronic communication between members f the health care team apply.

Communication between provider and patient

The role of the front line providers of healthcare involves listening, questioning, examining, provisionally diagnosing, investigating, confirming the diagnosis, caring, supporting, providing, educating and treating. The facts gleaned and the events which ensue must be recorded. These are required (a) for accurate management and resource allocation, and (b) for future reference by a person's current provider or future providers, and (c) for potential medico-legal requirements. Traditionally this record is the written medical record.

Medical records are certainly the hub of information management for each recipient of health care. They should form the hub of the entire healthcare system and possibly become health records. However, in their written form they are generally speaking illegible, disorganised, inaccessible, mislaid, and impossible to quickly digest. They offer to the provider, no timely, useful and current medical information. No prompts. No warnings. They cannot be easily communicated (in whole or in part) to others. On the other hand, they are as confidential as the inverse of their understandibility. They are cheap to initiate, and are created in a medium familiar to all.

There is a real potential for information technology, if applied to the medical record, to change the health service. The resulting *electronic medical or health record* (EMR) is the subject of another chapter of this book. It's functional needs will be considered in this chapter.

Good information management demands that data are recorded once only, by the person best qualified and at its source. Thus, for example, the office secretary should enter administrative and demographic information; the doctor the medical history, examination and orders; the nurse the daily observations and events; the pathologist (or technician or automated analyser) the laboratory results; the radiologists the interpretation of x-rays ... etc. Transcription of data should not occur.

In the Australian environment of unconscripted and unregimented private healthcare providers, the EMR must be desired by potential users if it is to succeed. This means it must offer the provider advantages over the competing alternative pen and paper system. It must generate less tedium, more speed, timely and current clinical information and, diagnostic and management assistance. It should prompt the provider so as to avoid excesses and omissions. Automated case summaries and views of any aspect of a patients previous records should be available. Form and letter generation should be semi-automated. It should manage financial matters. It should be accurate, secure, accessible yet confidential.

The following must come together to enable a significant improvement in healthcare information management:

- **Healthcare terminologies:** Consistent healthcare terminology, and standard coding and classification of concepts.
- **Resources:** Computer understandable, current, useful and useable information, data and knowledge resources.
- **Computing:** Adequate computing hardware, and software development tools
- **User interface & data capture:** Optimised user interface and data capture methods.
- **Electronic Communications:** Adequate electronic communications.
- **Safeguards:** Adequate security, confidentiality and integrity safeguards.
- **Benefits:** There must be benefits to all concerned - to the computing industry, the healthcare funding organisations, the healthcare providers, and to the patients.

Let us now glance at each of the components necessary for improved healthcare information management. Several of the topics will be considered at length in other chapters. This chapter is designed to offer an overview of the "science" involved. This should help in the formation of a perspective of the subject.

Healthcare terminologies

Components of a terminology

As we have seen, there is a need to describe healthcare concepts in a consistent manner. We, as humans, are able to assimilate, without confusion, many variations of description. Computers, on the other hand are very poor at recognising concepts from inconsistent descriptions. To manage this, a standard list of terms is collected and divided into *preferred terms* and *alternative descriptions* (or *synonyms*). For example, the preferred term "subacute thyroiditis" could have as its synonyms "de Quervain's thyroiditis" and "granulomatous thyroiditis." Within any terminology, the preferred terms should be unique, however each preferred term may have synonyms which are shared with other preferred terms. For example, "cold" could be a synonym for both "hypothermia" and the common viral infection "coryza."

The preferred term is an agreed short description of a *concept*. A concept is the image created by the words which describe it. A *definition*, of the concept may be needed. Too often the wording of the preferred term means something different to different users.

As healthcare is ever-changing, the preferred description of a concept may change. For example, an international terminology has recently changed the preferred term "maturity onset diabetes mellitus" to "non insulin dependent diabetes mellitus." The concept has remained unchanged (ie the disease is the same).

A unique identifier (or *code*) for each concept is required. Anything would do as long as it is unique and suitable. If the preferred term is used, its description could not change. As we have seen, this can happen. Thus the preferred term is not suitable as a code. In fact, words are not efficient ways to store identifiers in computers, as the computer may be required to store an identifier many times. The code should thus be reasonably compact - a "number" of some sort. It need not be seen by users, as the computer can always display its equivalent descriptive words. The process of matching a healthcare entity to a term in a terminology and assigning a code to it, is called *coding*. The terminology may be called a *coding system*. Some terminologies or coding systems describe themselves as (a) a *"controlled vocabulary"* such as MeSH (National Library of Medicine 1994)), (b) a "classification" such as the International Classification of Diseases 9th Revision (ICD9) (World Health Organization 1977), ICD9 with Clinical Modifications (ICD9-CM)(Commission on Professional and Hospital Activities 1978), International Classification of Procedure Codes (ICPC) (Lamberts & Wood 1987)), or (c) a "nomenclature" such as the Systematized Nomenclature of Medicine (SNOMED) (Cote & Robboy 1980).

Sometimes rules are offered to improve the accuracy of coding. These instruct the user to consider certain other concepts under various conditions.

Thus far we have seen that a terminology is composed of a *code*, a *preferred term* to describe each *concept* (ideally *defined*), and various *synonyms*. *Rules* may be provided to direct the user.

Classifications, hierarchies and terminologies

Knowledge enables us to refer to something and unconsciously all related things come to mind - thus if we consider the concept of "appendicitis" we think of those conditions which we have learnt are included, for example "acute, subacute, chronic and relapsing appendicitis." We also know that "diseases of the large intestine" include appendicitis. This organisation of concepts into groups (or classes) such that one group includes others is called *classification*. The resulting structure can be visualised as a branching tree (perhaps turned up-side-down!), and is described as an *hierarchy* (for which the adjective is *hierarchical*).

When the concepts of a terminology are arranged to form an hierarchy, some concepts, by their nature, will be found to belong in more than one place. For example, “tuberculosis of the lung” is both an “infectious disease” and a “lung disease.” Another finding, when building hierarchies, is that concepts belong in different classes depending on the *context* of the classification. For example, various types of apples can be classified according to their size, colour, or time of ripening etc. - and each grouping would be different. Many healthcare concepts can have *multiple hierarchies*. For example, “bleeding disorders” include some diseases from “congenital, infectious, allergic, nutritional and metabolic” conditions, and each of these groups of conditions belong to an hierarchy of their own.

The practical importance of classifications or hierarchies is in their ability to include all subordinate concepts. They can thus specify the detail (or *granularity*) of the data.

- An organisation may need to know how many operations were performed, while an association may be interested in the number of hysterectomies. If the obstetrician records the details of each type of hysterectomy, using an appropriate terminology, then the less detailed requirements of the association and the organisation can be automatically supplied by utilising the hierarchy of the terminology.
- Information resources which can be harnessed by healthcare computing systems need to be able to include various groups of concepts. For example, if the indications for a blood test are required, and they are specified in the resource as “bleeding disorders,” the hierarchy of the terminology should be able to determine which conditions are included in the concept “bleeding disorders.”

As computing systems become “smarter,” their internal use of hierarchies increases.

Specifying hierarchies in terminologies

A classification or hierarchy may be specified by a numbering system:

2	level one concept	(eg. “diseases of the large intestine”)
27	level two concept	(eg. “appendicitis”)
271	level three concept	(eg. “acute appendicitis”)
272	level three concept	(eg. “chronic appendicitis”)
273	level three concept	(eg. “relapsing appendicitis”)

In this example there can only be ten concepts at any level. This method is used by ICD9 and ICD9cm systems.

Note: An hierarchy can be described in terms of parents, siblings, and children. In the above example, the *parent* of “appendicitis” is “diseases of the large intestine.” The *children* of “appendicitis” are “acute appendicitis,” “chronic appendicitis,” and “relapsing appendicitis.” The *siblings* of “chronic appendicitis” are “acute appendicitis,” and “relapsing appendicitis.”

By allocating several digits to a level, more space can be reserved (eg. 200700100, where 100 concepts can occur at each level). The resulting string of characters may be inefficient and unwieldy.

Another method is to count using the base 16 (hexadecimal), and the numbers 0-9 and A-F. This compacts the string of characters, but lack of room between levels may continue to be a problem. This method is used by SNOMED-3 (Cote et al 1993).

Further compaction can be obtained by using the characters 0-9, A-Z, and a-z, making about 58 levels between each character (confusion between the numbers one and zero and upper and lower case “o” and lower case “l” (el), may need to be avoided). This method is used by RCC-1 and RCC-2 (Computer Aided Medical Systems 1993).

Another approach is to use a series of numbers of any size each one of which is separated by a *delimited* character (eg. “D.12.98.21.7”). This method is used by the “tree numbers” of MeSH (National Library of Medicine 1994).

It is possible to specify hierarchies in a computer file (or table). Two fields (or columns) are required. One for “the concepts,” and the other for “the parent of each concept.” By *directing* the computer in its search of this file, the parents, siblings and children of any concept can be determined. Repeated searches allow the full hierarchy to be displayed. Multiple hierarchies can be represented. A concept is simply given more than one parent. Certain *cyclic* references must be forbidden (eg. a concept cannot have itself as a parent). This computing structure (or *graph*) is called a *“directed acyclic graph”*. It is used by RCC-3 (Computer Aided Medical Systems 1993).

Single and multiple axes and terminologies

Healthcare concepts can be expressed by concepts from a single list (or *module* or *axis)*, however this list becomes extremely long, as more clinical detail is coded.

Consider the diagnosis of “acute infection of a bone.” The site of the infection will be required and will need to be coded. Thus, there needs to be a list of concepts for “acute infection” of “each part” of “each bone” in the body. Now consider a diagnosis of “fracture.” Again the bone will be required. An additional list of “fractures” to “each part” of “each bone” will be required. Thus two complete lists of sites for two diagnoses results.

Another approach is to use *multiple axes*, and combine them. Thus, if there is a single list of “site” codes containing the bones and their parts, then this can be combined with the code for “acute infection” or the code for “fracture” to form *compound codes*. Thus one complete list of sites can be used for all diagnoses.

The need to combine axes may be even more apparent when considering a laceration, an abrasion, a bruise, a rash, or an itch, to any part of the body.

The more *atomic* approach of combining *elements* from various axes to form *compound* codes, requires the definition of extensive *rules* or *relationships* to avoid the possible generation of nonsense (eg a “fractured eyebrow”).

Qualifiers (or modifiers) and terminologies

Some systems offer *qualifiers*. (RCC-3.1 (Computer Aided Medical Systems 1994)) or *modifierss* (SNOMED-2 (Cote & Robboy 1980)). These are modifying concepts which can be linked to a chosen concept . They enable increased clinical detail to be coded. In effect qualifiers are an additional axis. If they can be associated with only specified concepts they can be called *legal qualifiers* - otherwise *general qualifiers* would best describe them.

For example, legal qualifiers for bacterial agents using the SNOMED-2 terminology are:

"0 = not otherwise specified"
"1 = growth present"
"2 = growth absent"
"3 = growth contaminated," etc.

The qualifier code number (eg. "1") is added to the concept code number. (For example, if the code for "actinomyces israellii" is "E10730," then "E10730_1" would specify "actinomyces israellii growth present.")

Mapping terminologies

There have evolved several coding systems. Each has been designed for its specific purpose. For example ICD9cm (International Classification of Disease, 9th revision with clinical modification) is used by healthcare managers to code hospital discharge diagnoses and procedures which may then be grouped to become diagnosis related groups (DRGs). The Medicare Benefit Schedule (MBS) (Australian Government Publishing Services 1994)) is used by doctors to code their procedures for payment claims. MeSH is used to code the contents in biomedical literature in the National Library of Medicine, Washington DC. There is often a need to *translate* or *map* or *cross reference* the codes from one system to another. If the above systems were mapped, a procedure identified by a MBS code could be also described in terms of ICD9cm for the hospital manager, and MeSH for literature research. Some coding systems offer mappings to other systems. Maps to ICD9cm are virtually mandatory for hospital coding systems in Australia and the USA.

Mapping one terminology to another is not an easy task. Various types of cross references result. Terms are found which are exact matches, while others have the same meaning using different words, some are less specific, some more specific, other terms do not match at all.

The Unified Medical Language System (UMLS (Humphreys & Lindberg 1989)), as mentioned earlier, is an attempt to develop a terminology interface to biomedical literature, healthcare records, health related databanks and knowledge bases - in the USA. The UMLS has developed a Metathesaurus (Sherertz et al 1989), a Semantic Network (McCray 1989) and a Resource Map (National Library of Medicine 1990). These enable a single query to obtain information from multiple data sources. The *Metathesaurus* is a large compilation of several terminologies which are mapped to MeSH. The *Semantic Network* consists of hierarchies of "semantic types" which have heritable relationships. They form a network. Examples of these relationships are: "is a ...," "exhibited by ...," "carried out by ...," "forms ...," "is an evaluation of ...” etc. The *Resources Map* supplies the information required to access various data sources.

Subsets and terminologies

Healthcare terminologies are large. They often exceed 100,000 terms. Their use requires a coding process. A few words, or beginnings of words, are typed into a computer and used to search the terminology. A list results, from which the user *picks* the most suitable preferred term. If the entire terminology is searched every time, the resulting list may be long and contain many terms which are obviously unwanted, for example a medical practice is unlike;y to require veterinary terms. The contents of a terminology can be identified (by the use of a simple *scoring* system) to show the concepts more likely to be used in various *contexts*. Application programs (eg. medical record systems) can then use the *context score* to reduce the size of *picking lists*. Thus, for example, if a terminology identifies its concepts for age, and the age of the patient is known by the computer, then only those terms appropriate for the age of the particular patient need be listed. To give a specific example, if the patient is an adult, and “penicillin” is required, paediatric versions of penicillin could be omitted from the initial list of options. Likewise, female diseases and operations should be omitted from lists when the patient is a male.

The context scores applied to terminologies can include the type of provider (eg. nurse, general practitioner, surgeon) the component of the task in hand (eg. history, examination, diagnosis, investigation, procedure, medication, operating theatre reports, midwifery reports, etc.), and the race, age, and sex of the patient.

The scoring system can enable picking list to be graded to show “the probable,” “the possible,” “the improbable” and “the lot, regardless.”

Usage frequency may dictate the picking list. Systems may thus *learn-as-they-go*. This is achieved by scoring the frequently used terms so they sort to the top (where they are more accessible). This process may be made *user specific* - ie. those terms most used by a particular user (provider) can be shown first.

The application program should consider the previous history of the patient, and attempt to present previously used terms.

Thus presenting lists of terms appropriate for the provider, the recipent of care and the task, can be a non-trivial undertaking. Some systems offer *micro-glossaries* (SNOMED (Cote & Robboy 1980)) or *specialty subsets* (RCC-3), *subject types* (RCC-3), *sorting numbers*. (RCC-3) (Computer Aided Medical Systems 1993), *Sex* and *age groupings* (ICD9cm (Commission on Professional and Hospital Activities 1978)).

Resources

Information resources are essential ingredients of any quality healthcare information management systems. It is generally beyond the scope of system builders and vendors to generate and maintain these resources. Bodies with the appropriate authority and domain expertise are required for their creation. They need to be supplied in a form that is predictable and readily useable by the system developer, and potentially useful to the end user. Examples of resources include:

- Library contents containing indexes of authors, and keywords and text of the contents (all or a summary) of each publication.
- Medicines data which enable items to be searched from several aspects (eg, generic name, brand name, composition, indications, contraindications, interactions, side effects, etc.). Drug and disease interaction warning, description and explanation data files are required. Default values useful to the prescriber are essential. Prompts to remind the user of current rules and regulations are valuable.
- Laboratory data, which provides information about available tests, their indications, alternatives, costs, requirements, availability, default values, prompts and warnings.
- Procedure data which describes each procedure in a manner suitable for the patient (using lay language), the provider (using professional language), and the account (a shortened form). Indications, contraindications, default values, prompts and warnings should supplement the usually extensive instruction text. These should enable the provider to be prompted and helped at the moment of recording a procedure. They would reduce the time spent on the study of organisational and healthcare rules, regulation, preferences, or advice.
- Provider data, which list healthcare resources and providers, what they do and their contact details.
- Management protocols which remind the user of the preferred management for a disease or situation.

The above resources can vary from simple *text data files* to *knowledge* which can be utilised by a computer. The representation of knowledge (*knowledge representation*) is a complex field, and discussed later in this book. If the knowledge is mixed in with a computer program it is called *hard coded*, and maintaining it becomes impossibly difficult. The knowledge should be in a separate *knowledge-base*, where non-programmers who are *expert* in a field may create and change it. It may be presented in a *knowledge frame*. The knowledge is represented using a standard hierarchical terminology, and carefully thought out *rules*. These usually take the format of "*if* this, *then* that *or else* something." The creation of these rules requires expert domain knowledge and the help of a *knowledge engineer*. Often the *logic* of the rules is expressed in terms of probability or likelihood, requiring a numeric scoring system (eg. 0-4). It is thus somewhat *fuzzy* - hence the term *fuzzy logic*. A computer program (or *engine*) is specially designed to utilise the knowledge in the knowledge base. The required answer may have to be deduced or *inferred* from the rules and their fuzzy logic. A *knowledge base* or *expert system* is thus driven by an *inference engine*.

Healthcare providers are constantly making decisions about the management of an individual's health care. The numerous factors requiring consideration make the task difficult. Any system which aids in the decision process can be called a *decision support system*. For example the timely presentation of relevant information in an easily understood form can help - a graph for instance, or a case summary. Thus, the optimal organisation and presentation of data is a large component of a decision support system. They may also contain extensive expert knowledge and apply this to give results which rival that of a panel of experts. These concepts are discussed in more detail in another chapter.

Some *computer diagnostic systems* are designed to present a list of possible diagnoses based on a given history, examination and results of investigations. The differential diagnosis is shown in order of probability. An explanation as to why each diagnosis is included is available. Some of these systems are able to tell the user what is the best thing to ask or do next. They may consider the risks and the costs. They are useful training systems, but are no match for an experienced clinician. Their future may be to run in the background of a system, and provide helpful messages when it is appropriate, or when called upon. Examples of diagnostic systems include Iliad (Warner HR et al 1988), DxPlain (Hupp et al 1986), Quick Medical Reference (Miller et al 1986).

The electronic medical record is destined to evolve from a relatively simple database of coded events and facts, to an information resource centre and decision support system. As its complexity and expert knowledge increases, its ability to be created and maintained by one author or organisation will decrease. It may have to be built from *modules* (or be *modular in design*). Each module would be created and maintained by an organisation which has the necessary expertise. Modules could "plug into" a basic database system. They could be called, have messages passed to them, and return results to the calling system - ie. the basic database system. For this to occur, agreed communication protocols and message standards must be established. These may be *proprietary standards*, or a *national standard*. The latter is preferable, but usually later in coming. Each of the above resource examples could be a module.

Computing

Adequate computing *hardware* (the boxes and touchable parts), and *software* (the programs) must exist before computers can be used to their potential. The hardware manufacturers provide computers which can do many things. An operating *system* is used to make the hardware "go." A *programming environment* which contains *development tools*, allows the system designer and programmer develop *application programs*. These are sold to the public.

The evolution of programming environments and development tools have enabled complex programs to be written in a relatively short time and at little cost. This evolution will effect healthcare information management. It is a reasonable thought that doctors create their own simple electronic medical record systems using a modern database environment. Doctors and nurses in hospitals are now able to create their own prototype systems. The development and demonstration of their requirements is invaluable to the system designer.

User interface & data capture

The appearance of the computer screen, where things are, how they function and what happens next, can be described as the *human machine interface*, or *user interface*. It is a vital part of quality software, and is, in a way, an art-form and a science. If all else is equal, the interface sells the product. The design should be such that things are where expected, and happen as expected. This is called *intuitive* design. It involves human dynamics, taste and consistency of presentation. A well designed interface can greatly reduce the time taken to learn how to "drive" an application program. Help is usually provided in a manual. The better the software, the less the manual is used. Helpful instructions on the use of the program are provided by *on-line-help*. This should be *context sensitive* (ie. the help changes in accordance with the computer screen or screen object).

Computerised healthcare information management requires that data be captured by the computer. This is traditionally done via the *key board* by typing. This can be fast, accurate and quiet. However it does require keyboard skills. A *pointing device* is helpful and enables (by a "point and click and drag" action) the selection of an object or a part of the screen. Pointing devices include the *mouse*, *trackball*, *drawing pad*, *joy stick* and *touch sensitive screen*, etc.

The chore of typing can be greatly reduced by the automatic entry of "what is probably required" (eg. "today's date" in a date field). These probable answers are called *defaults*. Complex and intelligent defaults may be provided. These can greatly improve data capture accuracy and speed. For example, the list of medications used by a particular patient would make sensible defaults for the medications possibly required at each consultation, (as a repeat of a previous medication is a frequent event).

Complex management *protocols* can be designed to improve data capture, reduce typing, improve accuracy, and improve management. *User defined protocols* and *resource protocols* should be (or will be) a feature of any electronic medical record system.

Data may also be captured by the computer analysis and recognition of *handwriting* and the spoken word (*speech recognition*). These methods require much computing power, are inaccurate, and are currently in their infancy. Some systems however show promise. The smaller the vocabulary, the better the result. Systems can be designed so at any time only a small vocabulary is required, yet the total vocabulary is large.

Bar-coding is widely used in non-medical arenas. It is very easy to create the bar-code equivalent of any text. Reading the printed bar-coded text back into the computer, with the use of a *wand*, is just as easy.

Printed text may be captured from an electronic copy of the document (via a *scanner*) and the application of optical character recognition (*OCR*) software.

Electronic communications

Adequate electronic communications are necessary. More details regarding data communications, standards and privacy issues are provided in other chapters. Some basic requirements are listed here:

- Physical connection between hardware apparatus. Cables are commonly used (telephone lines, twisted pair, co-axial, optic glass fibre transmitting laser light). Rays are not uncommon (infra red, radio and microwaves).
- Organisation and infrastructure. Hardware joined together forms a *network*. If the area involved is *local* (eg. a hospital), then a *local area network* is formed. If the connections involved a *wide area*, then a *wide area network* is the result. Some times the data involved resides on a central computer to form a *central database*. It is possible to *distribute* the data so parts of it reside on many different and widely separated computers. A *distributed database* results. Provided the computer knows the whereabouts of the

required data, a distributed database can be made to behave very much like a central database. The creation, organisation, regulation and maintenance of complex national networks represents a major infrastructure. Expansion of these networks or *data highways* to carry vast quantities of information (eg. a moving picture to every home) has coined the phrase *super data highways*. Nations are currently negotiating their construction. They will have the capacity to carry all the required medical and educational data for the nation's healthcare.

- Standards. An electronic message must be *packaged* so the network knows who sent it and where it is to go. On arrival, the message must be correctly processed. This requires that specific instruction be included within the message. Standard *message formats* and *transmission protocols* are the essential ingredients.

International and national organisations devote their existence to the creation and maintenance of standards. Without them much of our complex civilisation could not function. In the healthcare communication field, the following standards may be encountered:

- OSI or Open Systems Interconnection standards. This specifies a seven layered structure with many options. Options may be specified to form an optimised profile for a specific purpose. One example is GOSIP.
- GOSIP, or the Government OSI Profile, has been specified for public sector use in Australia, Europe, and North America.
- EDI, or Electronic Document Interchange is a standard for message transmission. Currently EDIFACT standard messages are evolving.
- EDIFACT, or EDI For Administration, Commerce and Transport is a message standard adopted by the UK and of interest to Europe and New Zealand and Australia. It has a well defined set of message structures and rules for the development of new messages. Standard healthcare messages are largely undefined at this stage.
- MEDIX, or MEDIcal data eXchange is a language specification for the output of hardware used in healthcare. It has North American origins.
- HL-7, or Health Level-7 refers to the seventh level of the OSI standard. It is a language specification for the output of hardware used in healthcare. North America has widely adopted HL-7. It is of interest to Australia and New Zealand.
- MIB, or Medical Information Bus is a defined data exchange protocol to enable laboratory equipment and computers to be interconnected.
- ASTM or American Standards for Testing and Materials which has a Clinical Data Interchange Standard (E1238).

Safeguards

Adequate security, confidentiality and integrity safeguards must evolve with improved information management. *Security* refers to the ease with which non-authorised people can access data. *Confidentiality* refers to the anonymity of the patient or healthcare provider. *Integrity* refers to the accuracy of the data.

As the data held within a system become more secure and more confidential, the less accessible it becomes to those who have authority to access it. Any system is a compromise.

Security is attained by *locks* (on doors, windows and computers) and *user identification* (by the use of passwords, magnetic stripe badges, smart keys and PIN numbers, security beam emitting badges, thumb print recognition, voice imprint recognition). Once a system has granted access, only that information and those actions allowed by the access code will be made available to the user. Thus an accountant may see all financial data, but no clinical data. The data itself can be disguised by *encryption*. This process jumbles the data according to a formula which contains an access code supplied by an authorised user. The unjumbling process (*decryption*) can occur only if the access code is supplied. A system of *smart keys* (using “smart card technology”) and personal identification numbers (PIN) can be used to encrypt and decrypt data for secure transmission.

Confidentiality may appear to result if only authorised people access data, and patient and provider identifying information has been removed to create anonymity. However, it is possible to electronically *cross match data* from several sources and determine with a possibly high degree of accuracy, the name of the provider or patient. *Legislation* is designed to prevent government authorities unlawfully cross matching data.

Integrity of data involves the accuracy of the data received (eg. from the patient), and the accuracy of its interpretation (eg. by the doctor), and the accuracy with which it is entered into, and stored and transmitted by the computer. Elaborate mathematical data integrity checking occurs as a computer manages its data. During processing, the hardware is unlikely to change data by error without notifying the user. However the program (software) may be instructing the computer incorrectly. To ensure a program acts correctly it must be tested. It is not possible to test a complex program in all combinations of circumstances. Thus, no program is absolutely error free.

Benefits

As stated earlier, information technology should not be applied to healthcare information management unless all concerned benefit - the computing industry, the healthcare funding organisations, the healthcare providers, and the patients.

The computing industry should benefit by the increased sales of computer hardware, software and communication resources.

Funding and administration organisations should benefit by an improvement in the education of, and information available to - the public, students, providers, administrators and epidemiologists. If it is known what is happening and what is needed, resources can be better allocated. A better quality and less costly health service is a potential outcome.

The provider should benefit by a saving of time, by a reduction of tedium, and an increase in the quality of care (because continuing healthcare education becomes less time consuming and tedious, a reduction of errors, excesses and litigation should result).

The patient should benefit from a quality healthcare service which offers efficient communication between providers, organisations and patients. A better informed patient is more equipped for self-care.

References

Australian Government Publishing Services 1994 Medical benefit schedule book. Prepared by: Department of Community Services and Health

Commission on Professional and Hospital Activities 1978 International Classification of diseases, 9th revision, clinical modification (ICD-9-CM). Edward Brothers, Ann Arbor, Michigan

Computer Aided Medical Systems 1993 Read Codes system sevelopers' guide. CAMS, Leicestershire, UK

Computer Aided Medical Systems 1994 Read Codes version 3.1 system sevelopers' guide addendum. CAMS, Leicestershire, UK

Cote RA & Robboy S 1980 Progress in Medical Information Management, Systematised Nomenclature of Medicine (SNOMED). Journal of the American Medical Association, vol 243, no 8, pp 756-762

Cote RA, Rothwell DJ, Beckett RS, Palotay JL & Brochu L 1993 The systematised nomenclature of human and veterinary medicine - SNOMED, international introduction. College of American Pathologists

Humphreys BL and Lindberg DAB 1989 Building the unified medical language system. In: Kingsland L C (ed) Proceedings of the Thirteenth Annual Symposium on Computer Applications in Medical Care, Washington DC. IEEE Computer Society Press, pp 475-476

Hupp JA, Cimino JJ, Hoffer EF, Lowe HJ, Barnett GO 1986 DX-plain - a computer-based diagnostic knowledge base. In: Proceedings of the Fifth World Conference on Medical Informatics (MEDINFO 86). Amsterdam: Norther-Holland, pp 117-121

Lamberts H & Wood M (eds) 1987 ICPC Internationl Classification of Primary Care. Oxford University Press, Oxford

McCray AT 1989 The UMLS semantic network. In: Kingsland L C (ed) Proceedings of the Thirteenth Annual Symposium on Computer Applications in Medical Care, Washington DC. IEEE Computer Society Press, pp 503-505

Miller RA, McNeil MA, Challinor SM, MasarieFE Jr, Myers JD 1986. The INTERNIST-1 QUICK MEDICAL REFERENCE project - status report. Western Journal of Medicine, vol 145, pp 816-822

National Library of Medicine 1994 Medical subject headings, annotated alphabetical list. NLM, Bethesda, Maryland

National Library of Medicine, US Department of Health and Human Services and National Institutes of Health 1990 General description of META-1. In: UMLS knowledge sources experimental edition September 1990, Documentation Draft, pp 5-7

Sherertz D, Tuttle M, et al 1989 A hypercard implementation of meta-1: the first version of the umls metathesaurus. In: Kingsland L C (ed) Proceedings of the Thirteenth Annual Symposium on Computer Applications in Medical Care, Washington DC. IEEE Computer Society Press, pp 1017-1018

Warner HR, Haug P, Bouhaddou O, Lincoln M, Warner H Jr, Sorenson D, Williamson JW & Fan C 1988 ILIAD as an expert consultant to teach differential diagnosis. In: Greenes RA (ed) Proceedings of the Twelfth Annual Symposium on Computer Applications in Medical Care, Washington DC. IEEE Computer Society Press, pp 371-376

World Health Organization 1977 International classification of diseases - manual of the international statistical classification of diseases, injuries, and causes of death, ninth revision. WHO, Geneva