ABC of health informatics
# What is health information?

Jeremy C Wyatt, Frank Sullivan

> **This is the first in a series of 12 articles**
> **A glossary of terms is available on bmj.com**

Information is an ethereal commodity. One definition describes it as the data and knowledge that intelligent systems (human and artificial) use to support their decisions. Health informatics helps doctors with their decisions and actions, and improves patient outcomes by making better use of information—making more efficient the way patient data and medical knowledge is captured, processed, communicated, and applied. These challenges have become more important since the internet made access to medical information easier for patients.

This ABC series focuses on information handling during routine clinical tasks, using scenarios based on Pendleton's seven-stage consultation model (see box opposite). The articles cover wider issues arising from, and extending beyond, the immediate consultation (see box below). Questions on clinical information that often arise in clinical and reflective practice are dealt with, but discussion of specific computer systems is avoided (a glossary of terms appears on bmj.com, and other glossaries are listed in the box at the end of this article).
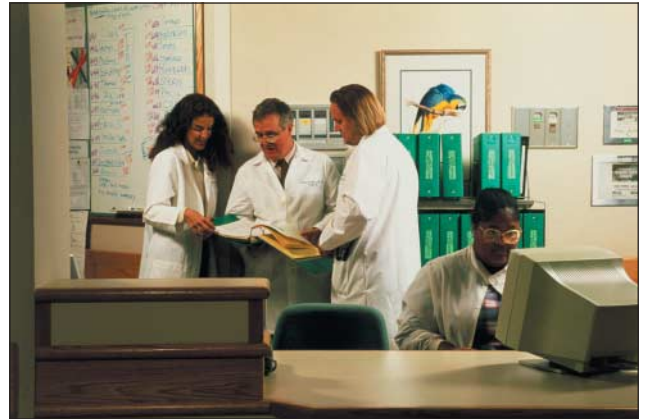
JOHN GREIM/SPL

Some questions on clinical information

Medical record keeping
- What records to keep?
- In what format?
- What data to enter, and how?
- How to store records, and for how long?
- With whom to share the record?

How to use the information records contain
- To manage my patients?
- To audit and improve my service?
- To support my research?
- To feed another information system?

How to communicate with my colleagues and patients
- Face to face?
- On paper?
- Using the internet?

Clinical knowledge sources
- What knowledge sources are out there, and how to select them?
- How to use these sources to answer my own, and my team's, clinical questions?
- How to keep knowledge and skills up to date?
- How to use knowledge to improve my own, and my team's, clinical practice?

**Pendleton's consultation model, adapted for ABC series**
- Discover the reason for the patient's attendance
- Consider other problems
- Achieve a shared understanding of the problems with the patient
- With the patient, choose an appropriate action for each problem
- Involve the patient in planning their management
- Make effective use of the consultation
- Establish or maintain a relationship with the patient

> **Ms Smith is a 58 year old florist with a 15 year history of renal impairment caused by childhood pyelonephritis who is experiencing tiredness and muscle cramps. She has sought medical attention for similar problems in the past, and is considering doing so again**

Capturing and using information

Consider the different forms that information can take, where each form comes from, its cost, and how to assess the quality of the information. These issues arise during a general practitioner's (Dr McKay) encounter with Ms Smith.

Dr McKay applies her own clinical knowledge and skill, perhaps augmented by a textbook or other knowledge source, to capture relevant data from Ms Smith. Dr McKay browses Ms Smith's record to check her medical history. She updates the record and either takes action herself, or telephones a consultant nephrologist (Dr Jones), who suggests 1α-hydroxy cholecalciferol 0.5 μg daily for Ms Smith. Dr McKay then follows up the telephone conversation with the consultant by issuing an electronic prescription. The prescription transfers through a secure local network to Ms Smith's usual pharmacist
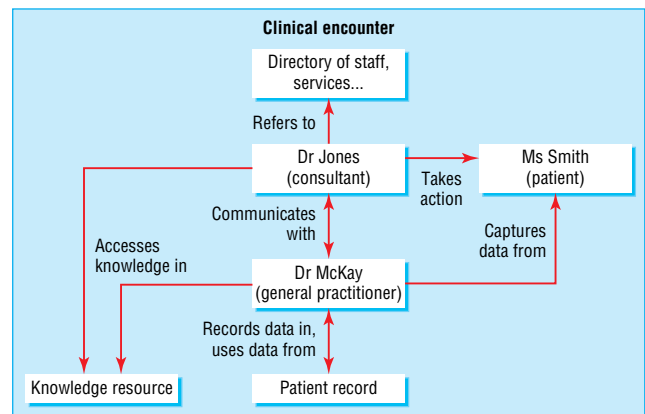
Information flows in a clinical environment

along with a formal online outpatient referral request. Dr Jones checks a hospital phone directory on the web before referring Ms Smith to the dietician for a low calcium diet. Ms Smith is kept informed of these developments by telephone before her appointment the next week.

Representing, interpreting and displaying information

When Dr McKay reads Ms Smith's patient record what she sees on the page is not actually information, but a representation of it. A "real" item of information, such as the fact that Ms Smith has hypercalcaemia, is distinct from how that item is represented in an information system (for example, by selecting Ms Smith's record and writing "Hypercalcaemia," or choosing a Read code that updates Ms Smith's computer-based record). The real information is also distinct from a person's interpretation of it, which might resemble a fragment in a stream of consciousness, "Remember to check on Ms Smith—calcium problem back again." These distinctions reflect common sense and semiotic theory: real things only exist in the physical world, and each person interprets them in private and associates their own images with them.

Back in the clinical world, the lesson is that we should capture and represent each item of information in a form that helps each user—whether human or computer—to find and interpret it. The next time Dr McKay logs into Ms Smith's computer record, although Ms Smith's serum calcium may be represented internally in the computer as the real number 2.8, on the computer screen it can be shown as a figure, a red warning icon, a point on a graph showing all her calcium results, or as the words "Severe hypercalcaemia" in an alert. These display formats can all be achieved with a paper record, but it would take more time and effort to annotate abnormal laboratory results with a highlighter pen, graph the values on a paper chart, or write an alert on a Post-it note and place this on the front of Ms Smith's record.

Selecting a format is important because it determines how to represent each item of information in a system, and in turn how each item is captured. When information is captured and represented on paper or film, it is hard to change the order in which each item appears or to display it in other formats. When information is captured and stored on a computer, however, it can be shown in a different order or grouped in different ways. When data is coded and structured, or broken down into simple elements, it can be processed automatically—for example, the computer can add the icon, graph the data, or generate the alert about Ms Smith.

Sources of clinical information

Clinicians use three types of information to support patient care: patient data, medical knowledge, and "directory" information. This description ignores two questions, however: where does the knowledge in a textbook come from, and how do we improve on the methods used to manage patients? Patient data are the source in both cases (see box opposite). Local problems—such as an adverse event or failure to implement a guideline that everyone agrees to apply to their patients—can be picked up by quality improvement activities such as clinical governance. In well organised clinical environments and specialties, a registry is used to capture patient experiences and monitor for adverse outcomes.

Sometimes, however, patient data are used to suggest, or even answer, more general questions—for example, about drug

### Common sense meets semiotic theory

In her shop, Ms Smith sells a kind of flower that grows on shrubs with prickly stems and serrated leaves. Humans use consistent symbols to represent these things (for example, "rose; roos"). However, each person privately adds their own connotations to these symbols

Some definitions of rose from *Chambers 21st Century Dictionary*
- An erect or climbing thorny shrub that produces large, often fragrant, flowers that may be red, pink, yellow, orange, or white, or combinations of these colours, followed by bright-coloured fleshy fruits
- The national emblem of England
- A light pink, glowing complexion (put the roses back in one's cheeks)
- A perforated nozzle, usually attached to the end of a hose, watering can, or shower head that makes the water come out in a spray

Possible formats to display information include informal or structured text, tables, graphs, sketches, and images. The best format for each item of information depends on who will use it, how they will use it, for what task, and on the formats readily available. With permission from Klaus Gulbrandsen/SPL
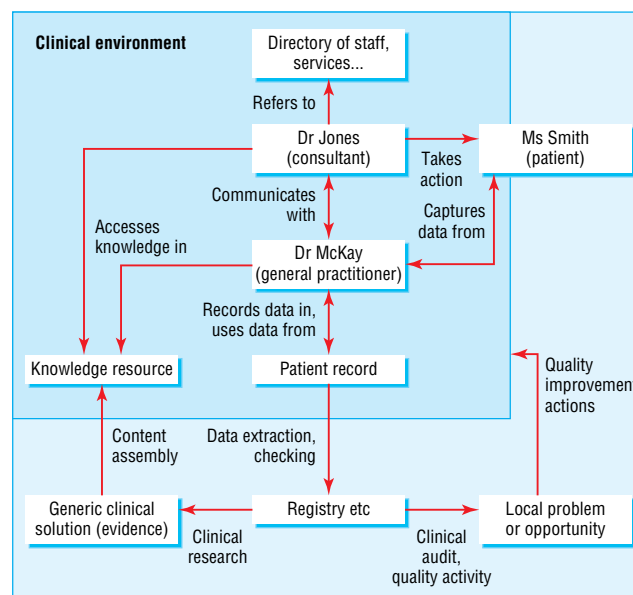
Information flows in clinical and non-clinical environments

effectiveness, disease aetiology, or the accuracy of tests. The results should be high quality, generic evidence that can be safely applied outside the specific clinical environment that is being studied. Often, this evidence is published as if it were the final word. Clinical epidemiology shows us, however, that the results of a single study often differ substantially from the "truth." Well conducted systematic reviews of all rigorous, relevant studies are a better approximation, and are an example of the content assembly methods used to develop good quality knowledge resources.

The costs of information

To a businessman, information must seem the ultimate product: once it is captured, it can be sold any number of times without using up the original supply. Unfortunately for clinicians, each item of information that is captured, processed, and displayed has an associated cost or risk. By choosing to code the current problem as chronic pyelonephritis only (see figure above), Dr McKay fails to record the endocrine dimension with potential loss of explanatory power for others looking at Ms Smith's records. Entering more than one code takes extra time and may cause difficulties in interpretation for secondary use of the data.

Information costs are especially high for data captured by health professionals in the structured, coded representation often required by computerised record systems. If the information is only ever going to be read by humans, it should not be captured as structured data because this will discourage doctors from recording useful free text that computers do not need to "understand"—for example, "Ms Smith is going to Spain for a holiday, her cat died last week." All patient record systems should allow easy entry of such unstructured text (perhaps by voice recognition) to support the human side of medicine, and to help maintain the therapeutic relationship with patients.

Assessing the quality of information

Imagine that Dr Jones is auditing outcomes in his hypercalcaemic patients and wishes to include Ms Smith's data. Is her data of adequate quality for this task?

Information only exists to support decisions and actions: if it fails to do this, it is irrelevant noise. The aims of clinical audit are to understand current practice and suggest appropriate actions for the future. If the data are full of errors or incomplete, refer to patients seen years ago, or cannot be interpreted by the user, they are unlikely to help. More subtly, if useful data items are present—for example, serum calcium—but vital context is omitted, such as serum albumin or current treatment, it is still hard to use the data. Without this context, information is often useless; with it, data collected for one purpose can often, but not always, be used for another.

Glossaries for informatics terms

- Coiera E. *Guide to health informatics.* 2nd ed. London: Hodder Arnold, 2003. www.coiera.com/glossary.htm (accessed 26 August 2005)
- Wyatt JC, Liu J. Basic concepts in medical informatics. http://jech.bmjjournals.com/cgi/content/full/56/11/808 (accessed 26 August 2005)

Jeremy C Wyatt is professor of health informatics, and Frank Sullivan is NHS Tayside professor of research and development in general practice and primary care, University of Dundee.

The series will be published as a book by Blackwell Publishing in spring 2006.
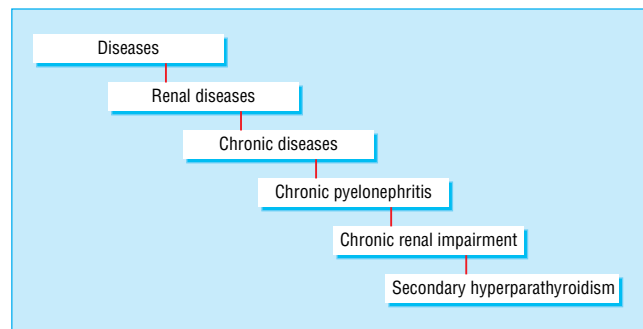
Partial hierarchy of diseases

Quality criteria for patient data

Criterion	How to test it	Comment
Accurate	Comparison with a gold standard source of data—for example, the patient	Technically, validity—does the data item measure what it is meant to? Reliability is a related concept—do two observers agree on the data item?
Complete	Per cent missing data at a given point	Often difficult to estimate without access to multiple sources of information
Timely	Delay from the event the data describes to its availability for use on the information system	Unless data are available at the point they are needed to inform decisions, fulfilling the other criteria is almost worthless
Relevant	Amount that data alter decisions or actions of the user; the impact of leaving an item out of the dataset	Unless data are relevant to information users, they contribute to information overload
Appropriately represented	Degree of structuring and coding of items	Depends on the user of the item and their needs
Relevant detail included	If data are detailed enough to support decisions	Highly dependent on the purpose and confidentiality of the information
Relevant context included	Is there enough context (for example, date patient seen, by whom) to support appropriate interpretation of data?	A key issue, only partially solved in current electronic patient records

Further reading

- Hersh W. What is Medical Informatics? www.ohsu.edu/dmice/whatis/index.shtml (accessed 26 August 2005)
- Pendleton D, Schofield T, Tate P, Havelock P. *The consultation: an approach to learning and teaching.* Oxford: Oxford University Press, 1987
- Nygren E, Wyatt JC, Wright P. Medical records 2: helping clinicians find information and avoid delays. *Lancet* 1998;352:1462-6
- Morris AD, Boyle DI, MacAlpine R, Emslie-Smith A, Jung RT, Newton RW, et al. The diabetes audit and research in Tayside Scotland (DARTS) study: electronic record linkage to create a diabetes register. DARTS/MEMO Collaboration. *BMJ* 1997;315:524-8
- Naylor CD. Grey zones of clinical practice: some limits to evidence based medicine. *Lancet* 1995;345:840-2
- Brody H. *Stories of sickness.* Yale: Yale University Press, 1987
- Tanenbaum SJ. What physicians know. *N Engl J Med* 1996.329:1268-71
- van Bemmel JH, Musen MA, eds. *Handbook of medical informatics.* London: Springer, 1997 www.mihandbook.stanford.edu/handbook/home.htm (accessed 26 August 2005)