

GLOSSARY

Basic concepts in medical informatics

J C Wyatt, J L Y Liu

This glossary defines terms used in the comparatively young science of medical informatics. It is hoped that it will be of interest to both novices and professionals in the field.

Medical informatics is the study and application of methods to improve the management of patient data, clinical knowledge, population data, and other information relevant to patient care and community health. It is a young science, which emerged in the decades after the invention of the digital computer in the 1940s. Mechanical computing in medicine had a much earlier origin, in the 19th century, with Herman Hollerith's "punched-card data-processing system" originally used for the US census and subsequently developed to support surveys in public health and epidemiology.¹ This example reflects the multidisciplinary nature of medical informatics, which interacts with various fields, including the clinical sciences, the public health sciences (for example, epidemiology and health services research) as well as cognitive, computing, and information sciences.

Given the diversity of backgrounds of medical informatics workers, newcomers can easily be confused by the jargon used in the field. An introduction to basic concepts would therefore be useful to those interested in learning more about medical informatics. In recent years, various branches of the discipline have appeared, including public health informatics, consumer health informatics, and clinical informatics. In a debate on whether medical informatics and these branches are distinct disciplines, Shortliffe and Ozbolt argued that "informatics is built on a re-usable and widely applicable set of methods that are common to all health science disciplines, and that 'medical informatics' continues to be a useful name for a composite core discipline that should be studied by all students, regardless of their health profession orientation."^{2,3} Our definitions of the various branches of medical informatics below reflect this.

One or more of the following criteria were used to select terms included in this list of concepts:

- The term is likely to be novel to epidemiologists and public health professionals.
- An existing word may have a widely accepted meaning but is used in a specific way in the medical informatics field.
- The term is relevant to epidemiology and public health.
- It is essential to understanding medical informatics.
- It is enduring—that is, not about some transient technology.

J Epidemiol Community Health 2002;**56**:808–812

- There is general agreement on what the term means and how it is used.
- There is value in attempting to define the term, even though there may be debate about it.

We trust this list of terms will be informative for novices and encourage discussion among medical informatics professionals, particularly on terms for which there is no widespread agreement*. For readers who are interested in pursuing the subject further, additional resources are cited at the end of the list.

ALGORITHM

A process for carrying out a complex task broken down into simple decision and action steps. Often assists the *requirements analysis* process carried out before programming.

BIOINFORMATICS

The use of *medical informatics* methods to facilitate research in molecular biology.

CHECKLIST

A type of *clinical decision tool*: a form listing one or more items of *patient data* to be collected before, during or after an encounter; can be paper or computer based.

CLINICAL CODING SYSTEM (CLINICAL THESAURUS, CONTROLLED VOCABULARY)

A limited list of preferred terms from which the user can draw one or more to express a concept such as *patient data*, a disease or drug name, etc. An alphanumeric code corresponding to the term is then stored by the computer. Synonyms or close matches to each preferred term are usually available and map onto the same internal codes. This approach makes it easier for a computer to analyse data than the use of free text words or phrases. Examples of clinical coding systems include SNOMED-CT (divergent codes used to capture *patient data*), MeSH (terms used to index biomedical literature) and ICD-10 (convergent disease codes for international comparisons, with specific rules to guide coders). Clinical coding systems play a key part in epidemiological studies and health service research, from the use of MeSH terms to conduct literature searches for systematic reviews to studies that use ICD codes to classify and compare diseases. To prevent information loss, it is vital that the terms and

* Notes to the list of concepts: *Italic* means "see also". Synonyms are mentioned in parentheses, after the core term.

See end of article for authors' affiliations

Correspondence to:
Professor J C Wyatt,
Department of Medical
Informatics, Academic
Medical Centre, University
of Amsterdam,
Netherlands;
j.c.wyatt@amc.uva.nl

codes are never changed or dropped, only added to. Obsolete terms can be marked as such to deter inappropriate use. Continuing maintenance is needed to incorporate new terms and codes for new concepts and synonyms as they arise.

CLINICAL DATA SYSTEM

Any information system concerned with the capture, processing, or communication of *patient data*.⁴

CLINICAL DECISION TOOL

Any mechanical, paper, or electronic aid that collects or processes data from an individual patient to generate output that aids clinical decisions during the doctor-patient encounter.⁵ Examples include *decision support systems*, paper or computer *reminders* and *checklists*, which are potentially useful tools in *public health informatics*, as well as other branches of medical informatics.

CLINICAL INFORMATION

Organised *patient data* or *medical knowledge* used to make clinical decisions (adapted from Shortliffe *et al*⁶); may also include *directory information*. Many activities in public health and epidemiology (for example, surveillance systems, cohort studies to assess the effects of a risk factor of disease, and clinical trials to estimate efficacies of new treatments) entail the organisation of such data (for example, case report forms for individual patients) into useable information (for example, incidence of notifiable cases of disease from surveillance programmes and summary evidence from cohort studies or clinical trials, expressed as odds ratios for certain harmful and beneficial outcomes). See also: *information*.

CLINICAL INFORMATICS

The use of *medical informatics* methods to aid management of patients using an interdisciplinary approach, including the clinical and information sciences.⁶

COMMUNICATION

The exchange of *information* between agents (human or automated) face to face or using paper or electronic media.⁷ Requires the use of a shared language and understanding or common ground.

COMPUTER VISION (IMAGE INTERPRETATION)

The use of computer techniques to assist in the interpretation of images, such as mammograms.

CONFIDENTIALITY (PROTECTING PRIVACY)

The policies restricting access to a person's data to those whom the patient agrees need access to them, except rarely in emergency and for the public good (for example, to contain epidemics, allow important research to be undertaken, or solve serious crime). In addition, other regulatory and institutional approval may be needed (for example, the need to seek consent from medical ethics committees or relevant national authorities). In recent years, leading public health researchers have warned that legislation enacted to protect patients' medical data in the UK, Europe, and US could potentially hamper observational research and medical record linkage studies.^{8,9}

CONSUMER HEALTH INFORMATICS

The use of *medical informatics* methods to facilitate the study and development of paper or electronic systems that support public access to and use of health and lifestyle information. For additional discussion on the scope of consumer health informatics, see Eysenbach.¹⁰ See also *eHealth*.

DATA QUALITY

The degree to which data items are accurate, complete, relevant, timely, sufficiently detailed, appropriately represented (for example, consistently coded using a *clinical coding system*), and retain sufficient contextual information to support decision making.

DATABASE

A collection of data in machine readable format organised so that it can be retrieved or processed automatically by computer. A flat file database is organised like a card file, with many records (cards) each including one or more fields (data items). A relational database is organised as one or more related tables, each containing columns and rows. Data are organised in a database according to a schema or data model; some items are often coded using a *clinical coding system*.

DECISION SUPPORT SYSTEM (COMPUTER DECISION AID)

A type of *clinical decision tool*: a computer system that uses two or more items of *patient data* to generate case specific or encounter specific advice.¹¹ Examples include computer risk assessors to estimate cardiovascular disease risk¹² and the Leeds Acute Abdominal Pain system, which aided the diagnosis of conditions causing such pain.¹³ Evidence adaptive decision support systems are a type of decision aid with a knowledge base that is constructed from and continually adapts to new research based and practice based evidence.¹⁴

DECISION TREE

A way to model a complex decision process as a tree with branches representing all possible intermediate states or final outcomes of an event. The probabilities of each intermediate state or final outcome and the perceived utilities of each are combined to attach expected utilities to each outcome. The science of drawing decision trees and assessing utilities is called decision analysis.

DEMONSTRATION STUDY

Study that establishes a relation—which may be associational or causal—between a set of measured variables. In epidemiology, cohort studies, randomised trials, and blind comparisons of a test with a gold standard are typical demonstration studies.^{6,15} See also *measurement studies*.

DIRECTORY INFORMATION

Information specific to an organisation or service that is useful in managing public health services, health care services, or patients. Examples include a phone directory, a lab handbook listing available tests and specimens to use, and a list of drugs in the local formulary.

eHealth

The use of internet technology by the public, health workers, and others to access health and lifestyle information, services and support; it encompasses *telemedicine*, *telecare*, etc. For discussion on the scope and security issues of eHealth, see a recent report by the National HealthKey Collaborative.¹⁶

ELECTRONIC HEALTH RECORD (EHR)

In the UK, the lifelong summary of a person's health episodes, assembled from summaries of individual *electronic patient records* and other relevant data.¹⁷

ELECTRONIC PATIENT RECORD (EPR)

A computer based *clinical data system* designed to replace paper patient records.

EXPLICIT KNOWLEDGE

Knowledge that can be communicated on paper or electronically, without person to person contact.¹⁸ Public health workers and physicians cannot use explicit knowledge if they cannot access it. There is thus a need to identify, capture, index, and make available explicit knowledge for professionals, a process called codification. Much of the work done by the Cochrane Collaboration entails codification of explicit knowledge. See also: *tacit knowledge*.

GEOGRAPHICAL INFORMATION SYSTEM (GIS)

Computer software that captures, stores, processes, and displays location as well as other data. The display may preserve distance ratios between data objects (for example, true scale maps) or link similar objects, ignoring distance (for example, topological maps such as that distributed to the public for the London Underground). GIS software is used in many ecological studies of disease. A famous example is Peto's study of diet, mortality, and lifestyle in rural China.¹⁹ Disease mapping studies have also been conducted to assess childhood leukaemia in areas with different radon levels,²⁰ the clustering of respiratory cancer cases in areas with a steel foundry,²¹ and socioeconomic gradients in infant mortality.²² GIS are also used for public health planning and surveillance purposes at local or national health departments. Care should be taken by policy makers in interpreting maps produced by GIS software, particularly in regard to the ecologic fallacy.²³

INFORMATION

Organised data or knowledge used by human and computer agents to reduce uncertainty, take decisions, and guide actions (adapted from Shortliffe *et al*⁶ and Wyatt²⁴). See also: *clinical information*, *patient data*, *medical knowledge*.

INFORMATION DESIGN

The science and practice of designing forms, reports, computer screens, etc, so that the *information* they contain can be found rapidly and interpreted without error (adapted from Sless²⁵). Information design is based on psychological and graphical design theories and empirical studies of human perception and decision making using alternative formats for *information*.^{25a}

KNOWLEDGE BASE

A store of knowledge represented explicitly so that a computer can search and reason with it automatically; often uses a *clinical coding system* to label the concepts. See also *decision support system*.

KNOWLEDGE BASED SYSTEM (EXPERT SYSTEM)

A computer *decision support system* with an explicit *knowledge base* and separate reasoner program that uses this to give advice or interpret data, often *patient data*.

KNOWLEDGE MANAGEMENT

The identification, mobilisation and use of knowledge to improve decisions and actions. In public health and medicine, much of this work involves the management of *medical knowledge* (from epidemiological studies, randomised controlled trials, and systematic reviews) so that it is actually used by the physician. This entails *practice innovation*²⁶ or narrowing the gap between what we know and what we do. The NHS is developing a program of knowledge codification to inform routine problem solving, for example, through the National Electronic Library of Health, guidelines from the National Institute of Clinical Excellence (NICE), and care pathways and triage algorithms used in the NHS Direct Clinical Advice System.²⁷

MEASUREMENT STUDY

Study of the reliability, validity, or ease of use of a measurement instrument or method in a defined population.¹³ See also *demonstration study*.

MEDICAL INFORMATICS (HEALTH INFORMATICS)

The study and application of methods to improve the management of *patient data*, *medical knowledge*, population data and other *information* relevant to patient care and community health. Unlike some other definitions of medical informatics (or example, Greenes and Shortliffe²⁸), this definition puts the emphasis on information management rather than technology. Branches of medical informatics include *bioinformatics*, *clinical informatics*, *consumer health informatics* and *public health informatics*.

MEDICAL KNOWLEDGE (CLINICAL KNOWLEDGE)

Information about diseases, therapies, interpretation of lab tests, etc, which is potentially applicable to decisions about multiple patients and public health policies, unlike *patient data*. This information should where possible be based on sound evidence from clinical and epidemiological studies, using valid and reliable methods. See also: *explicit knowledge*, *tacit knowledge*, *knowledge management*.

MINIMUM DATASET

A list of the names, definitions and sources of data items needed to support a specific purpose, such as surveillance of the health of a community, investigation of a research hypothesis or monitoring the quality of care in a *registry*.

OBJECTIVISTIC EVALUATION

An evaluation approach that uses experimental designs and statistical analyses of quantitative data.⁶ ¹⁵ Such an approach is never completely objective. See also *subjectivistic study*.

ONTOLOGY

A description of the concepts and relations in a domain, such as drug prescribing. Sample concepts here would be "patient", "prescriber", and "drug"; relevant relations might include "prescribes to", "requests prescription from", and "causes side effects to". A taxonomy or hierarchy is a simple kind of ontology in which concepts are arranged according to only one relation: "is a kind of". Note that ontology as used here has a different meaning from its use in the philosophy of science, an area of interest to theoretical epidemiologists.

PATIENT DATA

Information about an individual patient and potentially relevant to decisions about her current or future health or illness. Patient data should be collected using methods that minimise systematic and random error. See also: *medical knowledge*, *data quality*.

PRACTICE INNOVATION (RESEARCH IMPLEMENTATION, BEHAVIOUR CHANGE)

Applying usually *explicit knowledge* to innovate in public health and clinical practice by identifying barriers to change and applying appropriate *practice innovation methods* to overcome these (adapted from Wyatt²⁶). See also: *knowledge management*.

PRACTICE INNOVATION METHODS (RESEARCH IMPLEMENTATION, BEHAVIOUR CHANGE METHODS)

A range of methods that can potentially overcome barriers to change in the *practice innovation* process, such as clinical audit, outreach visits or *clinical decision tools*.

PUBLIC HEALTH INFORMATICS

The use of *medical informatics* methods to promote public health practice, research, and learning, using an interdisciplinary approach, including the public health sciences, for example, epidemiology and health services research, and the information sciences, for example, computing science and technology (adapted from Yasnoff *et al*²⁹). In a recent paper outlining an agenda for developing this branch of informatics, Yasnoff *et al*³⁰ argued for the need to construct, implement, and integrate public health surveillance systems at national and local levels, to enable rapid identification and response to disease hotspots (and more topically, bioterrorism). As Yasnoff rightly points out, methods of assessing costs and benefits of such systems are needed. Public health informatics can also contribute in other areas, for example, *reminders* have played an important part in prevention programmes such as smoking cessation advice to smokers³¹ and the use of preventive care for patients.³²

REGISTRY

A *database* and associated applications that collects a *minimum dataset* on a specified group of patients (often those with a certain disease or who have undergone a specific procedure), health professionals, organisations, or even clinical trials. Registries can be used to explore and improve the quality of care or to support research, for example to monitor long term outcomes or rare complications of procedures. Key issues in registries are maintaining *confidentiality*, coverage of the target population, and *data quality*.

REMINDER

A type of *clinical decision tool* that reminds a doctor about some item of *patient data* or *clinical knowledge* relevant to an individual patient that they would be expected to know. Can be paper based or computer based; includes *checklists*, sticky labels on front of notes, an extract from a guideline placed inside notes, or computer based alerts. There has been much interest in reminders as a *practice innovation method* recently because of the poor uptake of practice guidelines, even those based on good quality evidence. An example is the treatment of dyslipidaemia in primary care, where there is a big gap between recommendations and actual practice.³³

REQUIREMENTS ANALYSIS

The process of understanding and capturing user needs, skills, and wishes before developing an information system (adapted from Somerville³⁴). See also *software engineering*.

SECURITY

The technical methods by which *confidentiality* is achieved.¹⁶

SOFTWARE ENGINEERING

The process of system development, documentation, implementation, and upgrading (adapted from Somerville³⁴). In the classic or “waterfall” model of software engineering, *requirements analysis* leads to a document that serves as the basis for a system specification and *database* schema, from which programmers work to develop the software. However, increasingly, users and software designers work together from the start to develop and refine a prototype system. This helps to engage the users, educate the software development team, brings the requirements documents alive, and allows users to explore how their requirements might change as a result of interaction with the new software.

SUBJECTIVISTIC EVALUATION

An evaluation approach that relies primarily on qualitative data derived from observation, interview, and analysis of documents and other artefacts.^{6 15} The focus of such studies is on description and explanation; they tend to evolve rather

than be prescribed in advance. As we can never truly understand another person’s feelings, such studies always approximate the subjective. See also *objectivistic study*.

TACIT KNOWLEDGE (INTUITION)

Knowledge that requires person to person contact to transfer and cannot be communicated on paper or electronically.^{18 27} Over time, some tacit knowledge can be analysed, decomposed, and made explicit. See also: *explicit knowledge*.

TELECARE

A kind of *telemedicine* with the patient located in the community (for example, their own home); see also *eHealth*.

TELEMEDICINE

The use of any electronic medium to mediate or augment clinical consultations. *Telemedicine* can be simultaneous (for example, telephone, videoconference) or store and forward (for example, an email with an attached image). See also *eHealth*.

Additional resources

Readers who are interested in general coverage of the field of medical informatics are encouraged to refer to standard texts.^{15 35 36} Those who are interested in alternative or complementary definitions of the above terms can look up various sources.^{6 7 37–39}

ACKNOWLEDGEMENT

We thank Ameen Abu Hanna (Department of Medical Informatics, University of Amsterdam) and the *JECH* anonymous referees, who all provided useful comments on drafts.

Authors’ affiliations

J C Wyatt, Department of Medical Informatics, Academic Medical Centre, University of Amsterdam, Netherlands

J LY Liu, Centre for Statistics in Medicine, Institute of Health Sciences, Oxford University, UK

Funding: the NHS R&D Health Technology Assessment Board funded part of this work.

Conflicts of interest: none.

REFERENCES

- 1 Shortliffe EH, Blois MS. The computer meets medicine and biology: emergence of a discipline. In: Shortliffe EH, Perreault LE, Wiederhold G, *et al*, eds. *Medical Informatics—computer applications in health care and biomedicine*. New York: Springer-Verlag, 2001:3–40.
- 2 Masys DR, Brennan PF, Ozbolt JG, *et al*. Are medical informatics and nursing informatics distinct disciplines? The 1999 ACMI debate. *J Am Med Inform Assoc* 2000;**7**:304–12.
- 3 Shortliffe EH, Garber AM. Training synergies between medical informatics and health services research. *J Am Med Inform Assoc* 2002;**9**:133–9.
- 4 Wyatt JC. Clinical data systems I: data and medical records. *Lancet* 1994;**344**:1543–7.
- 5 Liu JLY, Wyatt JC, Altman DG. Exploring the definition and scope of clinical decision tools: focus on the problem, not the solution. Oxford: Working paper, Centre for Statistics in Medicine, Oxford University, 2002.
- 6 Shortliffe EH, Perreault LE, Wiederhold G, *et al*. Glossary. In: *Medical informatics—computer applications in health care and biomedicine*. New York: Springer-Verlag, 2001:749–820.
- 7 Van Bommel JH, Musen M. Glossary. In: *A handbook of medical informatics*. Heidelberg: Springer-Verlag, 1997:557–603.
- 8 Lawlor DA, Stone T. Public health and data protection: an inevitable collision or potential for a meeting of minds. *Int J Epidemiol* 2001;**30**:1221–5.
- 9 Walton of Detchant, Doll R, Asscher W, *et al*. Consequences for research if use of anonymised patient data breaches confidentiality. *BMJ* 1999;**319**:1366.
- 10 Eysenbach G. Consumer health informatics. *BMJ* 2000;**320**:1713–16.
- 11 Wyatt JC, Spiegelhalter D. Field trials of medical decision aids: potential problems and solutions. *Proc 15th Annu Symp Comput Appl Med Care* 1991;3–7.
- 12 Hingorani AD, Vallance P. A simple computer programme for guiding management of cardiovascular risk factors and prescribing. *BMJ* 1999;**318**:101–5.

- 13 **Adams ID**, Chan M, Clifford PC, *et al.* Computer aided diagnosis of acute abdominal pain: a multicentre study. *BMJ* 1986;**318**:101–5.
- 14 **Sim I**, Gorman P, Robert A, *et al.* Clinical decision support systems for the practice of Evidence-based medicine. *J Am Med Inform Assoc* 2001;**8**:527–34.
- 15 **Friedman CP**, Wyatt JC. *Evaluation methods in medical informatics*. New York: Springer-Verlag, 1997.
- 16 **National HealthKey Collaborative**. *Securing the exchange and use of electronic health information to improve the nation's health: a summary report to the community*. New York: The Robert Wood Johnson Foundation, 2001.
- 17 **Burns F**. *Information for health*. Leeds: NHS Executive, 1998.
- 18 **Wyatt JC**. Knowledge for clinicians 10: management of explicit and tacit knowledge. *J R Soc Med* 2001;**94**:6–9.
- 19 **Chen J**, Campbell TC, Li J, *et al.* *Diet, lifestyle and mortality in China*. Oxford: Oxford University Press, 1990.
- 20 **Kohli S**, Brage HN, Lofman O. Childhood leukaemia in areas with different radon levels: a spatial and temporal analysis using GIS. *J Epidemiol Community Health* 2000;**54**:822–6.
- 21 **Lloyd OL**. Respiratory-cancer clustering associated with localised industrial air pollution. *Lancet* 1978;*i*:318–20.
- 22 **Wong TW**, Wong SL, Yu TS, *et al.* Socio-economic correlates of infant mortality in Hong Kong using ecologic data 1979–1993. *Scand J Soc Med* 1998;**26**:281–8.
- 23 **Richards TB**, Croner CM, Rushton G, *et al.* Geographic information systems and public health: mapping the future. *Public Health Rep* 1999;**114**:359–73.
- 24 **Wyatt JC**. Medical informatics: artefacts or science? *Methods Inf Med* 1996;**35**:197–200.
- 25 **Sless D**. What is information design? In: *Designing information for people*. Sydney: Communications Research Press, 1994:1–16.
- 25a **Wyatt JC**, Wright P. Medical records. I: design should help use of patient data. *Lancet* 1998;**352**:1375–8.
- 26 **Wyatt JC**. Practice guidelines and other support for clinical innovation. *J R Soc Med* 2000;**93**:299–304.
- 27 **Wyatt JC**. *Clinical knowledge and practice in the information age: a handbook for health professionals*. London: Royal Society of Medicine Press, 2001.
- 28 **Greens RA**, Shortliffe EH. Medical informatics: an emerging academic discipline and institutional priority. *JAMA* 1990;**263**:1114–20.
- 29 **Yasnoff WA**, O'Carroll PW, Koo D, *et al.* Public health informatics: improving and transforming public health in the information age. *J Public Health Manag Pract* 2000;**6**:67–75.
- 30 **Yasnoff WA**, Overhage JM, Humphreys BL, *et al.* A national agenda for public health informatics. *J Am Med Inform Assoc* 2002;**9**:535–45.
- 31 **Law M**, Tang JL. An analysis of the effectiveness of interventions intended to help people stop smoking. *Arch Intern Med* 1995;**155**:1933–41.
- 32 **Dexter PR**, Perkins S, Overhage JM, *et al.* A computerized reminder system to increase the use of preventive care for hospitalized patients. *N Engl J Med* 2001;**345**:965–70.
- 33 **Monkman D**. Treating dyslipidaemia in primary care: the gap between policy and reality is large in the UK. *BMJ* 2000;**321**:1299–300.
- 34 **Somerville I**. *Software engineering*. 5th edn. Wokingham: Addison-Wesley, 1995.
- 35 **Shortliffe EH**, Perreault LE, Wiederhold G, *et al*, eds.. *Medical informatics—computer applications in health care and biomedicine*. New York: Springer-Verlag, 2001.
- 36 **Coiera E**. *A guide to medical informatics, the internet and telemedicine*. New York: Chapman and Hall, 1997.
- 37 **Bergus GR**, Cantor SB, Ebell MH, *et al.* A glossary of medical decision making terms. *Med Decis Making* 1995;**22**:385–93.
- 38 **Coiera E**. Glossary. In: *A guide to medical informatics, the internet and telemedicine*. New York: Chapman and Hall, 1997:339–50.
- 39 **Hammond E**. Glossary for healthcare standards. <http://dmi-www.mc.duke.edu/dukemi/acronyms.htm>, 1995.