

Article

Codes, classifications, terminologies and nomenclatures: definition, development and application in practice

A theme of the European Federation for Medical Informatics Primary Care Informatics Working Group (EFMI PCI WG)

Simon de Lusignan

EFMI PCI WG Chair; Senior Lecturer, Primary Care Informatics, Department of Community Health Sciences, St George's Hospital Medical School, London, UK

ABSTRACT

The Primary Care Informatics Working Group of EFMI is working to help develop the core theory of primary care informatics (PCI). Codes, classifications, terminologies and nomenclatures form an important part of the science of PCI, as they allow clinical information to be readily stored and processed in information systems.

This article provides definitions and a history of the International Classification for Primary Care

(ICPC), and of the Read code and the Systematized Nomenclature for Medicine (SNOMED).

The Working Group wishes to encourage shared definitions and an understanding of the practical application of structured data to improve quality in clinical practice.

Keywords: computers, family practice, feedback, primary health care, vocabulary controlled

Background

The Primary Care Informatics Working Group of the European Federation for Medical Informatics (EFMI PCI WG) sees primary care as unique; consequently its data, information and knowledge needs are different from those of other specialties. As a consequence we see primary care informatics as a unique subspecialty of health informatics.^{1,2} The group is working to share understanding in a number of key areas that have been previously set out in this journal and more recently updated on the EFMI website.^{3,4} One of these areas is ensuring that computer data quality reflects the standard of clinical care. Relevant to this theme are the following issues:

- how coded data can be used to improve the quality of chronic disease management

- the history and development of clinical coding systems (the term 'coding' is used loosely to include classifications, terminologies and nomenclatures)
- how to give feedback so that it is most effective in raising data quality.

The EFMI PCI WG is supporting workshops on these three themes. The first is planned for the EFMI Special Topic Conference (STC), which will be held in Athens in March.⁵ The second and third are due to take place at Medical Informatics Europe (MIE) 2005, in Geneva in August.⁶ This article provides definitions of the language likely to be used when describing clinical coding systems, and a historical overview of the two coding streams in Europe.

Codes, classifications, terminologies and nomenclatures

Worldwide clinical computer systems record data in two ways. Firstly, they allow the recording of coded data; this is usually done by selecting from a picking list or entering data into some sort of form or template. Secondly, most clinical computer systems also allow primary care professionals to enter 'free text' or narrative.

Coding is needed because there are so many ways that a clinical concept can be represented. For example, a patient with coronary heart disease can be represented or implied by any of the following free text labels: 'diagnosis of myocardial infarction', 'raised cardiac enzymes', 'myocardial ischaemia', 'triple vessel coronary artery disease', 'three vessel coronary artery bypass grafting' and so on. The clinical coding of the problem title using any of these labels should relate to the underlying disease process in any patient suffering from coronary heart disease, so that they could all be identified for audit and other clinical purposes. As yet natural language processing (NLP) has not reached the point where free text can be automatically turned into coded data.^{7,8}

A code is a simple representation (or label) given to a concept that allows it to be processed within an information system. Classifications provide a method of ordering information within a defined area or domain. For example, all the circulatory diseases in the Read system start with the letter G, and all operations with the number 7. The World Health Organization's (WHO) International Classification of Disease (ICD) is perhaps the best known classification.⁹ A terminology should provide comprehensive labelling of all the concepts within a domain.

Nomenclatures, which are the most sophisticated of all, allow concepts to be combined to enable more complex concepts to be created. Whereas in a classification system coronary artery bypass grafting (in Read chapter 7) does not convey any meaning that the patient has ischaemic heart disease (in Read chapter G), the same is not true within a nomenclature. In a nomenclature, coronary artery bypass grafting with vein allograft would be the 'sort of operation' performed on patients with coronary artery disease. The difference between codes, classifications, terminologies and nomenclatures is set out in Table 1. This does not mean that nomenclatures are without problems. They are much larger than classifications and terminologies, and much more complex. It is said that 'appendectomy' could be represented in 17 ways, making interpretation of data more difficult as it is necessary to search all 17 ways in which the concept might be represented.¹⁰

The history of coding, classifications and coding systems used across Europe

Europe does not have a standard approach to coding and classification.¹¹ Most of Europe uses ICPC (International Classification for Primary Care – a system owned by WONCA, the World Organization of Family Doctors), with various modifications and additions to meet different countries' needs.¹² United Kingdom (UK) primary care mainly uses Read at present and is due to migrate to SNOMED-CT (Systematized Nomenclature for Medicine – Clinical Terms).¹³ An overview of the history of these systems is set out below.

History of ICPC

1976: ICHPPC – International Classification of Health Problems in Primary Care. This was a list of common disorders encountered in primary care derived from ICD-8 (eighth revision of the World Health Organization's International Classification of Diseases).^{14,15} It was the forerunner of ICPC released in 1987 (see below).

1984: RFEC – Reason for encounter classification.¹⁶

1985: IC-process-PC – International Classification of Process in Primary Care.¹⁷ These classification systems reflect the need to code more than just the disease or problem title. This system and RFEC were brought together with ICHPPC to form ICPC. It has been represented formulaically as: (ICHPPC+RFEC+IC-process-PC = ICPC).

1987: ICPC – International Classification for Primary Care. Since 1972 WONCA had been looking to develop an instrument to support research in general practice. In 1987 they released the first version of ICPC.

1993: ICD-10 – International Classification of Diseases and related health problems. ICPC is used alongside ICD-10 (the World Health Organization's International Classification of Diseases) across the rest of Europe. These classifications, ICPC and ICD-10, are distributed at very low cost, removing the financial barriers associated with the use of Read codes or Clinical Terms. However, ICPC and ICD-10 have limitations in that they are less comprehensive and therefore can represent fewer concepts accurately. Although satisfactory at recording diagnosis,

they are much less good at recording detailed observations, investigations or actions taken (e.g. with ICD-10 there are codes for alcoholism, but not the ability to record the number of units of alcohol consumed). Although released in 1993, many countries did not migrate from ICD-9 (International Classification of Diseases – ‘and related health problems’ was added for version 10), which is still used in many parts of Europe. In common with other coding systems its size has increased. ICD-10 is larger than ICD-9; it has around 8000 codes, 3000 more than ICD-9.

1997: ICPC-2 – ICPC Version 2.¹⁸ In 1997 WONCA published ICPC-2. Like the first version of ICPC it is a biaxial coding system. There are 17 body system-related chapters and seven components covering patient-orientated aspects of primary care: diagnosis, reason for encounter, etc. ICPC-2 has been released in over 20 languages. It was extended to ICPC-Plus in 1998 in Australia; additional terms allow more detailed meaning to be provided than with the standard ICPC-2 release.¹⁹

History of Read and SNOMED

1983: Read codes version 1 (4-byte set). In 1983 Dr James Read’s codes were released, and these have gone on to become the UK national standard. The initial coding systems were about compactness, as early computer systems had so little memory. Later on it became more important that they were comprehensive, and version 1 has about 30 000 terms.

1988: Read codes version 2 (5-byte set). These earlier versions of Read codes (versions 1 and 2) are hierarchical, like a family tree. The version 2 code set was recommended for use by the Joint Computing Group of the British Medical Association, Royal College of General Practitioners and Primary Health Care Specialist Group. The 5-byte code set offers around 100 000 terms, and is in use in the majority of primary care computer systems in use in general practice in the UK.²⁰

1994: Read version 3, Clinical Terms. In 1994, a concept-based coding system was developed (Read 3), also known as ‘Clinical Terms version 3’ (CTv3).²¹ The intention was to develop a terminology that could include specialist practice as well as general practice. It has over 200 000 terms. It is used in a small minority of general practices in the UK. This will not

continue to be developed in the UK, but instead CTv3 has been merged with an American coding system called SNOMED (Systematized Nomenclature for Medicine). The new combined version is to be known as SNOMED-CT (CT for clinical terms).

1999: SNOMED-CT – Systematized Nomenclature for Medicine – Clinical Terms. In 1999, the UK Health Minister and the American College of Pathologists announced a joint venture to develop SNOMED-CT by late 2001. This was to be a combination of the SNOMED-RT (Reference Terminology) developed in the United States and Clinical Terms version 3 developed in the UK.^{22,23} The National Health Service (NHS) is due to migrate to SNOMED-CT as part of the implementation of the National Programme for Information Technology (NPfIT).^{13,24}

Workshop details

Despite the diversity of the coding systems described above, there is a great deal of commonality between those from different countries working in primary care. The issues for the Working Group are:

- 1 how to get primary care clinicians actually recording coded clinical data and using it as a tool to improve quality of care
- 2 how to promulgate understanding of the different coding systems and how they are likely to develop
- 3 what forms of feedback are known to be effective, and what theoretical models can be applied to achieve maximum change.

As mentioned previously, workshop 1 is planned for the EFMI STC in Athens, 19–20 March 2005. Workshops 2 and 3 are planned to take place at MIE 2005 in Geneva in August. Workshop 2 is being organised by Marc Jamouille and will include presentations by speakers from across Europe.

Please refer to the EFMI PCI WG website and the conference websites for details of these and other European activities for primary care informaticians.⁴⁻⁶ If you are interested in joining the EFMI PCI WG, please contact the group secretary (Neil Dhoul, email: adhoul@sghms.ac.uk) who will send you an application form (there is no charge to those from EFMI-affiliated countries – most of the European Union and a few others).

Table 1 Definitions of code, classification, terminology and nomenclature

	Definition	Examples
Code	A representation applied to a term so that it can be more readily processed.	Read <u>code</u> for asthma is H33
Classification	Arrangements of all elements of a domain, into groups according to established criteria.	International <u>Classification of Disease (ICD)</u>
Terminology	Language labels attached to a concept – all terms of a professional domain.	Read Clinical <u>Terms</u> version 3, CTv3
Nomenclature	Naming things: in a nomenclature, codes are assigned to medical concepts, and medical concepts can be combined according to specific rules to form more complex concepts. This leads to a large number of possible code combinations.	Systematized <u>Nomenclature of Medicine (SNOMED)</u>

REFERENCES

- Sullivan F. What is health informatics? *Journal of Health Service Research and Policy* 2001;6:251–4.
- de Lusignan S. What is primary care informatics? *Journal of the American Medical Informatics Association* 2003; 10:304–9.
- de Lusignan S. Introduction to the Primary Care Informatics Working Group of the European Federation for Medical Informatics (EFMI). *Informatics in Primary Care* 2003;11:175–6.
- European Federation for Medical Informatics Primary Care Informatics Working Group (EFMI PCI WG): www.efmi.org/efmi/wg.asp?page=groups2&wgid=6
- European Federation for Medical Informatics. *Special Topic Conference 2005: Health and Medical Informatics – Educational Aspects*. ghia.nurs.uoa.gr/efmi%2Ddstc2005/
- European Federation for Medical Informatics. *Medical Informatics Europe 2005 (MIE2005)*. www.mie2005.net/
- Liu H, Lussier YA and Friedman C. Disambiguating ambiguous biomedical terms in biomedical narrative text: an unsupervised method. *Journal of Biomedical Informatics* 2001;34:249–61.
- Lussier YA, Shagina L and Friedman C. Automating SNOMED coding using medical language understanding: a feasibility study. *Proceedings of the 2001 AMIA Annual Fall Symposium*. Philadelphia: Hanley and Belfus, 2001, pp.418–22.
- University of North Carolina. *International Classification of Disease*. www.cpc.unc.edu/dataarch/nhealth/icd.html
- Chute CG. Clinical classifications and terminology. *Journal of the American Medical Informatics Association* 2000;7:298–303.
- de Lusignan S, Mimmagh C, Kennedy J, Zeimet M, Bommezijn H and Bryant J. A survey to identify the clinical coding and classification systems currently in use across Europe. *Medinfo 2001 Proceedings*. 2001;10(Pt 1): 86–9.
- ICPC-2. *International Classification of Primary Care: an introduction*. www.ulb.ac.be/esp/wicc/icpc2.html
- NHS Information Authority. *SNOMED-CT (Systematized Nomenclature for Medicine – Clinical Terms)*. www.nhsia.nhs.uk/snomed/
- Classification Committee of WONCA. *An International Classification of Health Problems in Primary Care (ICHPPC)*. Occasional paper 1. London: The Royal College of General Practitioners, 1976.
- World Health Organization. *Manual of the International Statistical Classification of Diseases Injuries and Causes of Death* (8th revision). Geneva: WHO, 1967.
- Lamberts H, Meads S and Wood M. Classification of reasons why persons seek primary care: pilot study of a new system. *Public Health Report* 1984;99:597–605.
- Classification Committee of WONCA. *International Classification of Process in Primary Care (IC-Process-PC)*. Oxford: Oxford University Press, 1986.
- WONCA (World Organization of National Colleges and Academies of Family Physicians). *International Classification for Primary Care (ICPC-2)*. www.ulb.ac.be/esp/wicc/icpc2.html
- Britt H and Miller G. ICPC-Plus. An extended version of the International Classification for Primary Care for computerised clinician systems. *Proceedings of the Annual Conference of the Primary Health Care Specialist Group* 1996. Worcester: British Computer Society, 1996, pp.107–12.
- Robinson D, Wanger K and Price C. The Clinical Terms and ICPC: identifying equivalence and enabling compatibility for 'non-medical' terms. *Proceedings of the Annual Conference of the Primary Health Care Specialist Group* 1998. Worcester: British Computer Society, 1998. www.phcsg.ncl.ac.uk/conferences/cambridge1998/robinson.htm

- 21 O'Neil M, Payne C and Read JD. Read codes version 3 – a user led terminology. *Methods of Information in Medicine* 1995;34:187–92.
- 22 Stearns MQ, Price C, Spackman KA and Wang AY. SNOMED clinical terms: overview of the development process and project status. In: Bakken S (ed). *Proceedings of the 2001 AMIA Annual Fall Symposium*. Philadelphia: Hanley and Belfus, 2001, pp.662–6.
- 23 NHS Information Authority. *Global Standard for Healthcare Terminology*. Birmingham: NHSIA, 1999. www.nhsia.nhs.uk/snomed/pages/publications.asp
- 24 National Programme for Information Technology. www.npfit.nhs.uk

CONFLICTS OF INTEREST

None.

ADDRESS FOR CORRESPONDENCE

Simon de Lusignan
EFMI PCI WG Chair
Senior Lecturer, Primary Care Informatics
Department of Community Health Sciences
St George's Hospital Medical School
London SW17 0RE
UK
Tel: +44 (0)20 8725 5661
Fax: +44 (0)20 8767 7697
Email: slusigna@sghms.ac.uk
Web: www.efmi.org/efmi/wg.asp?page=groups2&wgid=6

Accepted January 2005