A framework for evaluating health classifications

Michelle Bramley

Abstract

Evaluation is important to evidence-based policy and practice in Health Information Management. Health classifications are important components of information systems and should be evaluated to determine their suitability for the task required. This paper provides a framework for evaluating health classifications that are used for statistical and reporting purposes. The framework revises and updates the fundamental principles that make health classifications effective. It also draws on other frameworks, where relevant, to reflect the influence that informatics has had on nosology. Principles are illustrated with examples, topical issues associated with some principles are discussed, and examples of evaluation in practice are provided.

Key words: *Evaluation framework; health classifications; nosology; health information management*

Introduction

Health information is the foundation for the policy, strategic and operational levels of the healthcare industry. Managing and translating health information efficiently is clearly important, and data standards and classifications are mechanisms through which this can be achieved (Gardner 2003; Information and Communications Technology Standards Committee 2004, pp.32–35; National Health Information Management Group 2002; Williams & Rowlands 2003).

Health classifications are systems that categorise terms used in healthcare and order them in a logical and methodical way (Hoffman & Chamie 1999). They are primarily used for statistical and reporting purposes, but may be used for other reasons, such as health services planning and funding. The most widely used statistical classification in healthcare is the *International Statistical Classification of Diseases and Related Health Problems (ICD)*, now in its 10th revision (World Health Organization 1992a). The ICD is used internationally to classify morbidity and mortality data. The ICD-10-AM, an Australian modification of the ICD (National Centre for Classification in Health 2004), is used nationally and internationally to classify diseases and procedures in in-patient settings. Another statistical classification used in Australia to classify general practice data for the BEACH©[1] data collection is the *International Classification of Primary Care 2nd edition–Plus* (ICPC-2 PLUS) (Family Medicine Research Centre 1998). Strictly speaking, the ICPC-2 PLUS is not a classification, but an electronic interface terminology.[2] Each term is uniquely identified and then classified to a category of a statistical classification, the *International Classification of Primary Care, 2nd edition* (WONCA 1998).

Not every domain in the Australian healthcare sector classifies its work, and nursing is a good example. There is no classification used on a systematic basis in Australia to identify what nurses do, why they do it and how well they do it (Hovenga 2002). While classifications do exist to capture this data, they need to be evaluated to determine how useful they are in an Australian setting (Hovenga 2002). In fact, all classifications should be evaluated before implementation to determine their suitability for the task required and the domain they are intended to serve.

This paper provides a framework for evaluating health classifications that are used for statistical and reporting purposes. The framework revises and updates the fundamental principles that make health classifications effective, as described in two seminal nosological papers (Hoffman & Chamie 1999; Price 1982). It also draws on other frameworks, where relevant, to reflect the influence that informatics has had on nosology. The principles are illustrated with examples, using the language of nosology to enhance understanding of the terms. Topical issues associated with some principles are discussed, and examples of evaluation in practice are provided. The examples and discussion have been shaped from my previous work for the National Centre for Classification in Health (NCCH), the Australian centre of expertise in health classification theory. The centre creates and maintains classifications and standards that uniformly describe health concepts such as diseases, injuries, contextual factors and clinical interventions.

Health classifications versus clinical terminologies

Health classifications differ from clinical terminologies in purpose, structure and output. Even though both manage clinical information, they manage it in different ways. Classifications are useful tools for aggregated statistical data analysis at all management levels, whereas terminologies are useful for operational decision making (for example, concept re-use in clinical decision support, point-of-care analysis) and semantically interoperable messaging. The once clear distinction between classifications and terminologies is becoming 'blurred' by the advent of electronic health records. The NCCH now frames health classifications as 'aggregate terminologies' — a 'functional subtype' of clinical terminologies (Scott 2003). This framework, then, is also relevant to aggregate terminologies that are used for statistical and reporting purposes because

[1] The BEACH© project collects data and reports information about general practice patients — their problems and treatments. It is conducted by the General Practice Statistics and Classification Unit, a collaborating unit of the Family Medical Research Centre, The University of Sydney, and the Australian Institute of Health and Welfare <http://www.fmrc.org.au/beach.htm>.

[2] See Family Medicine Research Centre – ICPC-2: Purpose. <http://www.fmrc.org.au/icpc2/purpose-d.htm>

they are conceptually equivalent to health classifications.

Before outlining the taxonomic principles of health classifications, it may be useful to define the key characteristics that distinguish health classifications from clinical terminologies: mono-hierarchies, comprehensiveness, aggregation, and meaningful coding systems.

Mono-hierarchies

Classifications are mono-hierarchical, that is, a concept appears at only one place in the hierarchy. Terminologies, particularly concept-oriented terminologies, tend to be multi-hierarchical, meaning a concept can appear at several places in the hierarchy.

Example: mono-hierarchies

Influenza illustrates this concept. Influenza is an infectious disease. Influenza is also a disease of the respiratory system. In a terminology, influenza can appear in two (or more) places in the hierarchy to reflect the different parent–child relationships. In a classification, it can appear in only one place in the hierarchy.
Placing a concept in only one place in the hierarchy is necessary for statistical analysis of data. Rogers (2003) explains why:
'Without the single parent rule, any statistical analysis of data could end up counting individual patients [in this example, the statistical unit] several times. If you are trying to pigeon-hole patients for the purposes of statistical aggregation, you have to be sure that each patient ends up in only one pigeon hole and you may need complex rules for deciding which pigeon hole a patient ultimately goes into if there is more than one reasonable choice'.

Comprehensiveness

Classifications are comprehensive (or exhaustive) and this feature is enabled through the use of 'other' and 'unspecified' (or residual) categories.[3] All concepts within a domain can be classified at any one point in time. Terminologies (generally) do not have residual categories.

Example: comprehensiveness

A general practitioner (GP) sees a patient and records the patient's signs and symptoms in the clinical record. Using ICPC-2, the GP can assign a code to every documented symptom and sign. Some concepts will be assigned to specific codes. The remaining concepts will be assigned to residual categories. Thus, all relevant information relating to the patient's illness can be coded at that point in time.

Aggregation

Classifications tend to aggregate several unique concepts under one code. Only certain concepts are separately identified by their own unique code; these concepts are important to distinguish for public health purposes, research or statistical reporting. Terminologies discriminate and allocate a unique code to each and every unique concept.

The residual categories in classifications allow the aggregation of several distinct concepts under one code. Aggregation is necessary and useful for the reporting of data at all management levels. At the operational level, however, the ability to retrieve specific *terms* from an information system that stores data by *codes* tends to be lost when concepts are classified to a residual category.

Example: aggregation

In ICD-10-AM, more than thirty unique concepts are indexed to N83.8 *Other non-inflammatory disorders of the ovary, fallopian tube and broad ligament.* 'Broad ligament laceration syndrome' is one of these concepts. When a clinical researcher asks you to identify the number of cases of broad ligament laceration syndrome treated in the facility over the past three years, the original clinical records must be retrieved for each instance of N83.8 to determine which records contain this term and thus meet the search criterion.

Terminologies work more efficiently than classifications in terms of data retrieval at the operational level because they allocate a unique identifier to each concept. It is primarily for this reason that terminologies are more suitable for decision support systems and point-of-care analysis. Whether terminologies are useful at the reporting level is open to debate. Bowman (2005) states that terminologies are 'inadequate' for aggregation 'because of their immense size, considerable granularity, complex hierarchies, and lack of reporting rules' (p.1). Roberts et al. (2004) note that 'In reporting, there are as yet few if any predetermined groupings to use for national or international data extraction and interpretation' (p.30). It seems then, that it is yet to be demonstrated whether terminologies can function as classifications in respect of aggregation and reporting.

Meaningful coding systems

Classifications apply coding systems[4] which generally reflect the hierarchical structure of the classification. In this sense, the codes have meaning and are best understood in the context of their relationship to other codes in the classification. In terminologies, the coding system does not (generally) reflect the hierarchical structure. In this sense, the codes are meaningless.

Example: meaningful coding systems

In ICD-10-AM, the codes are an essential part of the system and aid in understanding the structure of the classification. Users can get a sense of where they are in the classification from the coding system.

- O61.0 Failed medical induction of labour

O is used in the obstetrics chapter to identify conditions affecting pregnancy or occurring in childbirth or the puerperal period. The codes in block O60–O75 describe complications of labour and delivery.

3 Residual categories are further explained in the following section — Aggregation.

4 Codes are attached to concepts to help manage the data electronically. Codes can be either alphabetical or numerical, or alphanumeric.

The *Systematized Nomenclature of Medicine — Clinical Terms* (SNOMED-CT) (College of American Pathologists n.d.) is a concept-oriented reference terminology. In SNOMED-CT, concepts and terms and their relationship to each other are the primary focus. Users rarely see the codes; they are simply a management tool to enable computerised identification and retrieval. In this example, the concept identifiers (Conceptid) are the codes; they are randomly assigned to concepts and convey no meaning.

- Influenza (Conceptid 6142004)
- Influenzal pneumonia (Conceptid 39932903)
- Influenzal laryngitis (Conceptid 19463902)
- Influenzal pharyngitis (Conceptid 19452905)

(Source: Scott 2004)

Meaningful coding systems arose out of necessity to keep the hierarchical structure and format simple for users in a predominantly manual environment (Roberts et al. 2004). In the evaluation framework, some examples demonstrate the tensions that exist in maintaining a classification's structure and its clinical relevance. An electronic environment may offer solutions to some of the problems that arise with meaningful coding systems (Roberts & Chalmers 2002).

Evaluation framework

Health classifications can be evaluated by the characteristics that define a 'good' classification (Hoffman & Chamie 1999). These taxonomic principles are explained below and are themselves grouped into administrative, structural, content and usability principles.

1. Administrative principles

1.1 Purpose and scope (coverage)

A health classification must have its objectives, purpose and scope (or coverage) clearly stated so that its relevance to the domain it serves, or to other domains, can be measured (Hoffman & Chamie 1999).

Example 1.1a

A classification of chiropractic interventions may not be suitable for use by physiotherapists who specialise in manual therapy (manipulation and mobilisation). There may be some common interventions, but generally, physiotherapy interventions are different from chiropractic interventions, because of theoretical and philosophical differences between the two disciplines (Canadian Physiotherapy Association 1994).

Example 1.1b

Procedure classifications generally have a broad scope because they need to encompass all types of procedures performed in healthcare: diagnostic, therapeutic and preventive interventions; and invasive, non-invasive and cognitive interventions. They also need to be applicable to all clinicians, across all healthcare settings (Innes et al. 1997; National Committee on Vital and Health Statistics Subcommittee on Medical Classification Systems 1993).

1.2 Custodianship

The custodian is the organisation or body responsible for the development and maintenance of a health classification. They should be readily identified and their responsibilities clearly outlined (Hoffman & Chamie 1999).

1.3 Maintenance/updating

1.3.1 Maintenance plans

Health classifications must remain credible and relevant to users and so must be maintained and updated over time. The custodian must have a plan for regular updating and maintenance (Hoffman & Chamie 1999; Price 1982), which clearly documents the criteria against which a submission for change can be made. The plan should be well publicised and allow for all users and producers of statistical data to contribute to the process within an appropriate time frame (Hoffman & Chamie 1999).

Example 1.3.1

The NCCH updates the ICD-10-AM every 2 years. There is a public submission process advertised on its website <http://www3.fhs.usyd.edu.au/ncch/4.7.1.htm> which allows all stakeholders to participate in the update process. Similarly, the Australian Government Department of Health and Ageing updates the *Australian Refined Diagnosis Related Groups* (AR-DRGs) classification in line with updates to the ICD-10-AM and has a public submission process <http://www.health.gov.au/internet/wcms/publishing.nsf/Content/Casemix-1>.
The ICD has a ten year update cycle, though it is now more than 10 years since ICD-10 was introduced (1992). Plans to introduce ICD-11 have been extended to 2011 (Ustun 2004), primarily because the World Health Organization now has a mechanism through which the ICD is regularly updated — the WHO Update Reference Committee (National Centre for Classification in Health 2005).

Because stability is crucial to statistical analysis, the impact of any updates needs to be considered by all stakeholders (Hoffman & Chamie 1999). In Australia, there are diverging views on what constitutes 'regular' updating. Clinical coders, and the bodies responsible for maintaining classifications, prefer updates once a year or biennially so that the classification remains current, particularly in relation to surgical procedures. Researchers, policy analysts and those who maintain health information systems prefer a more stable classification with less frequent updates and suggest a timeframe of 3 to 5 years.

1.3.2 Concept permanence

Classification follows clinical research and discovery. As more is learnt about the nature of diseases and the efficacy of interventions, the better able we are to describe, define (or redefine), and classify them appropriately. Thus, the classification of clinical concepts can change over time.

Concept permanence relates to the maintenance of the meaning of concepts. The meaning of a concept in a health classification should be permanent and never deleted (International Organization for Standardization 2000). The terms representing the concept may change (even its preferred term or category descriptor), however, the concept's meaning should remain the same. Concepts can be retired or invalidated, but not deleted.

The principle of concept permanence does not refer purely to concepts, but to the codes as well. When a concept is denoted as inactive or is superseded, the code linked to the concept should not be re-used. The meaning of a code should also be permanent because of its link to a concept.

Example 1.3.2

This example nicely illustrates the tensions that can arise between adhering to the principles of classification development and maintaining a clinically relevant classification.
The changes to the rubrics classifying diabetes mellitus in ICD-10-AM (E10-E14) in the second and third editions were significant and timely. The NCCH replaced an outdated classification with a new classification that reflects the changes to clinical knowledge. Importantly, the changes made were based on the World Health Organization's (WHO) recommendations for reclassifying diabetes mellitus (World Health Organization 1999).
In making the changes it was necessary to delete existing codes and concepts and re-introduce the same codes with different descriptors (that is, overriding the concept permanence principle and introducing concept changes). The NCCH managed this change effectively by clearly identifying all concept changes in their mappings between editions. They ensured that the ability to compare data over time was not compromised by communicating these changes to stakeholders. Stakeholders could then interpret these changes with relevance to their work.
This example also illustrates a flaw in the structure of the ICD-10. The alphanumerical coding system reflects the hierarchy and, in most instances, does not allow for major restructuring of sections of the classification between revisions. Another problem is the long time lapse between revisions of the classification (currently more than 10 years).

1.3.3 Tracking changes over time

Mechanisms for tracking changes to health classifications over time should exist (Hoffman & Chamie 1999; Price 1982). These include version control and mapping.

Each update, revision or edition of a health classification should be clearly described and differentiated from previous versions. Documentation explaining the changes made, and listing the time frames in which they were made, should be readily available to all stakeholders (Hoffman & Chamie 1999).

Example 1.3.3a

The NCCH has developed the ICD-10-AM Chronicle, an electronic reference tool, with the aim of enhancing understanding about what changes were made, when they were made and why (National Centre for Classification in Health, n.d.).

With every update, the custodian should perform (or validate) mapping to help researchers compare data over time.

Mapping is the process of identifying equivalent concepts or terms in related health classifications or terminologies, or in different ones. It is performed to provide researchers, data analysts, governments and other authoritative bodies with information about the comparability between different health classifications or to demonstrate the relationships between changes introduced in updated versions or editions of a single health classification. The ability to compare data over time (tracking data) between different versions or editions of a classification is essential for epidemiologists and other researchers conducting longitudinal studies. It is also helpful to compare different health classifications to enable the translation of data from one health classification to another.

Mapping is a complex activity. Concepts infrequently match on a one-to-one basis. More often, there is some loss of meaning. Whether this loss of meaning is important will depend on the purpose of the mapping. For example, a mapping performed for the purpose of identifying the changes made to a parent classification and its country-specific modification (for example, ICPC-2 PLUS to ICPC-2) may require less precision than mapping performed between a terminology and a classification for the purpose of funding healthcare services (Bowman 2005). The effectiveness of any mapping must therefore be evaluated and will be dependent on a number of variables. Such variables include: reasons why the mapping was performed, the expertise of the mapper, criteria or rating scale applied (the degree to which the health classifications are comparable), and electronic tools used to assist the mapping process.

Example 1.3.3b

The *Unified Medical Language System* (UMLS) is a 'terminology mapping, translation and maintenance system on which broad-based information retrieval and domain modelling can be based' (Campbell, Oliver and Shortliffe 1998: p.15). In essence, it is a terminology interoperability effort which helps health professionals and researchers retrieve and use biomedical information from different sources (e.g. MeSH headings).
ICPC-2 has been incorporated into the UMLS to provide a mechanism through which ICPC-2 can be related to other classifications and terminologies included in the UMLS (Family Medicine Research Centre 2005).

2. Structural principles

2.1 Hierarchical organisation/theoretical framework

Health classifications should have a clinically logical, hierarchical organisation built upon a theoretical framework. The framework should clearly and simply define the principles of this organisation to enhance understanding by all users (Hoffman & Chamie 1999). The hierarchical organisation should facilitate data retrieval at different levels of specificity by enabling aggregation of data from subcategories to categories ('roll up/roll down') (Price 1982).

Example 2.1

The organising principles for the *Canadian Classification of Health Interventions'* (CCI) theoretical framework can be seen from analysing the codes themselves. The CCI is an electronic, multiaxial classification (Moskal 2002). The hierarchy is evident from the codes, as is the increasing level of specificity as you roll down through the list of codes. Codes are 'built' by combining an element from each relevant axis, and consequently the codes have meaning. In this example, '1' represents the therapeutic interventions axis of the classification, 'GE' represents the anatomical site axis, '89' represents the type of intervention, and 'LA' represents the approach. The qualifiers (XX and A,F,G,Q) provide further specificity about the graft or flap applied.

- 1.GE.89.^^ Excision total, larynx NEC
- 1.GE.89.LA using open approach (e.g. apron flap incision)
- 1.GE.89.LA-XX-A using open approach and autograft (e.g. full thickness skin graft)
- 1.GE.89.LA-XX-F using open approach and free distant flap
- 1.GE.89.LA-XX-G using open approach and pedicled distant flap (e.g. myocutaneous flap)
- 1.GE.89.LA-XX-Q using open approach and combined sources of tissue.

(Moskal 2002).

The organising principles for the ICD's framework may not be as obvious to users. The ICD also has variable or multiple axes; some chapters are based on aetiology (infectious conditions, neoplasms), some are based on status (pregnancy) and life stage (perinatal conditions), while others are based on topography (anatomical site). The problem with these chapters is that they are not mutually exclusive. Taxonomic concessions have been made in the framework to suit the ICD's theoretical epidemiological purpose and to ensure its practicality for users (World Health Organization 1993: p. 12). It is essential that users understand the implications of the structure when coding or interpreting statistics because the chapters based on aetiology, status or life stage take precedence over the anatomical site chapters (World Health Organization 1993: p. 13).

2.2 Expansion

'The only good classification is a living classification' (Bowker & Star 1999: p. 326). Health classifications need to be responsive to changes in clinical practice and new technology if they are to maintain their relevance. They also need to maintain their uniformity with related classifications (Roberts and Chalmers 2002). The structure of a health classification should be flexible to allow for expansion (Price 1982). The addition of new concepts into the hierarchy should not disrupt systematic code structures.

Example 2.2

The *International Classification of Functioning, Disability and Health* (ICF) (World Health Organization 2001) has linked a meaningful, alphanumeric coding system to concepts. There are at least five digits between each category to allow for expansion. Within categories, there is capacity to expand to ten digits for first, third and fourth level items, and one hundred digits for second level items.

b Body functions (a key component of the ICF)
b2 Sensory functions and pain (first-level item – one digit)
b210 Seeing functions (second-level item – two digits)
b2102 Quality of vision (third-level item – one digit)
b21022 Contrast sensitivity (fourth-level item – one digit)
(World Health Organization 2001: p. 220)

In the previous example, for the criterion 'concept permanence', the problems with meaningful, numeric coding structures were highlighted. Over time, the ICF may encounter the same problems as the ICD with expansion to accommodate changes in clinical practice and theory.

2.3 Comprehensiveness

A health classification must be comprehensive (i.e., exhaustive) to support the domain it serves (Hoffman & Chamie 1999). 'Other' or 'unspecified' categories (residual categories) must be provided so that all possible concepts within a domain can be classified somewhere, at any given point in time. Classifications, in this sense, have sensitivity at the expense of specificity.

Example 2.3

Current Procedural Terminology (CPT) (American Medical Association 1999a) is a mandated national code set used in the United States for coding, reporting and reimbursement of medical procedures and services performed by physicians in all settings. Although it is a terminology, one of the features it shares with classifications is the residual codes which are used in each section to identify a service or procedure that is not listed in the current edition. In the surgery section, the following are examples of residual codes:

- 21499 Unlisted musculoskeletal procedure, head
- 55899 Unlisted procedure, male genital system
- 64999 Unlisted procedure, nervous system

In the US, users are required to submit a special report when these codes are applied (American Medical Association 1999a: p. 51). The report provides a detailed description about the procedure and justification for its use to enable authorities to determine an appropriate payment. Information provided in the report may also be used to inform future updates.

Residual categories are essential to the purpose of a health classification. Nonetheless, they should be kept to a minimum because of the inability to retrieve specific data from these categories without reference to the original source. Management of residual categories through systematic analysis of the data should be performed regularly to determine whether the data demonstrate the need for a concept to have its own unique identifier. Analysis can also determine if concepts are being correctly classified to residual categories, which in itself implies that: (a) there may be a need to improve the index, or (b) there may be a need to improve clinical documentation.

3. Content principles

3.1 Mutual exclusivity

Categories or subcategories must be mutually exclusive (Hoffman & Chamie 1999). There must be only one code for any given concept with adequate indexing and guidelines to denote the boundaries. There cannot be two (or more) different codes for the same

concept. If there are, then the classification is said to have redundancy (needless repetition). The main problems with redundancy are data retrieval and statistical counts, particularly if data managers are unaware of the redundancy.

Example 3.1

Recall Roger's (2003) explanation for why statistical classifications need to be mono-hierarchical and apply that here: each statistical unit should only be counted once. Mono-hierarchies are one mechanism used by classification developers to facilitate mutual exclusivity. Terminologies, because of their multi-hierarchical structures, have problems with mutual exclusivity. SNOMED is a multiaxial system; codes can be built by combining elements from each relevant axis. In an earlier version, SNOMED III, Evans et al. (1994) found that appendicitis could be expressed in seventeen different ways. The following are three examples:

- D5-46210 Acute appendicitis, NOS
- D5-46100 Appendicitis, NOS
  G-A231 Acute
- M-41000 Acute inflammation, NOS
  G-C006 In
  T-59200 Appendix, NOS

In SNOMED III there were limited mechanisms for determining the equivalence of concepts built in this way, and hence the problems with data retrieval. Data managers needed to 'anticipate virtually all permutations and combinations of expression to ensure complete retrieval' (Chute 2000: p. 300). These problems with equivalency and redundancy are being addressed in SNOMED-CT through automated means, such as description logics (Spackman & Campbell 1998), and manual review and analysis (Sable, Nash & Wang 2001).

3.2 Unique, unambiguous and clearly expressed descriptors

To facilitate communication and understanding between all users and producers of statistical data, each category or subcategory descriptor should be unique (have one only meaning), unambiguous and clearly (not vaguely) expressed to convey meaning (International Organization for Standardization 2000; Price 1982). The meaning of each category or subcategory should be understood from its descriptor alone. Meaning should not be inherent in the concept's relationship within the hierarchy (International Organization for Standardization 2000).

Example 3.2a

The concept 'lobe' is *vague* and loses meaning when taken out of context. Does it mean 'lobe of the brain' or 'lobe of the lung'? The term 'von Recklinghausen's disease' describes two unique concepts; neurofibromatosis and osteitis fibrosa cystica generalisata. When used alone and out of context, 'von Recklinghausen's disease' is *ambiguous*. To avoid vagueness and ambiguity in these cases, the terms should be specified in the classification as 'lobe of the lung', 'lobe of the brain', 'von Recklinghausen's disease (neurofibromatosis)' and 'von Recklinghuasen's disease (osteitis fibrosa cystica generalisata)'.

Example 3.2b

The ICD-10 has a number of codes with the same descriptor. One evaluative study noted this aspect 'may cause some confusion' for users (General Practice Coding Jury 2000, p.16), because at first glance it appears that some of the categories are not mutually exclusive and the classification has redundancy (needless repetition). Two examples are: C18.6 *Descending colon* and D12.4 *Descending colon*. Both codes depict the same anatomical site, and the codes lose meaning when viewed out of context (for instance in a morbidity data report listing the frequencies of different types of neoplasms treated in a facility). Their true meaning is discovered only by looking at their place in the hierarchy of the classification. C18.6 identifies a malignant neoplasm, and D12.4 identifies a benign neoplasm, of the descending colon.
To overcome this problem in the ICD-10-AM, the NCCH has included the textual hierarchical information in the full descriptor that appears in its ASCII (American Standard Code for Information Interchange) code list which is distributed to software developers and authorities for use in software and information systems.

Example 3.2b is another illustration of the tension between practicality and principles. International classifications serve as a lingua franca; in the case of the ICD it is a common language for morbidity and mortality data. In interpreting this principle, be aware that in order to satisfy an international community and be a practical epidemiological tool, there has to be some ambiguity in some of an international classification's descriptors and definitions (Bowker & Star 1999).

3.3 Relevant and standardised terminology used in descriptors

Standardised use of language that is accepted and in common usage, and relevant to the domain and scope should be used to describe each category or subcategory (Price 1982). In the case of health classifications, descriptors should be clinically relevant.

Example 3.3a

Semantics are important. In healthcare, a patient's life may depend on certainty of meaning in communication between clinicians. O'Rourke (1997) nicely illustrates the importance of meaning when he discusses issues that patients have with some of the descriptors that practitioners use to describe certain conditions. He cites hypertension as one example in which the descriptor does not accurately reflect the condition and hence causes false perceptions in the public's mind. It is generally believed that tension equates to stress. Therefore 'hypertension' becomes a stress-related condition, and so stress is the important risk factor, rather than high blood pressure. O'Rourke goes on to say that high blood pressure, a more precise descriptor, is slowly becoming the preferred term. Public education should result, if more clinicians use the preferred term. Nosologists need to be sensitive to the nuances of language and the changes that occur over time in the way language is used within a domain.

Descriptors for procedures or interventions should be setting and provider neutral; they should not reflect the clinician who performed the procedure or indicate where the procedure was performed (Innes et al. 1997; National Committee on Vital and Health Statistics Subcommittee on Medical Classification Systems 1993). One logic behind this principle is redundancy; such additional information is captured elsewhere in

the information system. Another is to enable comparability across different clinicians and sites.

Example 3.3b

In the first edition of ICD-10-AM, the procedure codes for allied health interventions were included as a separate chapter and identified the clinician who performed the procedure. For example:

- 95254-00 [2095] Education/counselling, speech pathology
- 95054-00 [2056] Education and information, social work

This was the first attempt to classify allied health interventions consistently and include them in a national procedure classification (Innes and Bramley 1997). Some guiding principles had to be overlooked because of the timeframes for consultation, publication and implementation.

In the second edition of ICD-10-AM, the codes were restructured on the basis of provider neutrality and integrated into the existing chapters of the classification, thus removing any duplication of interventions (Bramley and Innes 1998). The timeframe between publication of the first and second editions allowed the NCCH to work constructively with the allied health professions to produce a classification that followed guiding principles and met the needs of key stakeholders.

3.4 Semantic and conceptual scope of descriptors

Category or subcategory descriptors should have a meaning (semantic scope) that corresponds to the idea (conceptual scope) being classified. When the meaning does not correspond to the idea, there is a concept mismatch.

Example 3.4a

If a category descriptor for an object producing an injury is 'furniture' then all subcategories should be items of furniture. 'Carpet' should not be classified to this category, as it is not considered a piece of furniture. If carpet was classified to this category, then the conceptual scope of the category would need to be changed to something more representative, like 'furniture and floor coverings'.

When a category descriptor states 'cough', expectations are that all subcategories are instances of 'cough', such as 'chesty cough', 'dry cough' or 'chronic cough'. If 'does not cough' is included as a subcategory, then there is a concept mismatch. The category descriptor is essentially saying the patient has a cough, and statistically, the count reflects the numbers of patients who had a cough at a particular point in time. It is illogical, then, to include in those statistics a patient that does not have a cough.

The *International Classification of External Causes of Injury* (ICECI Coordination and Maintenance Group 2004) is a multiaxial classification that captures data about the circumstances of an injury (external cause, mechanism of injury, place of occurrence, et cetera). Injury surveillance activities and injury prevention programs aim to reduce death and illness. To be effective, these initiatives require a significant level of detail and the importance of that detail is reflected in the ICECI (available online at <http:www.iceci.org>).

As part of a quality review process in the development phase of the ICECI, a taxonomic review was conducted to test the technical qualities of the classification. Part of this work was conducted by the NCCH. A few minor irregularities relating to semantic scope were identified and rectified in version 1.1.

Descriptors for procedures or interventions should not include diagnostic information, where possible (Innes et al. 1997; National Committee on Vital and Health Statistics Subcommittee on Medical Classification Systems 1993). The logic behind this principle is, again, redundancy. Diagnostic information is captured in the disease codes, so there is no purpose to capturing it again in the procedure codes.

Example 3.4b

The Australian Classification of Health Interventions (ACHI) (National Centre for Classification in Health 2004) classifies surgical procedures and other clinical interventions. One principle in developing procedure classifications is to avoid including diagnostic information in the category or subcategory descriptors. The descriptor should correspond with the procedure being performed, rather than the disease being treated, for example 'transplantation of liver' or 'lobectomy of lung'. Having said that, there are instances where the diagnostic information is inextricable from the procedure, for example 'haemorrhoidectomy', or warranted, such as 'excision of pilonidal sinus'.

3.5 Unique categories for important concepts

Concepts that have particular importance within a domain should have their own unique category in a statistical classification (Hoffman & Chamie 1999). These concepts are important for public health purposes, research or statistical reporting and so should be easily distinguishable.

Who determines which concepts are uniquely distinguished from others? In developing and maintaining health classifications, many 'voices' are taken into account, not just those of the developers. The influential voices are not purely based on clinical research and discovery. Political reasoning can dominate, as can technological, social, religious, legal, moral and ethical reasoning (Bowker & Star 1999). Balancing the needs of all stakeholders means compromise. In any process involving consensus, it is the strongest voices that have the most influence and will be heard above other voices who will generally be silenced (Bowker & Star 1999).

Example 3.5

Consider that in 1977, homosexuality was classified as a mental disorder in ICD-9. The guidelines clearly stated that it was to be coded to this category 'whether or not it is considered as a mental disorder' (World Health Organization 1977: p.196). The fact that homosexuality was considered a pathology then was probably underpinned by strong political, legal, social, moral and religious voices. Opinion must have been divided though and hence the inclusion of the guidelines.

By 1992, the term is absent in the ICD-10 because of: the strongest voices being now a reflection of the changing social norms; the emerging influence of gay and lesbian advocacy groups; and empirical data showing little support for its pathological classification (Herek 2005?; Narrain and Chandran 2002?; Smith, Bartlett & King 2004; Stakelbeck & Frank 2003; Wu 1998).

4. Usability principles

4.1 Definitions and instructional notes

The categories or subcategories in a health classification must be well defined and supported by definitions and explanatory notes (Hoffman & Chamie 1999; Price 1982).

Example 4.1

The *Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition – Text Revision* (DSM) (American Psychiatric Association 2000) is more than a classification, as its name suggests. The manual is used extensively by mental health clinicians, primarily as a diagnostic tool, but also as a statistical, research and educational tool, and is applicable in all healthcare settings. ICD-9-CM is the statistical classification applied and there are mappings to ICD-10 codes.
The categories in DSM are differentiated on diagnostic criteria and other defining attributes of a disorder. They are extensively defined, and support the practical use of the manual.

4.2 Indexes/thesauri

Indexes or thesauri are essential to provide users with easy access to the desired term and its classification. They should contain all terms relevant to the domain the health classification serves, including synonyms (Hoffman & Chamie 1999).

Example 4.2

The ICPC-2 PLUS is an electronic interface terminology, with links to its parent classification, the ICPC-2. To describe it very simply, it is an electronic index of terms used in general practice. Terms are managed in a database which is routinely updated from user feedback. Each term is given a unique identifier which is then linked to its relevant category in ICPC-2.
The end product is a clever electronic solution to two problems: (1) it overcomes the lack of specificity in the ICPC-2 for clinical research purposes; and (2) it simplifies the work of GPs when classifying or retrieving morbidity data. A demonstrator model of the ICPC-2 PLUS can be viewed on the Family Medicine Research Centre's website:
<http://www.fmrc.org.au/icpc2plus/demonstrator.htm>.

Indexes are also essential tools for maintaining classifications. They aid in reducing redundancy and detecting misclassifications.

4.3 Guidelines/training materials

Guidelines or rules on how the classification is to be used should be available to all users. Training materials should also be provided, particularly when revisions or updates are introduced (Hoffman & Chamie 1999).

Example 4.3

Volume 5 of the ICD-10-AM is the *Australian Coding Standards*. The aim of the volume is to achieve consistency in coding through guidelines on the use of the classification. The standards are written with the assumption that users have had training in the application of the classification, are trained in the abstraction of clinical data from clinical records, and familiar with the ICD conventions (National Centre for Classification in Health 2004). The Australian Institute of Health and Welfare has developed an Australian user guide for the *International Classification of Functioning, Disability and Health* (ICF) (Australian Institute of Health and Welfare 2003). Practical advice is also included for those preparing to implement the ICF.

4.4 Suitability/acceptability/appropriateness

Health classifications should be well matched to the users' work processes and information flow to enable integration (Price 1982). In this sense, health classifications can be described in terms of their specificity. The level of detail should match that of the domain it serves (International Organization for Standardization 2000).

Example 4.4a

The ICD-10-AM works well in acute care settings because of its specificity. However, the five-volume classification in its manual format is said to be too detailed for use in general practice (Britt, Beaton & Miller 1995; Parnanen, Kumpusalo & Takala 2000).

Health classifications can also be described in terms of their granularity. The term is often confused with specificity, generally because granular classifications allow more specificity. In informatics, the term 'granularity' is defined as 'the degree of modularity of a system. More granularity implies more flexibility in customising a system because there are smaller increments (granules) from which to choose' (Answers.com 2005).

Example 4.4b

The *Canadian Classification of Interventions* (CCI) is more granular than the *Australian Classification of Health Interventions* (ACHI) because the CCI has a flexible structure; it allows codes to be built by taking an element from each axis (or module or table). The codes in the ACHI are predefined; elements cannot be added to them for greater specificity.

Health classifications should not be too complex or difficult to use, nor should they require more detailed information than is available to the user (Price 1982).

Example 4.4c

The NCCH has worked collaboratively with the AIHW National Injury Surveillance Unit to include a greater level of detail in the external causes chapter of the ICD-10-AM. Guidance for some changes has been drawn from the *International Classification of External Causes of Injury (ICECI)*. Their rationale is that this specificity is necessary to make the external cause codes more useful for injury prevention and control (Harrison 2001).
The effects of this increased specificity on usability need to be measured. It is uncertain whether this high level of detail is currently captured in in-patient records and is able to be coded. Schmertmann & Wil-

liamson (2001) reviewed NSW injury data from 1995-1999 and found that residual categories were overrepresented in the data set. They could not determine whether this was because of poor source information or inadequacies of the ICD's theoretical framework (p.103).
Before any more specificity is added to this chapter, research needs to be undertaken to determine whether there is enough detail in the source information to satisfy the specificity. Only then will we know if these changes have indeed led to 'some major gaps in [injury] information … being filled' (Harrison 2001).

4.5 Compatibility/comparability

A health classification should be compatible or comparable with other similar health classifications in use, both nationally and internationally.

Example 4.5

Recall that mapping is a mechanism used for maintenance of a health classification. Mapping also measures comparability between health classifications. Mappings between different health classifications can demonstrate the quality of the relationship through the number of one-to-one mappings, representing minimal loss of detail. However, the quality of the relationship will always be dependent on the purpose of the mapping.
ICPC-2 is mapped to the ICD-10 to demonstrate the compatibility, and facilitate linkage, between the two classifications (Lamberts & Wood 2002; Okkes et al. 2002; Wood et al. 1992). The mappings are useful for conversion of data from ICPC-2 to ICD-10 and vice versa (Britt, Beaton & Miller 1995), and thus provide a connection or common language between primary care services and acute care services.

4.6 Subsets

If a health classification serves a wide domain, then mechanisms should exist to enable the production of a subset of the classification for a specific sector within that domain, or perhaps for other domains that deem the classification suitable for their purposes.

Example 4.6

The *Nursing Interventions Classification* (Iowa Intervention Project 2000) covers all nursing domains but, because of its broad scope, may not be suitable for use by community nurses working in community-based healthcare settings. A subset that includes only the interventions appropriate to their domain may be less cumbersome and more useable than the larger classification.
ICD-10-AM is the mandated classification for use in Australian acute and community-based mental health services. In community health centres, there are limited numbers of trained coding staff, so it is more likely that clinicians will code and report the data. To support clinicians in this task, the NCCH developed the *ICD-10-AM Mental Health Manual* (National Centre for Classification in Health 2002), a subset of the ICD-10-AM, Third Edition, as a diagnostic and coding tool. The manual is similar in theory to the DSM, however, the diagnostic criteria and other defining attributes of a disorder are based on a number of WHO publications that are all subsets of the ICD-10 (World Health Organization 1992b; 1996a; 1996b; 1997). Other features distinguishing this manual from the DSM are the inclusion of mental health interventions, and diagnostic instruments to supplement the diagnostic guidelines.

4.7 Adaptable to an electronic environment

Health classifications should be adaptable to an electronic environment to broaden their usability (Hoffman & Chamie 1999). Database management of health classifications can streamline the production, maintenance, updating, mapping and exchange processes. Subsets can be more easily created from databases.

Practical examples of evaluating health classifications

Evaluation is performed to determine the 'merit, value or worth of things' (Scriven 1991: p. 1). It is a rigorous process that employs both quantitative and qualitative research methodologies and should be performed at all stages of an information system's life cycle: planning, development, implementation, and operation. Evaluation enables informed decisions to be made based on scientific facts and thus is harmonious with evidence-based policy and practice in Health Information Management (Leys 2003; Rigby 2001). Health classifications are essential components of information systems. The following evaluations are all good examples of evaluation in practice. Some build on a theoretical framework, while others apply components of it. They each have different purposes and objectives.

The Australian Institute of Health and Welfare (AIHW) evaluates health classifications for inclusion in the Australian Family of Health and Related Classifications. Their aim is to endorse national classifications for specific purposes within particular health settings. As the Australian WHO Collaborating Centre for the Family of Classifications, the AIHW has aligned this work with the principles of the WHO's family of international classifications and the United Nation's family of international economic and social classifications (Australian Institute of Health and Welfare 2002). The AIHW General Practice Statistics and Classification Unit recently lodged a submission for inclusion of the Anatomical, Therapeutic, Chemical classification system with Defined Daily Doses (ATC/DDD) and the relevant documents can be viewed online at:

<http://www.aihw.gov.au/committees/ctwg/submissions/submissions.cfm>.

When the decision was made to implement ICD-10 in Australia, the next task was to determine the most appropriate procedure classification to be used in conjunction with the ICD-10 (Bramley 1994). Four candidates were assessed for their taxonomic quality and usability:

- 3M Health Information Systems ICD Procedure Coding System (ICD-10-PCS) (USA);
- Office of Population Censuses and Surveys Classification of Surgical Operations and Procedures, fourth revision (OPCS-4) (UK);
- Commonwealth Medicare Benefits Schedule (MBS) (Australia); and

- Current Procedural Terminology, fourth edition (CPT) (USA).

A rating system for each criterion was created in order to make a fair comparison. CPT was the best performer, MBS was the worst. The result for MBS was not surprising because it is not a classification, and it was perhaps unfair to include it in the assessment. However, the Australian Government saw merit in developing an Australian procedure classification based on the MBS, and the rest is history (Innes et al. 1997). In hindsight, the decision to create an Australian procedure classification was the best choice. The OPCS-4 was, and remains, outdated, and its use in the United Kingdom is under review (National Health Service Information Authority 2004). The ICD-10-PCS has not yet been implemented in the United States, and the future of the CPT is not certain as debate goes on (and on) about the value of having only one procedure classification for the US[5] (American Medical Association 1999b; Bowman 2005; Libicki & Brahmakulum 2004; National Committee on Vital and Health Statistics Subcommittee on Medical Classification Systems 1993; Rode 2005; Slee 2000).

A study conducted in the Netherlands (Van Achterberg et al. 2005) transcribed nursing diagnoses to the *International Classification of Functioning, Disability and Health* (ICF) to evaluate the fit between that classification and hospital-based nursing practice. Van Achterberg et al. (2005) found that the classification was useful to nursing practice in acute care settings, and because human functioning is at the heart of nursing care, they recommended its use in this domain (p.440). They found that aggregated levels of the classification were a better match with nursing diagnoses, than the more detailed levels, and that the principal focus was on the body functions and activities components of the classification. Negative points raised in the study were that some aspects of nursing observations were found to be missing from the classification. Nurses also had problems finding terms because of a lack of familiarity with the structure of the classification and the language used to describe the terms. For these reasons, the authors recommended that nurses participate in future revisions of the ICF and that the classification be introduced into nursing curricula (Van Achterberg et al. 2005).

The substantial task of determining a suitable classification for general practice in Australia was undertaken by the General Practice Coding Jury (2000), and supported by government funding through the General Practice Computing Group (GPCG). Bearing in mind developments in electronic health records, the scope of this evaluation included terminologies as well as classifications. Each criterion was prioritised and ranked and systems were compared on that basis. The Jury's recommendation was to adopt the ICD-10-AM in the interim (on a 5 year short-term basis), with a view to ultimately implementing SNOMED-CT, if further assessments supported such a move (General Practice Coding Jury 2000: p.4).

[5] In the US, two procedure classifications are used to code clinical procedures. ICD-9-CM is used for in-patient services. CPT is used in both ambulatory and in-patient settings. Both systems are used for payment (National Committee on Vital and Health Statistics Subcommittee on Medical Classification Systems 1993: pp.8-9).

The Jury's recommendations were not adopted because there was no widespread support for them among key stakeholders (General Practice Computing Group 2002: p.2). One response in particular was a critique of the methodologies applied by the Jury (Britt & Miller 2000). The Australian Government commissioned an independent review to make recommendations on the best way to proceed. The GPCG has decided to implement one of those recommendations, which is to develop a vocabulary for general practice, for the areas of diagnosis and problems (General Practice Computing Group 2002: p.2).

Not all evaluations achieve their goals, as can be seen in the Jury's evaluation. It is wise to bear in mind that evaluations are innately political, primarily because of the vested interests of various stakeholders (Leys 2003).

Conclusion

Evaluation is important to evidence-based policy and practice in Health Information Management. Health classifications are important components of information systems and should be evaluated. My aim in writing this paper was to provide a revised and updated framework for evaluating health classifications that are used for statistical and reporting purposes. Additional guidance has been provided, particularly to novices, by illustrating the criteria with examples from the field, and using the language of nosology in context to enhance understanding of these terms. Examples of evaluation in practice have also been included.

The principles listed in this framework are not intended to be exhaustive, nor will each criterion be relevant to every evaluation. Rather, these criteria can be used as building blocks on which to base an evaluation, because each evaluation will differ in purpose and objectives. One important point to bear in mind is that the basic principles on which this framework was built will change as classifications, and their relationship with terminologies and information technology, develop over time.

Acknowledgements

I am grateful to Beth Reid and Jean McIntosh for providing useful suggestions to improve this paper. I am also fortunate to have had such a wise mentor in Rosemary Roberts.

References

American Medical Association (1999a). *Current Procedural Terminology CPT 2000*. Illinois, AMA.

American Medical Association (1999b). *Issues related to the use of ICD-10-PCS as the single procedure code set*. Available at: < http://www.ama-assn.org/ama/upload/mm/362/icd_10_pcs.doc> (accessed 14 June 2005).

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders. Fourth edition. Text revision*. Washington DC, APA.

Answers.com (2005). *Definitions for granularity*. (Webpage of GuruNet Corporation). Available at: <http://www.answers.com/granularity&r=67 (accessed 14 July 2005).

Australian Institute of Health and Welfare (2003). *The ICF Australian users guide.* Available at: <http://www.aihw.gov.au/publications/dis/icfaugv1/icfugv1.pdf> (accessed 17 June 2005).

Australian Institute of Health and Welfare (2002). The Australian family of health and related classifications. Available at: <http://www.aihw.gov.au/committees/ctwg/public_resources/princ_australia_300802.doc> (accessed 15 June 2005).

Bowker, G. and Star, S. (1999). *Sorting things out: classification and its consequences*. Massachusetts, The Massachusetts Institute of Technology Press.

Bowman, S. (2005). Coordination of SNOMED-CT and ICD-10: getting the most out of electronic health record systems – White Paper. *Perspectives in Health Information Management* (Spring). Available at: < http://library.ahima.org/xpedio/groups/public/documents/ahima/bok1_027179.html> (accessed 9 June 2005).

Bramley, M. (1994). *A report describing and evaluating four potential candidates for a procedure classification system to be used in conjunction with ICD-10 in Australia*. NCCH, unpublished paper.

Bramley, M. and Innes, K. (1998). *ICD-10-AM – Implications for allied health professionals'* Presentation to the *3rd National Allied Health Casemix Committee Seminar*, Melbourne.

Britt, H., Beaton, N. and Miller, G. (1995). General practice medical records: Why code? Why classify? *Australian Family Physician* 24(4): 612-615.

Britt, H. and Miller, G. (2000). *A critical review of the final report of the General Practice Coding Jury*. Available at: <http://www.fmrc.org.au/classifi-i.htm> (accessed 15 June 2005).

Campbell, K., Oliver, D. and Shortliffe, E. (1998). The Unified Medical Language System: toward a collaborative approach for solving terminologic problems. *Journal of the American Medical Informatics Association* 5(1): 12-16.

Canadian Physiotherapy Association (1994). CPA commentary on the report 'The effectiveness and cost effectiveness of chiropractic management of low back pain'. Available at: <http://www.ndir.com/chiro/cpa-manga.html> (accessed 9 June 2005).

Chute, C. (2000). Clinical classification and terminology: some history and current observations. *Journal of the American Medical Informatics Association* 7:298-303. Available at: <http://www.j-amia.org/cgi/content/full/7/3/298> (accessed 14 June 2005).

College of American Pathologists (n.d.). SNOMED — CT (Webpage of SNOMED International). Not known when last updated. Available at: <http://www.snomed.org/snomedct/index.html> (accessed 14 June 2005).

Evans, D., Cimino, J., Hersh, W., Huff, S. and Bell, D. (1994). Toward a medical-concept representation language. *Journal of the American Medical Informatics Association* 1(3): 207-217.

Family Medicine Research Centre (2005). ICPC-2 : Purpose. Available at: <http://www.fmrc.org.au/icpc2/purpose-d.htm> (accessed 30 June 2005).

Family Medicine Research Centre (1998). *The International Classification of Primary Care, Second edition, PLUS (ICPC-2 PLUS)*. Sydney, The University of Sydney.

Gardner, M. (2003). Why clinical information standards matter. *British Medical Journal* 326: 1101-1102. Available at: <http://www.bmj.bmjjournals.com./cgi/content/full/326/7399/1101> (accessed 19 April 2005).

General Practice Coding Jury (2000). *Final report of the General Practice Coding Jury.* Available at: <http://www.gpcg.org/publications/docs/GP_Coding_Jury.pdf > (accessed 19 April 2005).

General Practice Computing Group (2002). *General Practice Vocabulary Project Communique.* Available at: <http://www.gpcg.org/publications/docs/TerminologyProjectCommunique.pdf> (accessed 16 June 2005).

Harrison, J. (2001). New external cause categories in ICD-10-AM, third edition. *Coding Matters* 8(3): 1-4.

Herek, G. (2005?). *Facts about homosexuality and mental health.* Available at: <http://psychology.ucdavis.edu/rainbow/html/facts_mental_health.html (accessed 14 July 2005)

Hoffman, E. and Chamie, M. (1999). *Standard statistical classifications: basic principles*. Available at: <http://unstats.un.org/unsd/class/family/bestprac.pdf> (accessed 9 June 2005).

Hovenga, E. (2002). Nursing classification and terminology systems. *Health Information Management Journal* 31:(1). Available at: <http://www.himaa.org.au/members/journal/31_1_2002/hovenga.asp> (accessed 11 May 2005).

ICECI Coordination and Maintenance Group (2004). *International Classification of External Causes of Injuries (ICECI) version 1.2*. Amsterdam and Adelaide, Consumer Safety Institute and the Australian Institute of Health and Welfare National Injury Surveillance Unit.

Information and Communications Technology Standards Committee (2004). *Foundations for the future. Priorities for health informatics standardisation in Australia, 2005-2008.* Available at: <http://www.ahic.org.au/downloads/Foundations%20for%20the%20future.pdf> (accessed 9 June 2005).

Innes, K. and Bramley, M. (1997). *Development of an allied health intervention classification.* Presentation to the 2nd National Allied Health Casemix Committee Seminar, Brisbane.

Innes, K., Hooper, J., Bramley, M. and DahDah, P. (1997). Creation of a clinical classification. The international statistical classification of diseases and related health problems – tenth revision – Australian modification. *Health Information Management Journal* 27(1): 31-38.

Iowa Intervention Project (2000). *Nursing Interventions Classification (Third edition),* McCloskey, J.C. and Bulechek, G.B. (Eds). St. Louis, Mosby-Year Book.

International Organization for Standardization (2000). *ISO/DTS 17117 Health informatics — controlled health vocabularies — vocabulary structure and high-level indicators*. Geneva, ISO.

Lamberts, H. and Wood, M. (2002). The birth of the International Classification of Primary Care (ICPC). Serendipity at the border of Lac Leman. *Family Practice* 19(5): 433-435.

Leys, M. (2003). Healthcare policy: qualitative evidence and health technology assessment. *Health Policy* 65: 217-226.

Libicki, M. and Brahmakulum, I. (2004). *The costs and benefits of moving to the ICD-10 code sets*. Available at: <http://www.rand.org/pubs/technical_reports/2004/RAND_TR132.pdf> (accessed 13 June 2005).

Moskal, L. (2002). The implementation of new morbidity classifications in Canada. *Health Information Management Journal* 31(1). Available at: <http://www.himaa.org.au/members/journal/31_1_2002/moskal.asp> (accessed 11 May 2005).

Narrain, A. and Chandran, V. (2002?). *It's not my job to tell you it's okay to be gay. Medicalisation of homosexuality: a queer critique.* Available at: <http://www.altlawforum.org/PUBLICATIONS/Sexuality/medicalizationfinal.rtf> (accessed 14 July 2005).

National Centre for Classification in Health — ICD-10-AM Chronicle (n.d.). (Webpage of the NCCH: not known when last updated). Available at: < http://www3.fhs.usyd.edu.au/ncch/4.6.htm> (accessed 21 June 2005).

National Centre for Classification in Health (2005). *Updating ICD-10.* Webpage of the NCCH: last updated March 2005. Available at: <http://www3.fhs.usyd.edu.au/ncchwww/site/3.2.htm> (accessed 21 June 2005).

National Centre for Classification in Health (2004). *The International Statistical Classification of Diseases and Related Health Problems, 10th revision, Australian modification (Fourth Edition).* Sydney, NCCH.

National Centre for Classification in Health (2002). *The ICD-10-AM Mental Health Manual*. Sydney, NCCH.

National Committee on Vital and Health Statistics Subcommittee on Medical Classification Systems (1993). Recommendation for a single procedure classification system. *Journal of the American Health Information Management Association* 64(8): 12-22.

National Health Information Management Group (2002). *Health Information Development Priorities*. Canberra, Australian Institute of Health and Welfare.

National Health Service Information Authority — Clinical coding (2004). Webpage of the National Health Service: last updated 14 June 2004. Available at: <http://www.nhsia.nhs.uk/clinicalcoding/pages/opcs_4.asp> (accessed 14 June 2005).

Okkes, I., Becker, H., Bernstein, R. and Lamberts, H. (2002). The March 2002 update of the electronic version of ICPC-2. A step forward to the use of ICD-10 as a nomenclature and a terminology for ICPC-2. *Family Practice* 19(5): 543-546.

O'Rourke, M. (1997). What's in a name? *Medical Journal of Australia* 166: 372-373.

Parnanen, H., Kumpusalo, E. and Takala, J. (2000). Primary healthcare ICD — a tool for general practice research. *International Journal of Health Planning and Management* 15: 133-148.

Price, J.A. (1982). Medical classification systems: an overview. *Topics in Health Record Management* (June): 77–82.

Rigby, M. (2001). Evaluation: 16 powerful reasons why not to do it — and 6 over-riding imperatives. In: Patel, V., Rogers, R. and Haux, R. (Eds.) *The 10th World Congress on Medical Informatics (Medinfo 2001).* Amsterdam, IOS Press, pp.1198-1202.

Roberts, R. and Chalmers, K. (2002). Modelling ICD-10-AM (Australian modification). In *Proceedings of the 18th Patient Classifications Systems Europe Conference.* Available at: <http://pcse.org/articles/index.asp?id=127> (accessed 30 June 2005).

Roberts, R., Innes, K., Walker, S. and Scott, P. (2004). Avoiding a war of words (and numbers). The uncertain future of terminologies and classifications. *Journal of the American Health Information Management Association* 75(9): 26-32.

Rode, D. (2005). Setting AHIMA's 2005 federal agenda. *Journal of the American Health Information Management Association* 76(3): 18-20.

Rogers, J. (2003). *Medical terminology course material – Enumerative schemes.* Available at: <http://www.cs.man.ac.uk/mig/links/RCSEd/terminology.htm> (accessed 11 May 2005).

Sable, J., Nash, S. and Wang, A. (2001). Culling a clinical terminology: a systematic approach to identifying problematic content. In *Proceedings of the American Medical Informatics Association Symposium*, pp.578-582. Available at: <http://www.amia.org/pubs/proceedings/symposia/2001/D010001605.pdf> (accessed 14 June 2005).

Schmertmann, M. and Williamson, A. (2001). *NSW injury profile. A review of injury data (1995-1999) and identification of areas requiring further study.* Available at: <http://www.irmrc.unsw.edu.au/Publications/injuryprofile.asp> (accessed 14 June 2005).

Scott, P. (2003). Reference terminologies. *Health Information Management Journal* 31(3). Available at: <http://www.himaa.org.au/members/journal/31_3_2003/scott.aspp (accessed 11 May 2005).

Scott, P. (2004). *Introduction to terminologies*. Tutorial conducted for *Standards Australia Joint IT-14 Meeting* in Melbourne.

Scriven, M. (1991). *Evaluation Thesaurus*. California, Sage Publications.

Slee, V. (2000). The medical record is the key. *Partnerships* 11. Available: http://www.healthcommons.org/News/2000-06/medical_record.html (accessed 14 June 2005).

Smith, G., Bartlett, A. and King, M. (2004). Treatments of homosexuality in Britain since the 1950s – an oral history: the experience of patients. *British Medical Journal* doi:10.1136/bmj.37984.442419.EE Available at: <http://bmj.bmjjournals.com/cgi/content/full/bmj;328/7437/427> (accessed 14 July 2005).

Spackman, K. and Campbell, K. (1998). Compositional concept representation using SNOMED: towards further convergence of clinical terminologies. In *Proceedings of the American Medical Informatics Association Symposium* pp.740-744. Available at: <http://www.amia.org/pubs/symposia/D005143.PDF> (accessed 14 June 2005).

Stakelbeck, F. and Frank, U. (2003). From perversion to sexual identity. Concepts of homosexuality and its treatment in Germany. *Journal of Gay and Lesbian Psychotherapy* 7(1/2): 23-46.

Ustun, T.B. (2004). ICD revision process – towards ICD-11. Available at: <http://www.nordclass.uu.se/WHOFIC/papers/ICD_revproc_to_ICD-11_Ustun.pdf> (accessed 14 June 2005).

Van Achterberg, T., Holleman, G., Heijnen-Kaales, Y., Van der Brug, Y., Roodbol, G., Stallinga, H., Hellema, F. and Frederiks, C. (2005). Using a multidisciplinary classification in nursing: the International Classification of Functioning, Disability and Health. *Journal of Advanced Nursing* 49(4): 432-441.

Williams, S. and Rowlands, D. (2003). Development of a national standard for the identification of healthcare providers. *Health Information Management Journal* 31(3). Available at: <http://www.himaa.org.au/members/journal/31_3_2003/williams.asp> (accessed 11 May 2005).

WONCA - World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians (1998). *The International Classification of Primary Care, (Second Edition)*. New York, Oxford University Press.

Wood, M., Lamberts, H., Meijer, J. and Hofmans-Okkes, I. (1992). The conversion between ICPC and ICD-10. Requirements for a family of classification systems in the next decade. *Family Practice* 9(3): 340-348.

World Health Organization (2001).*The International Classification of Functioning, Disability and Health*. Geneva, WHO.

World Health Organization (1999). *Definition, diagnosis and classification of diabetes mellitus and its complications. Report of a WHO consultation. Part 1:Diagnosis and classification of diabetes mellitus.* Available at: <http://whqlibdoc.who.int/hq/1999/WHO_NCD_NCS_99.2.pdf> (accessed 11 May 2005).

World Health Organization (1997). *Multiaxial Presentation of the ICD-10 for use in adult psychiatry.* Geneva, WHO.

World Health Organization (1996a). *Multiaxial classification of child and adolescent psychiatric disorders*. Geneva, WHO.

World Health Organization (1996b). *Diagnostic and management guidelines for mental disorders in primary care: ICD-10 Chapter V Primary Care Version*. Germany, Hogrefe and Huber Publishers.

World Health Organization (1993). *The International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Vol. 2 Instruction Manual.* Geneva, WHO.

World Health Organization (1992a). *The International Statistical Classification of Diseases and Related Health Problems, Tenth Revision*, *Vol. 1 Tabular List*. Geneva, WHO.

World Health Organization (1992b). *The ICD-10 Classification of Mental and Behavioural Disorders – Clinical Descriptions and Diagnostic Guidelines*. Geneva, WHO.

World Health Organization (1977). *The International Classification of Diseases, 1975 Revision*, *Vol 1. Tabular List*. Geneva, WHO.

Wu, J. (1998). Open debate on de-pathologizing homosexuality in China. *Chinese Society for the Study of Sexual Minorities (CSSSM) News Digest, Special Issue*. Available at: <http://www.csssm.org/English/e3.htm> (accessed 14 July 2005).

Michelle Bramley *BAppSc(HIM)*
Lecturer
School of Health Information Management
Faculty of Health Sciences
The University of Sydney
PO Box 170
Lidcombe, NSW 1825
Phone: +61 2 9351 9451
Facsimile: +61 2 9351 9672
Email: michelle.bramley@fhs.usyd.edu.au