

Modelling and encoding Therapeutic Guidelines: applying ICD-10-AM and European Article Number codes

Bryn Lewis, Ken Harvey, Jonathan Dartnell and Jennie Shephard

Abstract

This article reports on the assignment of ICD-10-AM and EAN codes to 2500 topics in Therapeutic Guidelines (TG). The analysis of the assignment of ICD-10-AM codes in this project has revealed that ICD-10-AM is not capable of describing the complete clinical information in the guidelines series. It is not likely that any existing single classification scheme will be capable of this and that a combination of schemes will be necessary. The TG data model was integrated with the prototype MCCA data model for drug products. This integration indicates that the representation of drugs, while not ideal, is an appropriate means of linking clinical drug reference information to drug product information.

Keywords: *Clinical classification systems; ICD-10-AM; ATC codes; data modelling; decision support systems.*

Introduction

Computerised clinical decision support systems have been shown to improve medical practice by comparing patient characteristics with credible information resources and offering clinicians succinct patient-specific and knowledge-specific advice at the time of clinical decision-making (Kidd & Mazza 2000). However, if disparate clinical information resources are to integrate seamlessly with each other and with decision support software, then standardised codes must be applied to their indexing and linking terms.

A General Practice Computer Group (GPCG) coding jury recommended the *International Classification of Diseases, 10th revision, Australian Modification (ICD-10-AM)* (World Health Organization 2000) as the interim clinical coding system for use within general practice in Australia (Commonwealth Department of Health and Aged Care 2000). In a development arising from the Better Medication Management System, the Medicines Coding Council of Australia (MCCA) is implementing European Article Number (EAN) codes (EAN International 1998) as the standard for electronic transmission of medicines data (Medicines Coding Council of Australia 2001).

Therapeutic Guidelines Limited (TGL) is a long established provider of best-practice guidelines concerning the treatment of choice for common conditions (Hemming 2000). TGL has recently converted their entire range of printed guidelines into first generation electronic products (in HTML format). This included extensive work on optimising indexes and hyperlinks in order to assist clinician look-up and problem solving.

TGL has recently embarked on a research program to optimise the production of their guideline products

and to enable the Guidelines to be used within electronic decision support systems (Lewis B 2001a). The research program is based around a Strategic Partnerships in Industry Research and Training (SPIRT) project entitled "Towards computerised clinical decision support systems: optimising the production of collaborative electronic guidelines". In addition, two GPCG sub-projects were carried out within the SPIRT research project. The program structure is as follows:

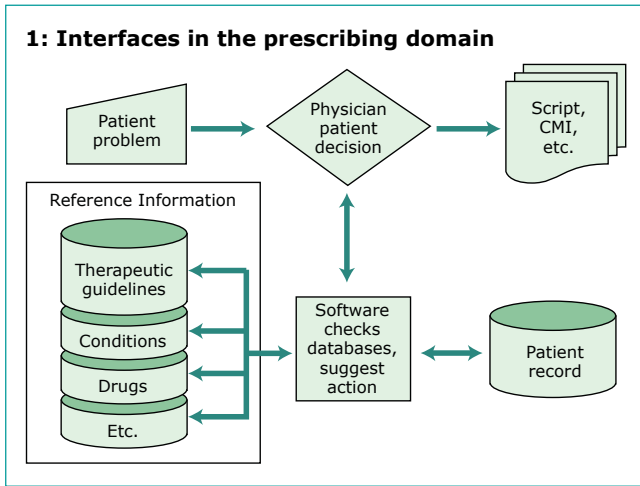
- SPIRT project
 - Coding sub-project
 - ICD-10-AM coding
 - EAN coding
 - Therapeutic Guideline (TG) data-modelling sub-project.

This article reports on the findings of the encoding sub-project within the SPIRT project. The motivation for the sub-project was to investigate means of integrating the guidelines information with external clinical software systems. While there are many candidates for a health coding and/or classification system for the purposes we describe, none completely satisfies all the requirements. The discussion has been restricted to those schemes that were directly examined.

The aim of the encoding sub-project was to investigate the viability of assigning ICD-10-AM codes to the topics, and EAN codes to the drugs, in the Guidelines. The assignment of an ICD-10-AM code would be judged successful if the interpretation of the meaning of the code was in accord with the meaning of the Guideline topic to which it had been assigned. A successful assignment in this sense would allow the code assignment to be used as part of a decision support mechanism. The assignment of an EAN code would be judged successful if the drug information referred by the code corresponded to the drug information in the MCCA database that was linked to that code. The assignment was carried out over the whole of the Guidelines, which at the time comprised nine titles and over 2500 clinical guideline topics. As the assignment of ICD-10-AM and EAN codes are two quite different activities, much of the discussion refers to one or the other.

The data-modelling sub-project within the SPIRT project has resulted in the implementation of a data model of the information in the Guidelines and a por-

A note on terminology: Throughout this article, the entity Therapeutic Guidelines Ltd is referred to as 'TGL'. The publications TGL produces are referred to as the Guidelines (when referred to as a collective), or Guideline topics (when referred to individually). In discussing the Guidelines in the context of clinical guidelines in general, they are referred to as the Therapeutic Guidelines to distinguish them from guidelines produced by other organisations.



2: Pack configurations for doxycycline

| | Product form | Active amount | Unit | No per pack | No of packs required for recommendation |
|---|--------------|---------------|------|-------------|---|
| 1 | capsule | 100 | mg | 7 | 1.0 |
| 2 | capsule | 100 | mg | 21 | 0.33 |
| 3 | capsule | 100 | mg | 28 | 0.33 |
| 4 | capsule | 50 | mg | 25 | 0.56 |
| 5 | tablet | 100 | mg | 7 | 1.0 |
| 6 | tablet | 100 | mg | 21 | 0.33 |
| 7 | tablet | 100 | mg | 28 | 0.33 |
| 8 | tablet | 50 | mg | 25 | 0.56 |

tion of the GPCG’s GP data model (prepared by Sim-ion Bowles and Associates for the Commonwealth Department of Health and Aged Care 2000). The resultant implementation defines the data structures necessary to create a database of the information in the Guidelines (Lewis B 2001b). This database structure is at a level of detail that allows the manipulation of discrete items of Guideline information, such as drugs, dosages and their relationship to Guideline topics. Provision for the ICD-10-AM and EAN codes was made in the data model. While the two sub-projects are related, this article does not discuss the data-modelling sub-project, except where it impacts the encoding project.

The article is structured as follows. We first present a simple scenario of the sort of application we would like to be achievable using the Guidelines information and briefly discuss the conceptual framework we are building to fulfil the requirements of the scenario. We then discuss the method by which ICD-10-AM codes, then EAN codes, were assigned to Guideline topics. The results of the assignment are then presented, followed by a discussion and conclusion leading from the results.

Drug coding and Therapeutic Guidelines

Box 1 depicts general interactions of the Guidelines in the health domain. The Guidelines are shown as an example of standardised reference information. In discussions of decision support (eg, Commonwealth Department of Health and Aged Care, National Health Information Management Advisory Council 2001; Fox & Das 2000; National Institute of Clinical Studies 2001; Wroe et al 2001), it has been suggested that developing a standard means of describing the various components of a guideline, such as the drugs involved, the conditions concerned and the treatments and recommendations advocated, is necessary. This is a requirement, as it allows clinical systems to manipulate the Guideline information to achieve electronic decision support. However, standard Guideline representation is only useful, or even possible, after more central concepts are consistently represented in a repository that can be accessed by clinical systems. An example component of such a repository is a regularly updated drugs database that is a comprehensive and public listing of all drugs available nationally. Such a data-

base needs to include product information at a level of detail that allows prescribing systems to specifically identify a product pack, for example by a barcode on prescriptions.

One means of realising interactions between the components of a repository is by utilising clinical coding schemes. These are the codes that uniquely identify individual elements in storage. In the case of drugs, Anatomical–Therapeutic–Chemical (ATC) drug classification codes (Department of Health and Family Services 1998) can identify drugs, and EAN codes have been developed to identify drug products. Clinical conditions can be identified with ICD-10-AM codes, although many candidates for this task have been developed. ICD-10-AM has been used in Australian hospitals since July 1998 as a means of describing a patient’s clinical classification. It is designed principally for epidemiological analysis and classification.

A simple scenario

A simplified scenario is as follows. An adult patient presents to a clinician with probable infective exacerbation of chronic bronchitis. This is entered in the electronic medical record as the provisional diagnosis. The computer assigns the relevant ICD-10-AM code and alerts the clinician that relevant clinical guidelines are available. Consulting the Guidelines, the clinician determines that the patient meets criteria for the commencement of antibiotic therapy and selects one of the recommended options:

amoxycillin 500mg orally, 8-hourly for 5 days

A warning alert notes that the patient has a suspected history of penicillin hypersensitivity and offers an alternative regimen:

doxycycline 200mg orally, initially, then 100mg daily for 5 days.

Doxycycline is available in eight pack configurations, as shown in Box 2. The recommendation specifies 6 doses, totalling 700mg. Configuration 1 or 5 is the most likely choice and is dependent only on brand/price and tablet/capsule preference.

Once a clinician has accepted a recommendation from the Guidelines, the available products could be found through reference to the unique record for each generic drug. The information could be easily pre-

sented, making the administrative side of prescribing more efficient and accurate. As these product configurations are specific to a product pack, the barcode for the pack is also available for inclusion on the prescription, with obvious benefits for dispensing.

Representing Guideline information

The information in the Guidelines series is presented as natural language text. As such, it is not inherently structured. It does, however, include relationships between highly structured information, as the example scenario demonstrates. These include relationships involving (but not limited to) clinical conditions, drugs and recommendations.

The data model of the Guidelines that has been developed in a related sub-project has decomposed those elements of the Guidelines, such as drugs and recommendations, which are inherently structured. Unstructured information, such as discussions in the text, has not had a structure imposed on it as yet, as it is not clear if this is possible or useful.

The reference information depicted in Box 1 is that information which is common to all the clinical information systems that need to access it. The drug product information used in the scenario needs to be available to the electronic guidelines, the prescribing system and the drug-dispensing system. It needs to have a guaranteed common format and content at a given time. This is true of other components of reference information, such as clinical conditions.

The ICD-10-AM and EAN projects described in this article were undertaken to advance the development of formats for these types of common reference models. In the case of the EAN project, we developed that portion of the TG data model so that it was compatible with the MCCA data model for drugs. The MCCA data model was not developed to make representing guideline information easier. However, because the elements involved in representing drugs are relatively well understood, the MCCA data model could be readily incorporated into the TG data model. Should the MCCA database project reach the stage of being a comprehensive, nationally available source of drug product information, the simple scenario we have presented can be realised. Without such an information source, even this simple application is likely to founder in attempting to cover all the drugs involved. This type of infrastructure is essential to the development of further electronic decision support mechanisms.

Method

ICD-10-AM code assignment

Extraction of topics

The extraction of the topics took place prior to the establishment of a Guidelines database; thus, extraction was from HTML files rather than from a structured database. The HTML files used were those that made up the content of the Guidelines electronic products at the time: *Antibiotic 11* (Writing Group for Therapeutic Guidelines: Antibiotic. Therapeutic Guidelines Limited 2000), *Analgesic 3*, *Cardiovascular 3*, *Dermatology 1*,

3: Extraction of guideline

| | <u>Number</u> | <u>Percentage of total</u> |
|---------------------|---------------|----------------------------|
| Total topics | 4593 | 100.00% |
| Non-clinical topics | 1657 | 36.08% |
| Drug topics | 444 | 9.67% |
| Clinical topics | 2492 | 54.26% |

Endocrinology 1, Gastrointestinal 2, Neurology 1, Psychotropic 4 and Respiratory 2.

Guideline topics are arranged in a hierarchy, with child topics being subject specialisations of the parent topics. A reduced example set of parent-child topics is

- Central nervous system infections
 - Meningitis
 - Hospital-acquired meningitis.
 - Brain abscess or subdural empyema

A topic is a section of the Guideline text that is considered to be an individual, isolated guideline. Each of these headings is a topic. The meaning of the topic is derived from the topic heading and the text contained within the topic. Editorial staff members identify an individual topic. Leaf topics (topics that have no child topics) represent the smallest amount of text that is allowed to occur on its own HTML page in the electronic product.

Extraction consisted of developing a simple parser that identified each of the topics and its relationship within the topic hierarchy. Each of the topics was then exported to an MS Excel file in a format that retained the identification of topics and the hierarchical relationship between them. The hierarchy relationship was indicated through a value placed in the "parent column" for each topic. All topics in the guidelines were extracted. The breakdown of extracted topics is indicated in Box 3.

Non-clinical topics: As the parser was not able to determine the meaning of the topics, the extraction included those topics that should not properly be assigned a code of any sort. For example, publication information is structured as a topic, but it is clear to a user that this is not clinical information. These topics were not removed from the topic listing, but were passed to the coders. The coders were made aware of this situation and asked to identify these "non-clinical topics".

Drug topics: These topics are descriptions of drugs or of the use of drugs. Some of these could have been assigned an ICD-10-AM code, but EAN and ATC codes provide a more rigorous approach to drug coding and so ICD-10-AM was not used. It should be noted that these topics provide advice on the use of drugs in the context of treatment determination. While it is difficult to identify a specific condition for which they would be referred to, they are still treatment guidelines and a means of retrieving them during a consultation is necessary.

Clinical topics: After eliminating the Non-clinical topics and Drug topics, 2492 topics were left for assignment.

Assignment of ICD-10-AM codes

A team of five experienced Health Information Managers undertook the assignment of ICD-10-AM codes to the Guidelines' topics. The list of topics was presented to the coders as a spreadsheet with several columns. The Guidelines titles were provided to each coder, allowing each topic heading and its text to be cross-referenced and the meaning of the topic to be determined. For each clinical topic, the coders assigned values under the following headings.

- ICD-10-AM code

The primary ICD-10-AM code was assigned in the standard manner, except that Australian Coding Standards were not applied. Coding standards provide guidance in situations where multiple codes are assigned to one episode of care. This was not the intention of the assignment exercise.

- Additional code

In cases of a dagger/asterisk combination, the first code assigned was that which was most appropriate for the topic. In these cases, the other code was assigned as an "additional code". Morphology codes were also assigned as additional codes.

- Alternative ICD-10-AM code

Alternative codes were those codes that could be used for the topic as well as the code that was assigned. These included site codes, or those codes that provide more specific detail.

- Basis of coding decision

The "basis of coding decision" comment explained the coding decision in cases in which alternative codes were available. Often this comment was that a default code was used, but in some cases it indicated that the code assigned was most appropriate for the topic according to the text of the guideline. Complete codes were assigned.

- Coding comment

A "coding comment" could be made by the coders to add any open-ended information that explained the coding decision, if an explanation was felt necessary.

- Code weighting

All cases in which one code did not sufficiently describe a topic were noted with a weighting system. The weighting system was used to enable a more systematic analysis of the results of code assignment, as well as to identify which topics might need re-examination. The weights were as follows:

- ◊ "S" — Specific when the code title or inclusion note exactly described the topic.

"S" was assigned when the code descriptor or the inclusion terms for the code were an exact match for the topic or the topic was encapsulated in the code descriptor. The following are examples of codes that were considered specific code allocations.

Example: Congenital (neonatal) hypothyroidism was coded to E03.1, *congenital hypothyroidism without goitre*, which has an inclusion term for *congenital hypothyroidism NOS*:

Acromegaly was coded to E22.0, *acromegaly and pituitary gigantism*.

- ◊ "A" — Appropriate when the code or inclusion term was not an exact match but was appropriate.

"A" was assigned when the assigned code was the default code listed in the index to diseases but there were other specific codes available.

Example: Pertussis is indexed to A37.9, *Whooping cough unspecified*. However it also has a "see also" instruction for whooping cough (*see also Whooping cough*) that directs the coder to several other specific codes for whooping cough with pneumonia or whooping cough due to bordetella.

If a topic included the term "acute" or "chronic" and the code was otherwise an exact match it was weighted as "A" rather than "S".

Example: Acute herpes zoster is assigned B02.9, *Zoster without complication*, on the basis of index look up Zoster (herpes zoster). But, as there is no listing for acute herpes zoster, it is assigned a weight of "A". Similarly "paediatric tinea" was assigned the generic code for tinea and assigned "A" rather than "S", and oral fungal infection was coded to B49, *unspecified fungal infection*, and assigned "A", as there was no specific entry for 'oral' fungal infection.

- ◊ "N" — Neither when the code approximated the topic but was neither of "A" or "S".

"N" was assigned when the coder felt that the topic included significantly more than the code described.

Example: Subungual warts was assigned code B07, *viral warts*. However, the coder felt that the term "subungual" sufficiently changed the topic to make a code of B07 not appropriate.

Example: Tuberculosis and HIV co-infection required two codes to fully describe the topic. As both components of the topic were clinical entities in their own right, and only one code could be allocated, it was impossible to assign one specific code.

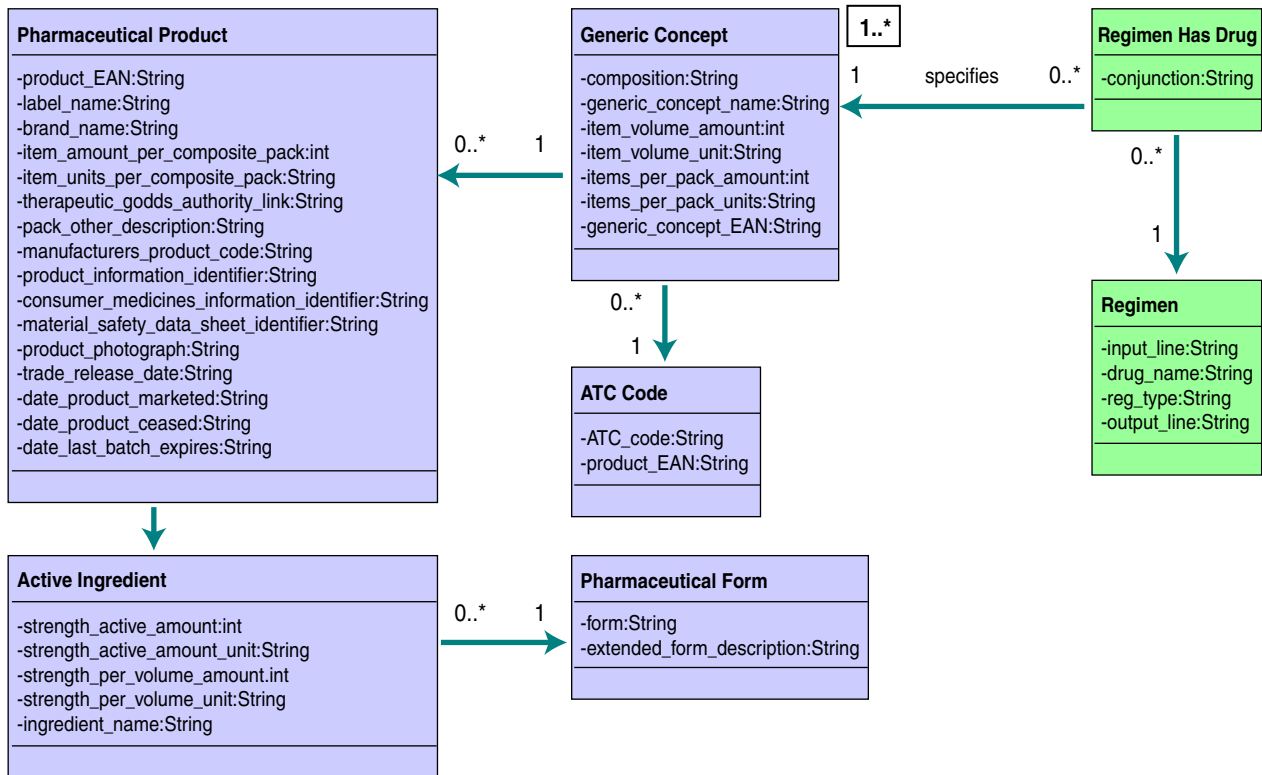
Example: Oral white patches was assigned code B37.0, *candidal stomatitis*. In this case, the coder has assigned a code from the index entry, *white mouth*. However, reference to the appropriate Guidelines text indicated that oral white patches could be due to several causes and therefore the weighting was assigned as "N".

The final data file was reviewed by the project manager and some changes made to achieve consistency across the file in allocation of "no code", weight assignment and accuracy of code assignment. The assessment of the allocation, its specificity and its appropriateness was judged in terms of whether the allocation would allow the concept being classified to be incorporated within a decision support mechanism. This is not the purpose for which the ICD-10-AM codes were designed and is a different view of the "success" or otherwise of code assignment.

EAN code assignment

EAN codes are unique numbers that can be placed as barcodes onto the packaging of products. They are designed for supply-chain management and are not intended to semantically classify products, merely to indicate a unique identifier.

4: EAN/ATC portion of the TG data model



The assignment of EAN codes was an exercise in integrating data models rather than a code-assignment exercise. The integration was achieved by assessing the semantics of the two data models and determining where they were in accord. Box 4 is a portion of the TG data model, indicating how the MCCA prototype data model was modified and connected to the TG data model. The purple/grey classes are adaptations from the MCCA data model and the green classes are from the TG data model.

A principal difficulty in the MCCA data model is that there is no provision for a unique record for each generic drug. The Generic Concept class in the model includes an item_volume or items_per_pack amount. This has repercussions when linking records. Referring back to the simple scenario outlined above, the entries for doxycycline are as listed in Box 2. There is no single representation for doxycycline, as the representation is made up of the collection of six records.

The cardinality of the association between Regimen_Has_Drug and Generic_Concept is one to many (1..*). The more appropriate association would be a single link (1). The actual cardinality is highlighted in a box in the diagram, with the ideal cardinality being the alternative. While this merging of the TG and MCCA data models is not ideal, the uniqueness of drug names would make the arrangement workable. Extensions to the data models to include, for example, drug interactions would benefit from the development of a unique record for each generic drug.

The linking of the data values involved determining the generic_concept_EAN for each drug name in the Guidelines. Each drug in the guidelines has a name that corresponds to the generic_concept_name for one

or more Generic_Concept. Thus, each drug in the guidelines corresponds to at least one generic_concept_EAN. The remainder of the information in the MCCA data model can be related to the generic_concept_EAN.

Results

ICD-10-AM assignment results

Box 5 summarises the number of results of the allocation exercise: of the 1513 topics that were allocated a code, 53% (806) were assigned a specific code (S), 39% (589) were allocated an appropriate code (A) and 7% (104) were assigned a code that was neither specific nor appropriate (N). This means that of the 2492 clinical topics in the guidelines, only 32% could be assigned an ICD-10-AM code that reflected the meaning of the topic; of the remainder that were assigned a code, 27% could be assigned an appropriate code, but the assignment did not fully describe the meaning of the topic. This indicates that, if ICD-10-AM were to be applied successfully across all Guideline topics, considerable reinterpretation of ICD-10-AM codes and their method of assignment would be required.

There was a 39% failure rate in code assignment. It is difficult to characterise all of these topics, but an attempt follows. Some of the topics relate to a specific clinical activity rather than a diagnosis. For example, "measuring blood pressure in pregnancy". Some have a diagnostic component and an administrative component — for example, "mild to moderate hypertension – general recommendations". Some of these topics provide generic descriptions, such as "trauma pain".

A large number of the topics that could not be allocated a code are descriptions of treatment management, such as “Principles of management of lung cancer” and “Cessation of anti-epileptic therapy”.

Many of the topics that have been allocated “no code allocated” could be assigned a code if it were imperative to do so; however, the assigned code would lose its relevance as a pointer to a particular area of the guidelines. For example, “pain” is included in over 200 Guidelines topics. Many of these describe a type of pain rather than a site of pain (eg, *trauma pain, cancer pain and postoperative pain*). ICD-10-AM contains about 50 codes for pain, four of which are generic pain codes (R52.0, acute pain; R52.1, chronic intractable pain; R52.2, other chronic pain; and R52.9, pain unspecified). The other pain codes in ICD-10-AM are site-specific codes.

Discussion

ICD-10-AM assignment

The principal difficulty in using ICD-10-AM as a means of representing all of the information in the Guidelines arises because it is not designed for this purpose. ICD-10-AM can be applied successfully when the concept being coded is presented in a simple manner and the context needed for its interpretation is clear and unambiguous. However, ICD-10-AM is not consistent in the application of the axes of classification it uses, and it is not explicit as to what those axes are.

The following is a summary of how this problem manifests itself:

- In some cases ICD-10-AM has a greater level of granularity than the Guidelines.
- In some cases, the Guidelines have a greater level of granularity than ICD-10-AM.
- In some cases ICD-10-AM matches the granularity of the guidelines, but the concepts being described are orthogonal to those in the guidelines.

The following discussion of the assignment exercise is structured around several issues related to the nature of ICD-10-AM.

ICD-10-AM as an epidemiological tool

ICD-10-AM is a classification designed specifically for epidemiological purposes and therefore not designed to have a unique code for every condition. Many ICD-10-AM codes have been used more than once in this exercise.

Multiple coding

The assignment of multiple codes is the only way in which the meaning of a great deal of the Guideline topics can be fully described. If a policy of assigning multiple codes was to be adopted it would be necessary to develop a consistent policy for their application. The examples given below highlight some of the difficulties in developing such a policy.

- ICD-10-AM has a dual coding convention known as the “dagger and asterisk” convention. This allows for cases in which the aetiology and manifestation of certain conditions occur in different chapters to

be coded using two codes. For example, female genital herpes is coded to A60.0 and N77.x, and male genital herpes is coded to A60.0 and N51.x.

- In addition to the ICD-10-AM dual coding convention, many topics contained more than one clinical concept and therefore required more than one code to describe the statement.
- Some topics required an external cause code to fully describe the topic (eg, iodine-induced hypothyroidism in infancy coded to E03.2, *hypothyroidism due to medicaments*, and Y57.6, *other diagnostic agents (iodine)*).
- Neoplasm coding in ICD-10-AM convention requires the use of two codes, one for the site and behaviour, and one for the morphology. Default codes were assigned for most neoplasms, as ICD-10-AM had site codes but the Guidelines did not. For example, melanoma was coded to C43.9, *malignant melanoma of skin, unspecified (site)*, and BCC was coded to C44.9, *malignant neoplasm of skin, unspecified (site)* and 9 site codes were assigned as alternative codes for each of them. Additional codes were then assigned for the morphology (from the *International Classification of Diseases*

5: Statistical analysis of coding assignment

| | <u>Number</u> | <u>Percentage of clinical topics</u> | <u>Percentage of codes allocated</u> |
|--|---------------|--------------------------------------|--------------------------------------|
| Clinical topics available for assignment | 2492 | | |
| No code allocated | 979 | 39% | |
| Codes allocated | 1499 | 60% | |
| Specific codes (S) | 806 | 32% | |
| Appropriate codes (A) | 589 | 24% | 39% |
| Neither specific nor appropriate codes (N) | 104 | 4% | 7% |
| Codes used more than once | 305 | | |

for *Oncology – ICD-0*, which is included in ICD-10-AM). For example, M8720/3 for melanoma, M8090/3 for BCC.

- Bacterial or viral infections in nappy rash has an “and/or” combination of codes to describe the topic – L22, *diaper dermatitis*, and A49.9, *bacterial infection, unspecified*; or L22 and B34.9, *viral infection, unspecified*. L22 was assigned to the topic but was weighted “N” (nappy rash by itself does not describe the infection related to the nappy rash and, as not all nappy rash results in infection, the code was considered inappropriate).
- ICD-10-AM allows for the assignment of additional codes to describe the organism responsible for certain infections. Thus, streptococcal perianal dermatitis, vulvitis or balanitis require an additional code for the organism and also have an “and/or” component to the topic. Thus codes could be L30.8, *other specified dermatitis*, and B95.5, *unspecified streptococcus* as the cause of diseases classified to other chapters; or N76.2, *acute vulvitis*, and B95.5; or N48.1, *balanoposthitis*, and B95.5.

Where more than one code was required to describe the topic, additional code(s) were listed in the "additional ICD-10-AM code" column and the weighting was generally assigned as "Appropriate", although some were assigned "N" for neither specific nor appropriate.

Severity

ICD-10-AM generally does not describe the severity of a condition, although there are some exceptions. Many of the topics in the Guidelines refer to "mild" "moderate", or "severe" conditions. These were generally coded to the generic code for the condition and were weighted as being appropriate.

Site axis versus aetiology axis

ICD-10-AM codes and the Guidelines differ in the axes used to describe various conditions. This is complicated further by inconsistency in approach across the range of conditions discussed in both information sources.

- ICD-10-AM has site codes for many conditions where the Guidelines do not. For example, herpes zoster in ICD-10-AM has several code combinations for the nervous system and for ocular sites; sinusitis is coded in ICD-10-AM to ethmoidal, frontal, maxillary, etc.; cellulitis is coded to finger, toe, upper limb, lower limb, etc. The Guidelines have topics for orbital cellulitis but no other site-specific topics for cellulitis.
- Some Guidelines topics describe sites in a different way to ICD-10-AM. For example, the Guidelines describe "right ventricular infarction" and ICD-10-AM has codes for "anterior", "posterior" and "lateral" myocardial infarction.

Other code options

- ICD-10-AM "Z" codes have been used for some topics. These codes usually describe either a reason for contact with the health service or a particular health status of the patient. They have been used with the discretion of the coders, but have mostly been assigned a weighting of "A" or "N".

Example: *Pre-treatment screening and monitoring for pulmonary tuberculosis* was assigned Z13.82, with a weighting of "A".

Alternative codes

There has been no weighting assigned to the alternative codes. However, there was an attempt to include only those codes that could be considered appropriate upon reference to the Guidelines text.

Some neoplasm behaviour was not described in the Guidelines topic. The default code provided in the index was used to assign the code. For example, *Phaeochromocytoma* has index entries for D35.0 (benign) and for C74.1 (malignant).

EAN assignment

The MCCA prototype structure is both simple and adequate, two qualities that recommend it as a data

model. The challenge in maintaining a database of this set of information is not technical from a database design point of view. The challenge is organisational, as the information needs to be updated regularly and access to it must be maintained. Should the MCCA database reach general acceptance as an authoritative database of drugs, the Therapeutic Guidelines database will be able to link its drug and recommendation information to other systems that also include MCCA structures in their systems. The research into development of the Guidelines for electronic decision support will keep abreast and contribute where possible to the development of a public drugs database.

Dealing with drug substances in this objective fashion is appropriate from the point of view of standards development and software system development. The approach does not, however, deal completely with the complexity of drug information as it is presented in the Guidelines. For example, the topic "Cessation of anti-epileptic therapy" can not be completely described by assigning a code to indicate the drugs relevant to "anti-epileptic therapy". There is no ICD-10-AM code to deal with this topic either. This is far from an isolated case, demonstrating that it is likely that a combination of coding schemes would be required to cover all the concepts discussed in the Guidelines.

Conclusion

This article has reported on the assignment of ICD-10-AM and EAN codes to the topics in the Guidelines. Sixty percent of the Guideline topics involved in the ICD-10-AM assignment could be assigned a code, of which 32% were successfully assigned a code and 28% were assigned a code that was not specific or neither specific nor appropriate. The analysis of the assignment of ICD-10-AM codes in this project has revealed that ICD-10-AM is not capable of describing the complete clinical information in the Guidelines series. It is not likely that any existing single classification scheme will be capable of this and that a combination of schemes will be necessary. In order to utilise ICD-10-AM it will be necessary to establish a set of procedures to allow the assignment of multiple codes to Guideline topics, thus allowing a more complete description of all topics.

The project has established a means of assigning codes to the Guidelines topics. Maintenance of code assignment will be required as the topics are developed in future Guidelines. In addition, this project has established a means of investigating additional classification schemes. It is yet to be determined what utility ICD-10-AM will have when it is used as a means of linking the Guidelines to information systems in a general practice or a hospital setting. These issues will be investigated as TGL continues its research in the development of electronic decision support.

The integration of the TG data model with the prototype MCCA data model for drug products has been achieved. The ease of integration indicates that the representation of drugs utilising EAN and ATC codes, while not ideal, is an appropriate means of linking clinical drug reference information to drug product information. Developing linkages between the electronic guidelines and third-party prescribing systems re-

quires the ongoing development and maintenance of a national drug database compatible with the MCCA structure.

Acknowledgements

The article describes research funded by the Australian Research Council, Therapeutic Guidelines Limited, the National Prescribing Service and the General Practice Computing Group.

References

- Cancer Research UK (2001). *Methods for the Representation of Clinical Guidelines*. Methods for the Representation of Clinical Guidelines: One-day Workshop. London: Cancer Research, UK.
- Commonwealth Department of Health and Aged Care (2000). *Final report of the General Practice Coding Jury: 67*. Retrieved from www 14/07/04 <www.gpcg.org/publications/docs/GP_Coding_Jury.pdf>.
- Commonwealth Department of Health and Aged Care. National Health Information Management Advisory Council (2001). *Setting the Standards: a national health information standards plan for Australia*. Canberra: Commonwealth Department of Health and Aged Care.
- Department of Health and Family Services (1998). *Australian statistics on medicines*. Canberra: Australian Government Publishing Service.
- EAN International (1998). An Introduction to EDI, EAN International 84. Retrieved from www 14/07/04 <www.ean-int.org/index800.html>.
- Fox J, Das S (2000). *Safe and sound: artificial intelligence in hazardous applications*. Menlo Park, Calif: AAAI Press/MIT Press.
- Hemming M (2000). Therapeutic guidelines: an Australian experience. *International Journal of Pharmaceutical Medicine* 14(4): 259-264.
- Kidd MR, Mazza D (2000). Clinical practice guidelines and the computer on your desk. *Medical Journal of Australia*. 173: 373-375.
- Lewis B (2001a). *Data modelling of Therapeutic Guidelines: final report*. Melbourne: General Practice Computing Group.
- Lewis B (2001b). Therapeutic Guideline development for integration with General Practice Information Systems. Realising Quality Healthcare: Health Informatics Conference.
- Medicines Coding Council of Australia (2001). MCCA Home page. 2002. Retrieved from www 00/00/00 <http://www.cecc.com.au/mcca/default.html>.
- National Institute of Clinical Studies (2001). *NICS-NHIMAC Electronic Decision Support Governance Workshop: Melbourne, 23 November 2001: responses from participating organisations*. NICS-NHIMAC Electronic Decision Support Governance Workshop, Melbourne: National Institute of Clinical Studies.
- Simsion Bowles and Associates for the Commonwealth Department of Health and Aged Care, in close collaboration with the General Practice Computing Group (2000). *General Practice Data Model & Core Data Set Project: Final project report*. Writing Group for Therapeutic Guidelines: Antibiotic. Therapeutic Guidelines Limited (2000). *Therapeutic Guidelines: Antibiotic*. North Melbourne, Vic: Therapeutic Guidelines Limited.
- World Health Organization (2000). *The International Statistical Classification of Diseases and Related Health*.
- Wroe CJ, Roberts AK, et al (2001). Drug information sources: linking a clinical drug ontology to product information databases. *Healthcare Computing 2001 (HC2001)*. London: BCS Health Informatics Committee. Retrieved from www 04/06/04 <www.cs.man.ac.uk/mig/projects/old/drugontology/papers/Databse-mapping.zip>.

Bryn Lewis BSc, GDipLibInfSc, GDipCompSc, GDipAdvCompSc, PhD

Research Fellow
Department of Information Systems
University of Melbourne, VIC
and
Software Research and Development Consultant
Therapeutic Guidelines Ltd
North Melbourne, VIC
Tel: +61 3 9326 9959
Fax: +61 3 9326 5632
e-mail: blewis@tg.com.au

(At time of writing: Research Fellow
Department of Computer Science and Computer
Engineering, La Trobe University, VIC)

As a research fellow in the Department of Computer Science at La Trobe University, Bryn Lewis was the chief investigator of a SPIRT project, 'Towards computerised clinical decision support systems: optimising the production of collaborative electronic guidelines'. Dr Lewis' principal research interests are in health informatics, decision support and management and implementation of electronic guidance. His background is in software system design and development.

Ken Harvey MBBS, FRCPA, MASM

Senior Lecturer, School of Public Health
La Trobe University, Bundoora, VIC
Tel: +61 3 9479 5773
Fax: +61 3 9479 1783
e-mail: k.harvey@latrobe.edu.au

Jonathan Dartnell BPharm, PhD

Production Manager, Therapeutic Guidelines Limited
North Melbourne, VIC
Tel: +61 3 9326 9959
Fax: +61 3 9326 5632
e-mail: jdartnell@tg.com.au

Jennie Shephard RMRL, GDipHthAdmin, CertHthEco

Clinical Coding Consultant and Sessional Lecturer
Department of Health Information Management
School of Public Health
La Trobe University, Bundoora, VIC
e-mail: J.Shephard@latrobe.edu.au