Anatomical Characterisation of Surgical Procedures in the Read Thesaurus

Colin Price FRCS, Timothy E Bentley BDS MSc, Philip JB Brown MRCGP,
Erich B Schulz MB BS, Mike O'Neil MRCP (UK).
NHS Centre for Coding and Classification, Loughborough, United Kingdom

Each concept in the surgical operations chapter of the Read Thesaurus has been analysed to determine its anatomical site component. The underlying structure of this chapter and its relationship to the anatomy chapter are explored. The defined anatomical sites have been included as atomic maps in the Read Code template table, one of the key component files of the Thesaurus, relevant features of which are described. The analysis methodology is outlined and the value of an anatomically characterised surgical procedure terminology is discussed together with the implications of semantically defining a wider range of characteristics of surgical procedures.

INTRODUCTION

The Read Thesaurus (Version 3 of the Read Codes) is a controlled clinical vocabulary produced during the Terms Projects (1992-95),[1,2,3] a series of major collaborations between all the healthcare professions and the National Health Service (NHS) Executive in the United Kingdom. Over 2000 clinicians from all disciplines were involved and, of more than fifty specialist clinical working groups set up at the start of the Terms Projects, fifteen have contributed to the operative procedures section, representing a broad range of general and specialist surgical practice[4].

The project management was undertaken by the NHS Centre for Coding and Classification (NHS CCC) which now has continuing responsibility for the maintenance and further development of the Thesaurus. An attenuated working group structure has been retained to advise the NHS CCC on updates so that the Thesaurus can continue to reflect current clinical practice.

This paper discusses the complete anatomical characterisation of the surgical procedures chapter that has been undertaken and subsequently issued in the Read Code template table.

THE READ THESAURUS

The Read Thesaurus is a comprehensive user-led terminology[5,6] and supports a range of uses by clinicians in primary, secondary and tertiary care. Quarterly updates are released as a series of data tables, one of which, the *template table*[7], includes fields for object, attribute and value triples. This allows both the addition of qualifying detail to concepts and also the representation of intrinsic characteristics.

Procedures hierarchy

The preliminary authoring for surgical operations was undertaken by specialist clinicians. A significant amount of editing was required at the NHS CCC to recast these initial hierarchies into a framework that was able to be applied coherently across the entire domain of clinical practice. A subtype hierarchy with a principally anatomical axis was felt to best fulfil this requirement (Figure 1).

Operations and procedures
Nervous system procedure
Eye procedure
Ear procedure
Respiratory tract procedure
Cardiovascular procedure
Lymph spleen and bone marrow procedures
Oral dental and salivary procedures
Gastrointestinal and digestive procedures
Urinary procedure
Male genital procedure
Female genital and obstetric procedures
Breast procedure
Endocrine procedure
Body wall and cavity procedures
Skin procedure
Musculoskeletal procedure
Fetal procedure
Transplant procedure
Multisystem procedure

Figure 1: First level headings of the operative procedure hierarchy

In addition to this primary anatomical axis, two non-anatomical subdivisions, transplant and fetal

procedures, are present in the upper levels of the operations hierarchy. All concepts are, however, also placed in the anatomy-based branch of the procedures hierarchy as appropriate (Figure 2).

Transplant procedure
Renal transplant
   Cadaver renal transplant
	Kidney operation
	Renal transplant
	   Cadaver renal transplant

Figure 2: Multiple classification of *Renal transplant*

In total, there are over 16 800 unique procedures within the chapter.

Read Version 3 adopts a directed acyclic graph hierarchy structure, independent of the identifying alphanumeric code, to support multiple placement of concepts (so-called multiple parentage) and provide greater flexibility for both refinement of the existing hierarchy and integration of additional concepts. The hierarchy is carried in a separate table of linkages between superordinate and subordinate concepts.

Anatomy chapter
During the Terms Projects an anatomy panel was established of clinicians and anatomists, to oversee the development of a rational, structured anatomical hierarchy. Source material included the existing Version 2 Read Codes, Nomina Anatomica[8], standard anatomical texts and substantial clinical input from specialty working groups.

The resulting anatomy hierarchy, which continues to evolve, now includes over 7600 unique concepts, in a subtype hierarchy. Multiple parentage has allowed both systemic and regional organisation. The high level headings are parallel with those in both the operations and disorders chapters of the Thesaurus.

Template table structure
A further feature of the Version 3 data structure is a template table which supports two important functions. Firstly, it enables pre-determined value sets to be added as qualifiers to object concepts by means of specified object-attribute-value triples. This restricts the options available to a user to an appropriate set and facilitates user interface design. Secondly, the same mechanism may be used to characterise concepts by attributed definitions, referred to as *atomic mapping*. This is essential in ensuring that the terminology supports information retrieval by detection of redundant forms.

Additional control flags within the template table govern the behaviour of the applicable values and warrant further discussion.

Characteristic status. This specifies whether the value is a qualifier (Q) or an atom (A).
Cardinality. This numeric field specifies the maximum allowable number of values for the attribute.
Semantic status. This applies constraints to the manipulations on the stated value that are allowable in an implementation. All values reside within the Thesaurus in hierarchies and the specified values in the template file may, therefore, have subordinate concepts. Additionally, values themselves may have qualifiers such as *right* or *left* for lateralised anatomical structures. The controls are:

- F - no further refinement allowed
- Q - the qualifiers of the value may be accessed
- C - the hierarchical descendants (children) of the stated value may be accessed
- R - descendants or qualifiers may be accessed
- M - if the stated value is used as a qualifier, it is mandatory to refine it - usually because it is an organisational concept. This is illustrated by laterality qualifiers that are specified using the concept *side* which must be refined to its useful subordinates if selected by the user (Figure 3):

Extract from template table for *Breast*			

Object	Attribute	Value	Semantic status
7N12.	XM0Rs	7NB3.	M
Breast	*Laterality*	**SIDE**	*Mandatory*

Side hierarchy

7NB3. | SIDE
  7NB31 | Right
  7NB32 | Left
  7NB33 | Unilateral
  7NB30 | Bilateral

The side hierarchy retains the codes from Read Version 2

Figure 3: Mandatory refinement of values

Others. Other fields allow the specification of the type of qualifiers (Read coded, date, numeric), the order in which attributes and values might be presented on screen and guidance whether the system should display the attribute.

111

SURGICAL PROCEDURE DEFINITION

Several groups have studied the formal semantic descriptions of concepts from surgical procedure vocabularies[9,10,11] and proposed a number of different models. A range of characteristics are identifiable in the definition of a procedure. These include anatomical site or topography (e.g. kidney, thyroid), surgical action or deed (e.g. excision, decompression, implantation), approach (e.g. open, percutaneous), instrumentation and devices, materials, pathological processes, intention (e.g. therapeutic or diagnostic) and extent (e.g. partial or total).

Specification of action and site are usually considered mandatory in the definition of a clinical procedure[10,11] The action must be directed at an anatomical site even though the focus for the action may be a prosthesis or implant, e.g. *removal of stent from bile duct* in which the **target** is the stent and the **site** is the bile duct.

An operative procedure may be conceptualised as a clinical event and an *event definition (ED) model* used to analyse its component concepts based on attributed definitions of concepts.[12,13,14,15] This involves applying to each concept a formal frame of semantically defined *attributes,* and specifying values for each applicable attribute. This template is variously described as an attribute or event-definition frame. The definition frame for the concept *transluminal fragmentation of pulmonary embolus* is shown in table 1:

Table 1: Definition of the object
Transluminal fragmentation of pulmonary embolus

Attribute	Value
Site	Pulmonary artery
Action	Fragmentation
Approach	Transluminal
Pathology	Embolus

ATOMIC MAPPING

The long-term strategy for development of the Read Thesaurus includes atomic mapping of concepts throughout the vocabulary. The reasons for selecting anatomical site as the first characteristic to be systematically defined in the template table are:

- Recognition of anatomical site as essential for the definition of a surgical procedure.
- Selection of a predominantly site-based hierarchy when structuring the procedures chapter.
- In the Thesaurus, the important qualifier *laterality* is always assigned by refinement of the relevant anatomical site value.

Authoring methods
Lexical matching of text strings between the surgical procedure terms and the anatomy chapter proved to be of limited utility for several reasons. Visual inspection of the resulting automated maps proved to be more time-consuming than inheritance of manually authored template entries to subordinate concepts. This was especially true of anatomical values applied to an anatomy-based hierarchy.

In addition, the natural clinical terminology applied to procedures often fails to identify any anatomical structure, e.g. Hartmann's operation, or structures may be ambiguous, e.g. Nissen's fundoplication (a number of structures have fundi including the stomach, gallbladder, uterus and eye). Synonyms may lead to conflicting site assignments for the same concept (*Median nerve* decompression = *Carpal tunnel* release).

For these reasons, atomic mapping was achieved by manual editing of the templates of the operations chapter from the main Read code database. The characterisation of all 16800 surgical procedures in the Read Thesaurus required 1250 different anatomy concepts (16.4% of the 7600 total).

DISCUSSION

A number of issues have been highlighted as a result of this anatomical mapping including the alignment of hierarchies for both procedures and anatomy, the unambiguous definition of procedures involving multiple sites, and the constraints imposed by the Version 3 Read file structure in specifying nested values.

Alignment of hierarchies
The integrity of the procedures hierarchy - a subtype structure based primarily on anatomy - can be tested by observing whether its anatomical atoms accurately reflect the anatomy hierarchy (Figure 4).

```
┌─────────────────────────────────────────────┐
│ Extract from heart procedures hierarchy     │
│                                             │
│ Heart procedure                             │
│   Heart valve operation                     │
│     Mitral valve operation                  │
│       Operation on mitral papillary muscle  │
│                                             │
│ Extract from anatomy hierarchy              │
│                                             │
│ Cardiac structure                           │
│   Cardiac valve structure                   │
│     Mitral valve structure                  │
│       Mitral papillary muscle               │
└─────────────────────────────────────────────┘
```
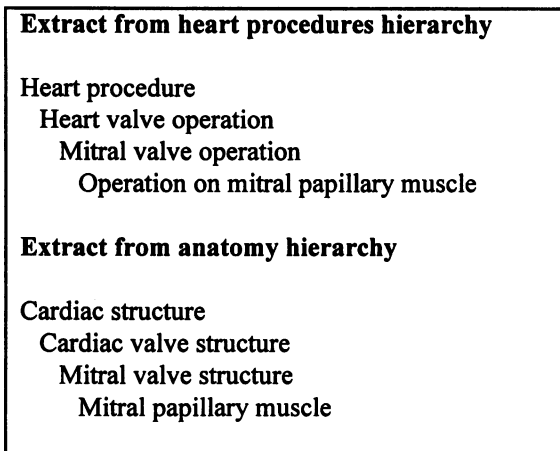
Figure 4: Parallel hierarchies for procedures and anatomy

The mapping exercise enabled hierarchical inconsistencies within both chapters to be identified and corrected.

Expressive capacity of template table
Nested qualifiers. As the value field in the template table is only able to reference enumerated concepts, the decomposition of object concepts in which values need further qualification requires an alternative strategy. This is illustrated by the concept *left salpingectomy* that contains the site atom of *left Fallopian tube*. In the Read Thesaurus, the laterality of symmetrical body structures is generally represented as a qualifier. One solution to this limitation is to instantiate lateralised anatomical concepts such as *left Fallopian tube* for the purpose of atomic mapping lateralised procedures where they occur, though this has generally been avoided:

Fallopian tube
    Left Fallopian tube
    Right Fallopian tube

Left Fallopian tube itself has an atomic map - *Laterality: Left.*

Multiple site elements. Sometimes several specific site attributes may be identified within a procedure. Examples include vascular grafting procedures (e.g. *femoropopliteal bypass graft*) where a donor and recipient site may always be identified. In order to represent these, we have used explicit attributes, subordinate to the attribute *site* in the hierarchy (*donor site, recipient site*).

Template table functions
The specification of permissible qualifiers in the template table affords the Read Thesaurus the functionality of a constrained compositional terminology and confers a manifold increase in expressivity. However, the need to recognise equivalence is crucial when qualifiers added to a superordinate concept may generate an identical construct to an enumerated subordinate, e.g.

Cholecystectomy + laparoscopic
    Laparoscopic cholecystectomy

Atomic definition of the subordinate concept allows this equivalence to be recognised.

The completed atomic map of the Read Thesaurus potentially provides useful functionality in a number of other areas. Formal characterisation of concepts will assist analysis by supporting filtering and classification based upon the atoms. The use of semantic definitions in inter-vocabulary translation has been recognised by other workers[9,12,14] and definition of concept fields for a characteristic may support comparison with other vocabularies via an intermediary reference structure.[16] Definition also allows the vocabulary to be tested for compatibility with emerging standards for concept representation.[11]

Formal definition of clinical concepts is a useful adjunct to the authoring process. The integrity of the subtype hierarchy can be tested and detection of duplicate concepts facilitated. When atomic mapping for other characteristics has been completed, it will be possible to automatically generate alternative hierarchies based on different axial characteristics[17], for example, all procedures with the action *biopsy* could be grouped.

CONCLUSION

The atomic mapping of anatomy for surgical procedures within the Read Thesaurus has enhanced its functionality and provided useful experience in developing a methodology for the planned larger-scale map. Work is progressing on defining other characteristics of procedures within the Read Thesaurus (action, approach) and also in characterising disorders (site, causation, etc.).

References

1. Buckland R. The Language of Health. *BMJ* 1993; **306**: 287-8.

2. Williams JG, Morgan JM, Severs MP, Howlett PJ. Let there be light. *Br J Healthcare Computing* 1993; **10**: 30-32.

3. Severs MP. The Clinical Terms Project. *Bulletin of Royal College of Physicians* (London) 1993; **27**(2): 9-10.

4. Tabaqchali MA, Venables CW. The clinical terms project: its potential for computerised surgical audit. *Ann Roy Coll Surg Eng (Suppl)* 1995; **77**: 124-9.

5. Benson TJR, Neame R. Healthcare Computing. Harlow: Longman, 1994.

6. O'Neil MJ, Payne C, Read JD. Read Codes Version 3: A User Led Terminology. *Meth Inform Med* 1995; **34**: 187-192.

7. NHS Centre for Coding and Classification. *Read Codes File Structure Version 3.1 - The Qualifier Extensions. Technical Report.* Loughborough: NHS CCC, 1994.

8. International Anatomical Nomenclature Committee. Nomina Anatomica. 6th Edition. Edinburgh: Churchill Livingstone, 1989.

9. Cimino JJ, Barnett GO. Automated translation between medical terminologies using semantic definitions. *M.D. Computing* 1990; **7**: 104-9.

10. Burgun A, Botti G, Lukacs B et al. A system that facilitates the orientation within procedure nomenclatures through a semantic approach. *Med Inform* 1994; **19**: 297-310.

11. NV 1828. Medical Informatics - Structure for Classification and Coding of Surgical Procedures. Brussels: CEN, 1995.

12. Masarie FE, Miller RA, Bouhaddou O et al. An Interlingua for electronic interchange of medical information: Using frames to map between clinical vocabularies. *Comp Biomed Res* 1991; **24**: 379-400.

13. Rada R, Ghaoui C, Russell J, Taylor M. Approaches to the construction of a medical informatics glossary and Thesaurus. *Med Inform* 1993; **18**: 69-78.

14. Rocha RA, Rocha BHSC, Huff SM. Automated translation between medical vocabularies using a frame-based interlingua. In: Safran C (Ed). *Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care* New York: McGraw Hill, 1994: 690-4

15. Rothwell DJ, Côté RA, Cordeau JP, Boisvert MA. Developing a Standard Data Structure for Medical Language - The SNOMED Proposal. In: Safran C (Ed). *Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care* New York: McGraw Hill, 1994: 695-9

16. Cimino JJ, Hripcsak G, Johnson SB et al. UMLS as a Knowledge-base - A rule-based expert system approach to controlled medical vocabulary management. In: Miller RA (Ed). *Proceedings of the Fourteenth Annual Symposium on Computer Applications in Medical Care.* New York: IEEE Computer Society Press, 1990; 175-9.

17. Cimino JJ, Clayton PD, Hripcsak G, Johnson SB. Knowledge-based Approaches to the Maintenance of a large Controlled Medical Terminology. *J Am Med Inf Assn* 1994, **1**: 35-50.