

**A Fuzzy Ontology Approach to Coding Clinical Notes: a Web-based
system**

Tsung-Chun Tsai

**A thesis submitted to Auckland University of Technology in fulfilment of
the requirements for the degree of Master of Philosophy (MPhil)**

2014

School of Computing and Mathematical Sciences

Primary Supervisor: Associate Professor Dave Parry

Secondary Supervisor: Professor Nikola Kasabov

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iv
ABSTRACT	v
LIST OF FIGURES	vi
LIST OF TABLES	viii
Chapter 1: Introduction	1
1.1 Background	1
1.2 Problem statement	1
1.3 Research components, goal and question	2
1.4 Paper produced	3
1.5 Remaining chapters	3
Chapter 2: Literature Review	5
2.1 Introduction	5
2.2 SNOMED CT	5
2.2.1 Description	5
2.2.2 Browsers	7
2.3 Fuzzy logic	10
2.4 Fuzzy subsets	12
2.5 Ontology	13
2.6 Fuzzy ontology	13
2.6.1 FO construction	13
2.6.2 FO use	14
2.7 Similarity measurement	17
2.8 Parser	20
2.9 Related approach	22
Chapter 3: Methods	26
3.1 Design of study	26
3.2 Methods for system development	31

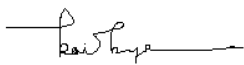
3.3 Selection of experimental techniques	32
Chapter 4: Experiment and Results	33
4.1 Prototype of the experimental system	33
4.2 Experiments	36
4.2.1 Experiment 1 and result	36
4.2.2 Experiment 2 and result	43
4.2.3 Experiment 3 and result	47
4.2.4 Experiment 4 and result	48
Chapter 5: Discussion and Conclusion	51
5.1 Discussion	51
5.2 Further study	52
REFERENCES	57
APPENDICES	61

ATTESTATION OF AUTHORSHIP

"I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning."

Tsung-Chun Tsai

2014

A handwritten signature in black ink, appearing to read 'Tsung-Chun Tsai', with a horizontal line extending to the right.

ACKNOWLEDGEMENTS

I wish to express my gratitude to my supervisors, Dr. Dave Parry and Prof. Nikola Kasabov, for their patience, support, guidance and advice throughout this research.

I would like to thank my wife Linda, my daughter Jasmine, my mother and family in Taiwan, and my parents-in-law in China for their support, patience and love in the process of completing this work.

Sincere thanks are also extended to helpful friends, participants and all staff of the School of Computing and Mathematical Science for their kindness and encouragement during my studying period at the Auckland University of Technology.

Finally, I would like to dedicate my thesis to my father who passed away in 2009. His everlasting spirit has guided me through my difficult times.

ABSTRACT

In clinical research, knowledge and data are often recorded in free-text format which is difficult to access reliably because the variety of expression is vast. The most comprehensive clinical vocabulary currently available is SNOMED CT (**S**ystematized **N**omenclature of **M**edicine-**C**linical **T**erms) which has become widely used in the clinical field. However, a critical issue that accompanies this extensive clinical terminology; an accurate complete set of concepts for classification purposes is difficult to obtain from the provided clinical information. In addition, clinical notes are provided in free-text format, which may be very difficult to match directly to SNOMED CT concepts.

This thesis proposes a novel fuzzy ontology approach that codes clinical notes for SNOMED CT. The aim is to realise effective mapping of clinical notes to SNOMED CT. The fuzzy ontology methodology creates a fuzzy subset that reduces the ontology size and then combined with the fuzzy ontology approach. To render the fuzzy subset of SNOMED CT suitable for professional clinical use, the study incorporates several other tools and methods. The clinical notes are processed by a parser to reduce their length and thereby increase the accuracy of the related concepts found in SNOMED CT. To improve the quality of the fuzzy subset, the similarity is checked by calculating the Levenshtein distance. A future area of the study in crowdsourcing approach which has a potential to increase the accuracy and the usability of the fuzzy subset is also proposed and discussed in the end of this thesis.

LIST OF FIGURES

	Page
Figure 1: Concepts in the SNOMED CT containing the term “Femur Fracture”	2
Figure 2: Information flow from text sources to concepts selected by SNOMED CT	6
Figure 3: CliniClueXplore	7
Figure 4: NPEx online browser Website	8
Figure 5: The interface of SNOMED CT Lite (Screenshot, n.d.).....	10
Figure 6: Visual interpretation of fuzzy logic	11
Figure 7: The four principal facets of fuzzy logic (Zadeh, 2008a)	12
Figure 8: Construction of fuzzy ontology for specific domains as conceptualised by Al shayji & El Kadhi (reproduced from Al shayji & El Zant El Kadhi, 2011)	14
Figure 9: Structure of T2FS-based IDRA system based on T2FO (Lee, Wang & Hagraas, 2010)	15
Figure 10: The structure of a fuzzy ontology on liver disease (Sun et al., 2010).....	16
Figure 11: An example of the Smith-Waterman algorithm (Durbin et al., 1998).....	19
Figure 12a: Output of the tool (categorization mode) (Ruch et al., 2008)	23
Figure 12b: Output of the tool (browsing mode) (Ruch et al., 2008).....	24
Figure 13: The Concepts, Qualifiers and Negations Identified From the Sample Note (Patrick et al., 2010)	25
Figure 14: Design of science research guidelines (Hevner, March, Park & Ram, 2004)	27
Figure 15: Outline of the operational system	28
Figure 16: The method of the parser	28
Figure 17: The procedure of the proposed solution	29
Figure 18: The procedure of the training section	30
Figure 19: Temporal refinement of membership.....	31
Figure 20: Front page of the system	33
Figure 21: Example of an entered testing clinical note and its output	34
Figure 22: Example of a new page for a potential concept.....	34
Figure 23: ERD of the three main tables of SNOMED CT	35
Figure 24: Flowchart of experiment 1	37

Figure 25: Results returned by Stanford NLP Parser to the sentence “Active fetus with normal growth and liquor volume”	38
Figure 26: Plotting of table – Zipf’s law	40
Figure 27: Example of keyword and phrases in a clinical report	41
Figure 28: Flowchart of experiment 2	43
Figure 29: Flowchart of experiment 3	47
Figure 30: Flowchart of experiment 4	48
Figure 31: Example of a task implemented in MTurk (Alonso et al., 2008)	54
Figure 32: An example of a reCAPTCHA test (CAPTCHA example, see a live demo of a CAPTCHA, n.d.)	55

LIST OF TABLES

	Page
Table 1: Example of DP MRI data	36
Table 2: The definition of “Keyword” and “Related concept”	36
Table 3: Sample of parsing result of the DP MRI data	38
Table 4: Assorted keywords with the numbers of related concept found in SNOMED CT	39
Table 5: Result of sorting keywords by concept count	40
Table 6a: Result of searching for phrases in SNOMED CT	42
Table 6b: Result of searching for phrases in SNOMED CT	42
Table 7: Subset of the parsing result of reports 401 – 500 archived in the DP MRI data set.....	44
Table 8: Examples of keywords with the numbers of related concepts	44
Table 9: Result of searching for (keyword + left neighbouring word in SNOMED CT).....	46
Table 10: Result of searching for (keyword + right neighbouring word in SNOMED CT).....	46
Table 11: Result of the “Support” value for each found diagnosis concepts	50

Chapter 1: Introduction

1.1 Background

In clinical research, knowledge and data are often recorded in free-text format. Currently, the textual data are extensively available in electronic format (Shortliffe & Cimino, 2006). Furthermore, these data are available as narrative text thereby providing a natural description. However, narrative text is not really decipherable by the computer during the information retrieval process. In the words of Shortliffe & Cimino (2006) “narrative text is difficult to access reliably because the variety of expression is vast”. An obvious solution is to manually retrieve the information, but this approach is prohibitively time consuming and costly. A better solution is to employ automated systems which can process enormous data quantities in a short time, and which are based on rules imposed by experts, who incorporate the best and most recent knowledge into the rules. Moreover, such systems are generally more objective and consistent than humans (Shortliffe & Cimino, 2006). However, if the system is used to help manage and improve the quality of clinical care and to facilitate research, its performance should not be significantly worse than that of medical experts. To be of use in clinical applications, the system must achieve an adequately high level of accuracy, sensitivity and (Shortliffe & Cimino, 2006).

1.2 Problem statement

As the most comprehensive clinical vocabulary current available, SNOMED CT (**S**ystematized **N**omenclature of **M**edicine-**C**linical **T**erms) is increasingly being adopted as in standardising medical information, and has also been adopted as New Zealand standard. The latest release of SNOMED CT includes more than 300,000 concepts linked through over one million relationships. However, because of the very extensive range and rapid evolution of these concepts, an efficient means of extracting relevant data from the entire set of concepts is not easily achieved.

Many clinical reports and notes are provided in free text form. Because free text is often ambiguous and rarely matches directly to SNOMED CT concepts, it is not really matched to related information from SNOMED CT. Therefore, standard diagnoses or observation codes should be integrated into a clinical decision support system. However, since each term in the free text can potentially code to multiple SNOMED CT concepts, the text may not be meaningful to an interpretation system. Even if the meaning is clear, the “BEST” code is not necessarily selected. For instance, in Figure 1, the SNOMED CT concepts all contain the medical term “Femur Fracture” in their description section. Because “Femur Fracture” is a specific medical term, some of the identified concepts may be outside of the relevant medical domain.

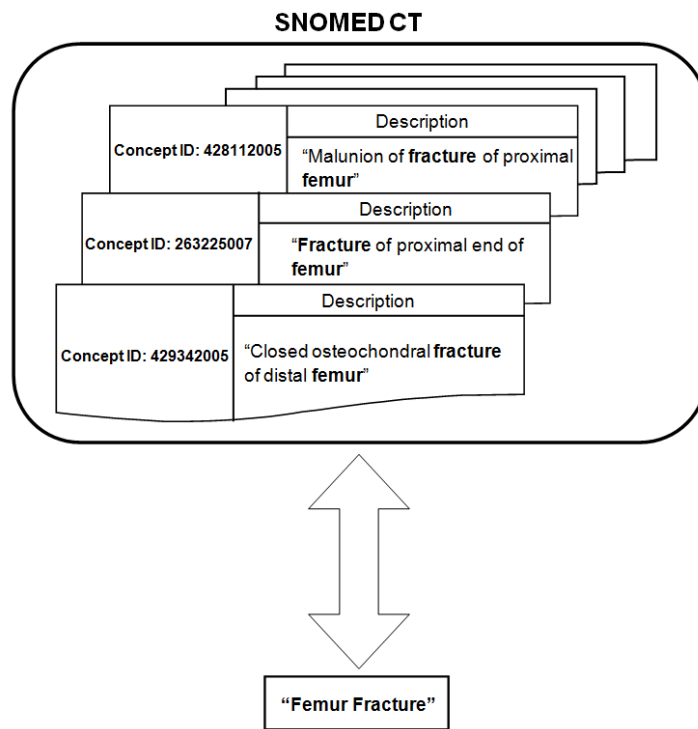


Figure 1: Concepts in the SNOMED CT containing the term “Femur Fracture”

Chiang et al. (2006) showed that clinicians using SNOMED CT browsers to code clinical document perform poorly, with very large inter-coder variances. Patrick et al. (2007) developed automated techniques for coding, but subset selection is often laborious, with high potentials for ambiguity. Therefore, this study focuses on implementing a system that employed fuzzy logic techniques. The goal of this system is to automatically extract and code medical text data to SNOMED CT in a novel manner.

1.3 Research components, goal and question

This study consists of four main parts. First the original document is parsed for potential keywords and key phrases. In the second part, each of the recognised keywords is then mapped to potential concepts from SNOMED CT. Third is generating the extent to which each keyword belongs to its corresponding concepts. Finally likely concepts are identified from the degree of membership of the fuzzy subset. The overall goal of this study is to implement a system that automatically code, and creates sub-ontology for optimal use of SNOMED CT. More specifically, the research identifies answers to the following questions:

- How can parsers reduce the number of SNOMED concepts that relate to clinical notes?
- How do phrases and single keywords extracted from clinical notes differ in their capacity to identify related SNOMED concepts?
- Can fuzzy subset be derived from examples?

1.4 Paper produced

The paper is “Crowdsourcing techniques to create a fuzzy subset of SNOMED CT for semantic tagging of medical documents” was published in Soft Computing Volume 16, Number 7, July 2012(Parry & Tsai, 2012). The four main sections in this paper are: *knowledge representation in healthcare, fuzzy subsets of ontologies, crowdsourcing and the semantic web, and system description*. Each section is summarised below:

- **Knowledge representation in healthcare**

The first section discusses on clinical vocabularies and coding. Clinical vocabularies are used to resolve communication of complex information in the healthcare professions while coding schemes are imposed to interpret diagnoses, procedures, laboratory results and other healthcare related issues. The paper emphasises that the coding schemes may not be compatible with the workflow of the hospital or clinic. It then provides examples of the use of the clinical information coding system and reveals that some coding systems require relatively complex and specific procedures to map clinical notes to generalized sources. This section concludes with examples of SNOMED CT in clinical use and associated problems in particular when selecting subsets of concepts, the edge of a domain that is which concepts should be included and excluded is not clearly defined.

- **Fuzzy subsets of ontologies**

This section conceptualises the fuzzy subsets of ontology with emphasis on subsets, fuzzy logic and fuzzy ontology. Next the approaches used to create the fuzzy subsets of ontologies are discussed. Document-based approaches and Expert- Based Approaches are described.

- **Crowdsourcing and the semantic web**

The concepts of the crowdsourcing and semantic web are introduced in this section followed by the crowdsourcing use. Next the learning algorithm that identifies the degree of membership of concepts in the subset is described. This algorithm is based on a simple algorithm that updates the membership function according to the response of the users in the crowdsourcing.

- **System description**

This section introduces a system that creates the fuzzy subsets and can be adopted in crowdsourcing. The system combines data handling methods, user interface and semantic tagging of documents.

1.5 Remaining chapters

This thesis consists of the paper produced outlined above, and five further chapters after this section in this paper. Chapter 2 provides a literature overview of SNOMED CT, Fuzzy logic, Fuzzy subsets, Ontology, Fuzzy ontology, and Similarity Measurement. Chapter 3 focuses on the methodology including the study and system development. Experiments and results are presented in Chapter 4. First, the system prototype is described, followed by description and

results of three experiments. Chapter 5 presents a discussion and conclusion, and identifies avenues for future study.

Chapter 2: Literature Review

2.1 Introduction

This chapter is the literature reviews on all the methods, the related approach and the techniques that been used and considered in this study. The methods and the approaches are Fuzzy logic, Fuzzy subset, Ontology and Fuzzy Ontology. The techniques been used as tools in this study to achieve and evaluate these approaches are similarity measurement and parser. Because human health is a topic of prime importance, healthcare has become a vital area of ontology and semantic web research. The medical ontology database SNOMED CT contains a vast number of clinical concepts, from which users must select appropriate concept or terms. This is a difficult task because single clinical information is often related to numerous items in SNOMED CT. A popular solution is to create SNOMED CT subsets that simplify the ontology by selecting the relevant concepts. Therefore, in this study, the keywords generated from clinical notes are used and such approach relies on similarity measurement between the keywords and SNOMED CT.

Clinical vocabulary handling is presented in the section one of the literature review. SNOMED CT introduces the clinical vocabulary used in this study, and the benefits it has brought to clinical practice. Sections two introduce fuzzy logic with examples of their clinical use which is employed in this study to improve the usability and interpretation accuracy of SNOMED CT. The ontology and examples of fuzzy logic in the ontology field are presented and described in section four. Because similarity measurement is used to improve information retrieval between clinical notes and SNOMED CT, common similarity measurement methods are presented next. The first and most important step in similarity measurement is to analyse the sentence in the two compared documents, and output the optimal similarity result. Here we adopt a parser to analyse the sentences for improving the keyword-generated results. Literature related to parsers is reviewed to identify the most appropriate parser for the study.

2.2 SNOMED CT

2.2.1 Description

The clinical vocabulary used in this study is SNOMED CT, a standardised clinical terminology providing extensive coverage of diseases, clinical findings, therapies, procedures and outcomes. Since its inception in 1965, SNOMED CT has been developed from a pathology-specific nomenclature into a logic-based healthcare terminology (Cornet & de Keizer, 2008). It is available in US and UK English, Spanish, Danish and Swedish, and is currently being translated into French, Lithuanian, and several other languages. To date, more than fifty countries have adopted SNOMED CT in their electronic health record systems. SNOMED CT allows cross-mapping to other clinical international standards that contain health-

related classifications and coding schemes (About SNOMED CT, n.d). Figure 2 shows the information flow from text sources through SNOMED CT.

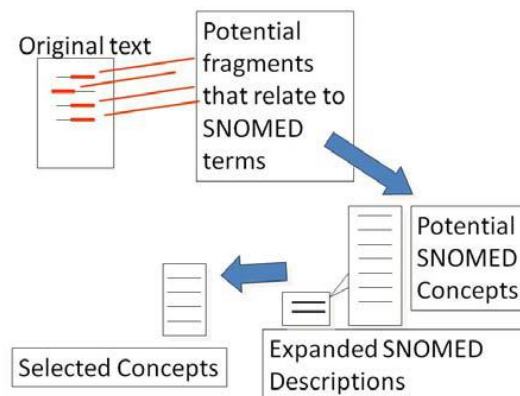


Figure 2: Information flow from text sources to concepts selected by SNOMED CT

SNOMED CT is designed and used for interpreting clinical data for electronic clinical decision support, disease screening and patient safety enhancement (Ceusters et al., 2007). SNOMED CT has provided extremely beneficial to healthcare professions. The key benefits are listed below (International Health Terminology Standards Development Organisation, n.d.):

Clinical providers

- SNOMED CT provides a standardised coding of clinical concepts.
- Important up-to-date data is available for all patients at the time of care.
- Accurate and comprehensive data obtained from SNOMED CT can improve diagnoses and treatments for individual patient.
- SNOMED CT is compatible with other tools (software), enabling patients' progress to be tracked electronically rather than by less efficient paper trail.

Clinical administrators

- SNOMED CT decreases transcription errors and improves data quality by avoiding repetitive data entry, which reduces waste of human resources.
- Access to accurate real-time clinical data improves the quality of clinical planning.
- SNOMED CT enables a cost-effective standardised approach.
- Provides a comprehensive database for research, service planning and identifying clinical indicators.

Patients

- Patients can access improved care-delivery at the place of the service because standardised patient clinical data.
- All practitioners can access health care providers can update patient clinical data.

- Patient data and records are more accurate and complete, and patient safety is improved.
- All clinicians can access an exhaustive evidence-based knowledge and decision support database enabling more accurate diagnosis and more appropriate treatment.
- In information accuracy and privacy are enhanced by reducing the exchanges of hard-copy data.

2.2.2 Browsers

To search for clinical terms in SNOMED CT, keywords phrases or concept IDs are input to the SNOMED CT browsers. The search results assist clinical students, researchers and professionals to understand or learn specific clinical information. SNOMED CT browsers may be software based, web-based and may operate from smart phone applications. A number of free to use SNOMED CT browsers are described below:

- **CliniClue Xplore**

First released at the end of 2009, Cliniclue Xplore is regarded as a significant enhancement for those interested in exploring SNOMED CT. CliniClueXplore also contains an online update program which is easier for users to select the updates they want to install and the online update can make CliniClueXplore to have a fully update easier at anytime. The interface of CliniClueXplore is displayed in Figure 3. URL for downloading and details of the software are available at: http://www.cliniclue.com/cliniclue_xplore

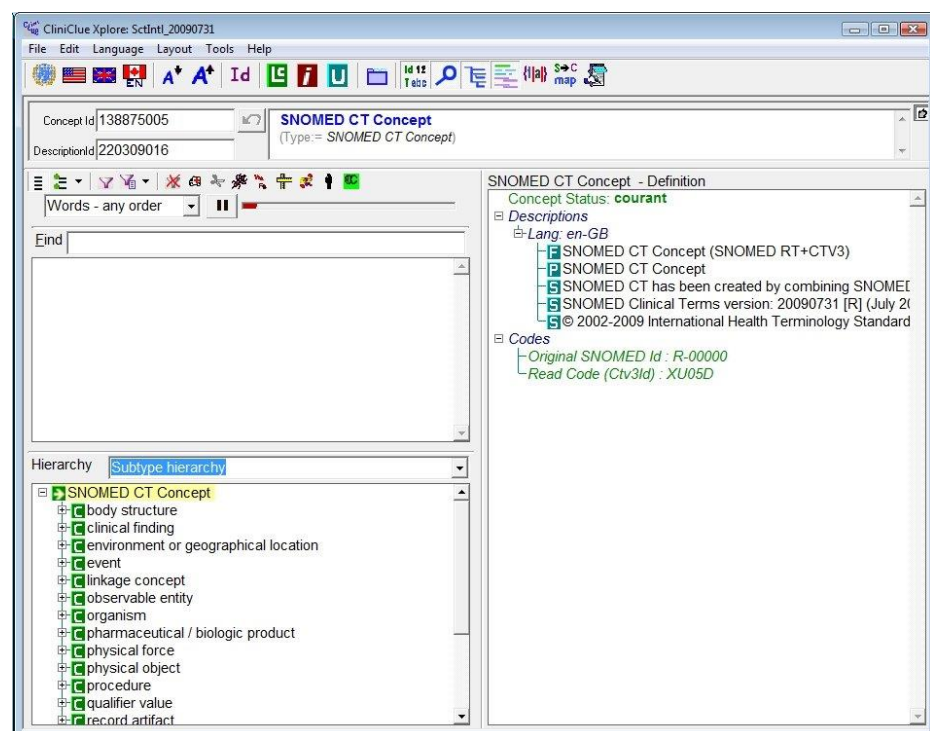


Figure 3: CliniClueXplore

- **SNOB**

Besides its role as a SNOMED CT browser, SNOB allows users to edit existing descriptions, add new descriptions, export the content to file, import files, import multiple releases of SNOMED into the same database, and visualise a changelog for each concept. SNOB also supports a multi-lingual user interface. SNOB uses a lexicon to improve search results and translate descriptions to other languages. URL for downloading and details of this software are provided at: <http://snob.eggbird.eu/>

- **Snapper**

The Snapper Platform software supports the description of terms in existing clinical terminologies using concepts or expressions from SNOMED CT. However Snapper is not exclusive to SNOMED CT, and is compatible with alternate target ontologies. URL for downloading and software details are available at: <http://research.ict.csiro.au/software/snapper>. An abridged version of Snapper, the Minnow edition is available through the URL: <http://research.ict.csiro.au/news/snapper-minnow-edition-now-available>

- **NPEX Browser**

This browser allows users to browse SNOMED CT online. The National Pathology Exchange (NPEX) is a national service for NHS pathology managers in the UK. It connects pathology laboratory systems to a national network, so that laboratories can send and receive test requests and results to/from other laboratories connected to the NPEX browser. Figure 4 shows the Website homepage of the NPEX online browser, whose URL and more details are available at: <http://www.snomedbrowser.com/>

The screenshot displays the NPEX online browser website. At the top, there is a blue navigation bar with links: NPEX, NLMC, THIS, and X-Tab. Below this, on the left, is the NPEX logo and text: "SNOMED-CT Browser", "UKTC SNOMED 15.0.0.0 (April 2013)", "Concept Search", and "About SNOMED-CT". The main content area has a green header with the text "Concept Search" and a description of NPEX. Below this, there are two search sections. The first section, "Search by Name", has a "Search Term:" input field, a "Concept Type:" dropdown menu set to "[All types]", and a "Search" button. The second section, "Search By Code", has a "SNOMED ID:" input field and a "Search" button.

Figure 4: NPEX online browser Website

- **VTSL**

The VTSL online browser designed for searching and browsing of SNOMED CT, was first released in January, 2011 which is maintained by the Veterinary Terminology Services Laboratory (VTSL). VTSL web browser includes history of concept changes, description changes additions, retirements and preference changes. The search results can be set up as searching by “All description”, “Fully Specified Name Only” or “Concept Identifier”. URL and details of this online browser can be obtained from: <http://snomed.vetmed.vt.edu/SCT/menu.cfm>

- **Snoflake**

The online browser Snoflake uses a variety of techniques to find, sort and retrieve codes. It includes a multi parent/child relationship viewer that enables easy traversal of SNOMED CT relationships with support for ICD10 code search. This online free SNOMED CT browser must be registered prior to first use. SnAPI, the application programming interface for SNOMED CT is now available which can be used to enable interaction with other software. URL and details of Snoflake are available at: <http://snomed.dataline.co.uk/snoflake.htm>

The following SNOMED CT browsers are compatible with smartphone:

- **IMO**

IMO Terminology Browser is a free iPhone application available from the App Store. It contains a speech-recognition feature by which users can seek the correct problem and diagnosis codes in clinician-friendly terms by simply speaking into the phone. It allows users to browse capture clinical intent with ICD-9, ICD-10, and SNOMED CT codes. URL of this app in apple store and details of this online browser can be obtained from: <https://itunes.apple.com/us/app/imo-terminology-browser/id425733061?mt=8>

- **SNOMED CT Lite**

SNOMED CT Lite deposits complete, latest SNOMED Codes onto the user's iPhone. The SNOMED codes are instantly searchable and also browsable by their parents. The entire SNOMED codes can be searched by SNOMED code or description. The user simply taps to view the SNOMED code and the full text of its standard long description, and can add the information to his/her favourites list for future immediate access. SNOMED CT Lite, whose interface is shown in Figure 5, is freely downloadable from the App Store which can be found in: <https://itunes.apple.com/app/id448278890>. The full version, SNOMED CT Premium 2011, is also available from App Store for US\$3.99 which the can be obtained from Apple store: <https://itunes.apple.com/us/app/snomed-ct-premium-2011/id448278408>



Figure 5: The interface of SNOMED CT Lite (Screenshot, n.d.)

- **Snomobile**

Snomobile SNOMED Browser is a mobile phone version of Snoflake, and is available on the App Store for iPhone users. The original price was US\$30.99, but now it's reduced to a cost of US\$0.99. The Snomobile browser uses a variety of techniques to find, sort and retrieve codes. It includes fast find functionality, allowing text and code search of full detail pages. It also offers an auto History feature that allows users to easily re-access their most used SNOMED CT codes and descriptions. Snomobile can be purchased in Apple store from: <https://itunes.apple.com/us/app/snomobile-snomed-browser/id354982762?mt=8>

- **MedTalk™**

MedTalk is a free SNOMED CT application for Android based smartphones. Users can search for an exact SNOMED-CT concept, or for a word that might be part of the concept. They also can search for a related term and then browse the SNOMED CT hierarchy until the exact concept is found. URL for downloading and details of this application are available at: <http://www.pyebrook.com/medtalk%E2%84%A2-snomed-ct%C2%AE-browser/>

2.3 Fuzzy logic

As mentioned previously, SNOMED CT is an extensive clinical terminology that is difficult to use efficiently. Problems are too complicated for a quantitative interpretation may be solvable by fuzzy logic (Mittal et al., 2010). Fuzzy logic is based on fuzzy set theory. A fuzzy set is characterised by a real-valued membership function that specifies the degree of membership of each point in the set. It begins with a set of user supplied human language rules, which it converts to their mathematical equivalents. Therefore, fuzzy logic can handle problems with imprecise and indistinct data (Mittal & Bhatia, 2009), and can potentially reduce the size of a

dataset without loss of articulation. Figure 6 illustrates the fuzzy logic concept by analogy to three temperature terms; cold, warm and hot. The points A and B along the vertical dashed line represent two truth values that might be interpreted as "fairly cold" (A) and "slightly warm" (B) in this analogy.

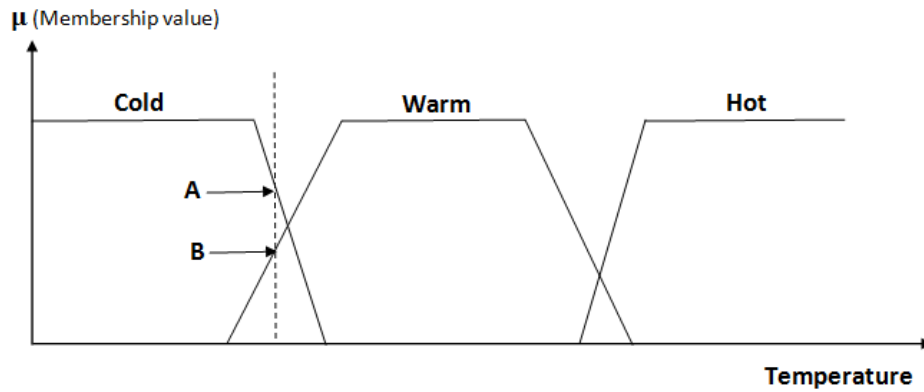


Figure 6: Visual interpretation of fuzzy logic

What is fuzzy logic? Fuzzy in this content does not mean inaccurate or imprecise, rather fuzzy logic is precise in a large measure (Seising, 2207). Fuzzy logic was first proposed in 1965 by Lotfi A Zadeh (Zadeh, 1965). The four principle facets of fuzzy logic, namely logical, fuzzy-set-theoretic, and epistemic and relational are presented in Figure 7 (Zadeh, 2008a). Most practical applications of fuzzy logic are associated with its relational facet. The non-traditional view of fuzzy logic precisiation adopts the precisiation position, described by Zadeh (2008a) as follows: "Informally, precisiation is a procedure the concerts an object p into an object p^* , where p^* is defined more precisely than p ". He also mentions the important role of generalized constraint in fuzzy logic. Possibilistic, probabilistic and veristic are the primary constraints, and bivalent possibilistic, probabilistic and bivalent veristic are the standard constraints. Zadeh (1994) also states that in many cases, the term fuzzy logic is actually used in two different senses; a fact that is generally not recognised. In the first sense, fuzzy logic is an extension of multi-values logic and is intended to serve as a logic of approximate reasoning. In the other sense, fuzzy logic is more or less synonymous with fuzzy set theory, which treats classes with non-sharp boundaries. Consequently, fuzzy logic may be viewed as an attempt to formulate or mechanise two remarkable human capabilities (Zadeh, 2008b). Important features of fuzzy logic are the linguistic variables and fuzzy if-then rules (Zadeh, 2008b). Combined these features realise a powerful modelling language that is widely applied in fuzzy logic problems. In fact, fuzzy logic formulism enables information to be summarised and compressed through the use of granulation. Thus, fuzzy logic imbues bivalent logic with a critical capability from which accurate inferences can be made from deficient information (Zadeh, 2009). In particular, fuzzy sets can provide a semantically sound means of representing granularized information. The use of fuzzy if-then rules in place of classical logic allows for the imprecision needed in real situations (Hong & Langari, 1999). Moreover, by applying fuzzy logic to navigation and use of an ontology

(Widyantoro & Yen, 2001; Parry, 2006), “fuzzy ontologies” can be realised. However, learning the fuzzy relations from documents is a complex task (Tho, Hui, Fong & Cao, 2006) and is unlikely to succeed for complex and specific medical terms. Consequently, when the fuzzy logic technique is applied to remove ambiguities from free text, and thereby efficiently retrieved information from SNOMED CT, the information remains complex because a single phrase or word extracted from free text may map to multiple concepts in SNOMED CT. Accordingly, learning large numbers of fuzzy relations directly from users in an onerous task, and may be rendered unreliable by inter-operator variability. To select the optimal information, fuzzy logic must be combined with other techniques.

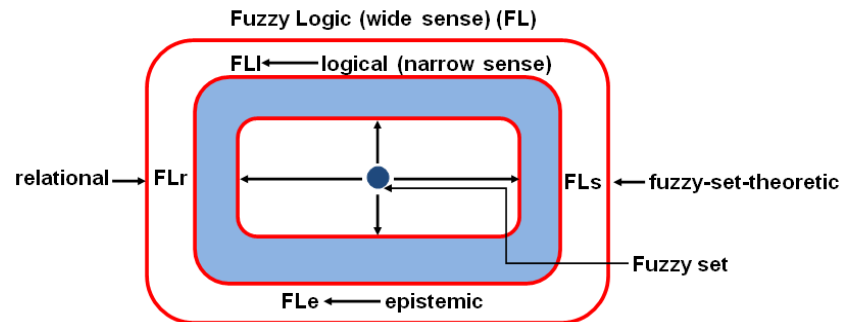


Figure 7: The four principal facets of fuzzy logic (Zadeh, 2008a)

2.4 Fuzzy subsets

In an ontology, subsets may overlap, and users may differ in their degree of belief that a concept should be included in a particular subset. A well-known method of dealing with such problems is the use of fuzzy logic (Zadeh 1965). In this research, a fuzzy subset is an approach for this issue and it is constructed from SNOMED CT. Our primary concern is the application of fuzzy set theory to ontologies. Traditionally, whether element B is an element of a set A is decided by a characteristic function that assigns a value of “1” if B belongs to set A and “0” otherwise. However, in fuzzy set theory, an object can partially belong to a set. The degree of membership is decided by a generalized characteristic function (Kasabov, 1996). Linguistic terms are often used when specifying and identifying a problem, or when articulating heuristic rules (Kasabov, 1996). This assignment of linguistic terms is called conceptualisation. Through conceptualisation, the problem can be represented in fuzzy terms; this is an essential step on fuzzy logic for problem-solving. For example, the linguistic terms “strong”, “very strong”, “slow”, “higher”, “lower”, “much dependent”, “less dependent”, “good”, “bad”, and many others are fuzzy concepts that can be represented as fuzzy sets. Linguistic terms in fuzzy sets can also be regarded as linguistic variables by which to indicate variables that take fuzzy values (also called fuzzy labels) with linguistic meanings. Kasabov (1996) notes that “fuzzy discretization does not lead to loss of information if the fuzzy labels are correctly chosen”. In addition, the fuzzy subsets approach is particularly useful when the aim is to generate a fuzzy subset of an existing ontology, rather than creating from documents (Parry & Tsai, 2012). In this study, we seek to reduce the complexity of the ontology for selecting the “BEST” code. Accordantly, we adopt fuzzy set theory to reduce the size of the ontology and assign a membership value to each

concept of the fuzzy set. Moreover, the clinical notes and concepts archived in SNOMED CT are in linguistic form. Therefore, we can use linguistic variables to assign the membership values, and thereby describe the degree of similarity between a potential keyword in a clinical note and a potential concept from SNOMED CT. Accordingly, when the clinical professionals search the concepts by keywords in a fuzzy subset, they can easily identify the relationship and importance between keywords and concepts by only checking the membership value and also fuzzy subset can be seen as a streamlined ontology to provide users a better use for an ontology. This goal is realisable through building a fuzzy subset for SNOMED CT based on fuzzy set theory. As mentioned above, once the fuzzy labels are correctly chosen, the degree of similarity will be precisely selected by the experts. The fuzzy labels chosen in this study will be described in section 3.1, “design of study”.

2.5 Ontology

Gruber (1993) defines ontology as “an explicit specification of conceptualization”. Ontologies are concerned with concept representation, often in compute readable format and relating concepts in a particular domain. Accordingly, information is not easily retrieved from domain ontology because such ontology can include numerous concepts and relations. Moreover, an ontology information retrieval system usually returns large result sets, from which users must expend much time extracting the relevant items (Leite & Ricarte, 2008). Therefore, several researches and applications have adopted fuzzy logic in the information retrieval systems of different ontology domains.

2.6 Fuzzy ontology

2.6.1 FO construction

The ontology used in this study is the clinical vocabulary SNOMED CT. As previously mentioned, because of the importance placed on human health, healthcare has enthusiastically embraced ontology and semantic web research. In addition, SNOMED CT is the most representative and widely uses ontology in the health domain. However, its effective use is limited by the several issues addressed in the preceding sections. Consequently, ontology in the health domain has attracted criticism as well as acclaim. This study focuses on solutions to that will improve the usability of SNOMED CT.

Fuzzy ontology is an extension of domain ontology that solves uncertainty problems. Classical ontologies are not appropriate for handling fuzzy information, and real-world information is often vague or ambiguous (Lee, Wang and Hagrais, 2010). Therefore, it is necessary to integrate fuzzy logic with ontology into a concept we call fuzzy ontology. Some existing examples are shown as in the next section.

2.6.2 FO use

- Al Shayji and El Zant El Kadhi (2011) propose a fuzzy logic prototype called Ontology architecture for intelligent decision support systems. Ontology architecture is targeted at top political decision-makers. Decision-making is a complex procedure requiring completion of fragmentary elements and imprecise information, and synchronous consideration of multiple objectives. The construction of fuzzy ontology for specific domains is shown in Figure 8. Firstly, the system extracts key concepts and relationships between sets of information, and then applies fuzzy logic to ontology. Thus, the fuzzy ontology includes information on important concepts in each domain. As an extension of domain ontology for solving political uncertainty problems, fuzzy ontology assumes that each index object is related to every other object in the ontology with a degree of membership assigned from fuzzy set theory.

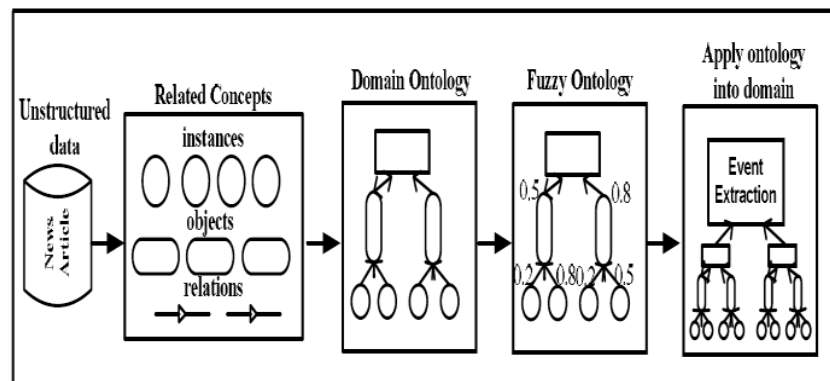


Figure 8: Construction of fuzzy ontology for specific domains, as conceptualised by Al shayji & El Kadhi (reproduced from Al shayji & El Zant El Kadhi, 2011)

- Lee, Wang and Hagraas (2010) developed a fuzzy ontology approach to personal diabetic-diet recommendation. They reported that classical ontology cannot handle the vague and imprecise knowledge inherent in real world applications. However, fuzzy ontology can realise effective processing of uncertain information and knowledge. This application called type-2 fuzzy set based (T2FC-based) intelligent diet-recommendation agent (IDRA), employs type-2 fuzzy logic to create a type-2 fuzzy ontology (T2FO) composed of three specified sub-domains; namely a type-2 fuzzy food ontology (type-2 FFO), a type-2 fuzzy-personal food ontology (type-2 FPFO), and a type-2 fuzzy personal profile ontology (type-2 FPPO). The original concept of fuzzy logic, proposed by Lotfi A. Zadeh in 1965, is based on type-1 fuzzy set theory. A decade later, Zadeh introduced type-2 fuzzy set theory. In this theory, the degree of membership is also fuzzy and the sets are essentially “fuzzy fuzzy” sets (Castillo, Melin, Kacprzyk & Pedrycz,

2007). By contrast, because the membership functions in type-1 fuzzy systems are precise, they remove any uncertainties from the data. In type-2 fuzzy sets, uncertainties in word meaning are modelled using type-2 membership functions. Thus type-2 fuzzy sets are used if uncertainties exist in the membership function; if no uncertainties are present the type-2 fuzzy sets are reducible to type-1 fuzzy sets. An overview of this system structure is shown in Figure 9. First, diet goals are prepared for the participations diabetic patients and the nutritional contents of common Taiwanese foods are compiled by dietary experts. Second, the participating diabetics are requested to regularly to input their food consumption. Third, a T2FO is created by the ontology-creating mechanism, comprising a type-2 FFO, a set of type-2 FPFOs, and a type-2 FPPO. Finally, the T2FS-based IDRA recommends a personal diabetic meal plan based on the information retrieved from the T2FO.

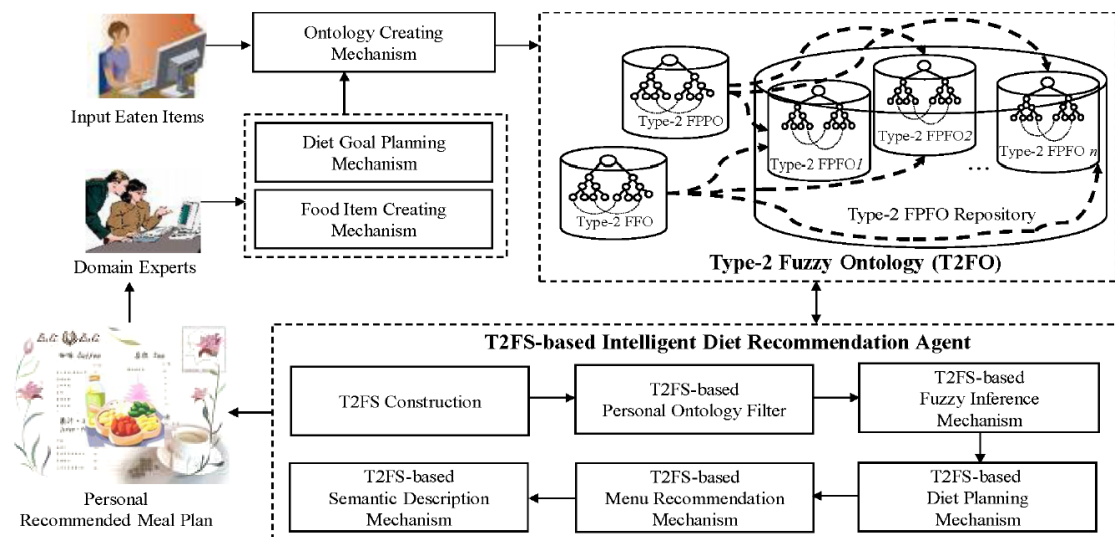


Figure 9: Structure of T2FS-based IDRA system based on T2FO (Lee, Wang & Hagra, 2010)

- Calegari and Ciucci (2007) proposed a framework for constructing a fuzzy ontology for semantic webs based on fuzzy description logic (DL) and Fuzzy Web Ontology Language (OWL). This framework has become one of the metadata standards developed by the W3C. Via the framework, users can annotate the data in a machine comprehensible way online. It also allows extended inferences and enables the data to be automatically integrated over its ontological structure (Fensel et al., 2006). Calegari and Ciucci (2007) noted that semantic web applications are based on ontologies, formal conceptualizations of a domain of interest shared between various applications. In the semantic web, users express their requirements or queries as linguistic adjectives, which generally constitute non-sharp data. Unfortunately, the conceptual formalism of the ontology structure is insufficient for representing or handling this type of information. As mentioned by Calegari and Ciucci (2007), fuzzy sets provide a possible solution to the difficulty of handling imprecise and vague data. This

fuzzy ontology concept is based on the semantic web ontology, which presents the knowledge in a DL-based ontology. The framework of this fuzzy ontology is inspired by the most complete and coherent definition of fuzzy description logic proposed to date introduced by Straccia (2006). Calegari and Ciucci (2007) added fuzzy relations with modifiers to Straccia's model, thereby extending the model to accommodate concept modification. As a result, all of the concepts extracted from the original ontology can be more comprehensively analysed. They also modified the semantic of cardinality restriction in Straccia's work and represent the degree of the relation used by a specified value. Finally, they use a fuzzy language called Fuzzy-OWL to implement the fuzzy ontology. In Fuzzy-OWL, the fuzzy SHOIN (D) syntax is incorporated in the fuzzy ontology definition, and fuzzy value is assigned to the relationships and entities of the ontology.

- Sun et al. (2010) constructed a fuzzy ontology for Traditional Chinese Medicine (TCM) accompanied by identified related problems and known facts for improved diagnosis and treatment. They mentioned standardisation and normalisation of TCM terms as the key to developing and maintaining TCM. The complexity of the TCM ontology impedes the integration, exchanging and sharing of knowledge required for development and maintenance of TCM. Thus, the TCM ontology is a very suitable candidate for Fuzzy Logic. In Sun et al's (2010) method, the fuzzy ontology combines fuzzy sets with fuzzy relations. The fuzzy sets describe the concepts-term of TCM, while the fuzzy relations specify relations between concepts. Thus, the fuzzy ontology processes the uncertain information from TCM and presents it in an effective format that also simplifies the clinical decision-making process. Applying this method, Sun et al. created a fuzzy ontology on liver disease, whose structure is shown in Figure 10. The top level (Level A in the figure) is the concept class (liver disease), the second levels (B, C, D, E) are four objects of the concepts, denoting different types of liver disease. At the bottom of the hierarchy are eleven typical symptoms of the various disease types, these are the attributes of the related objects of the concept. The arrows indicate relationships between the concept, objects and attributes. The words along each arrow describe the relationships between the two connected nodes and the degree of the relationship is specified at the node.

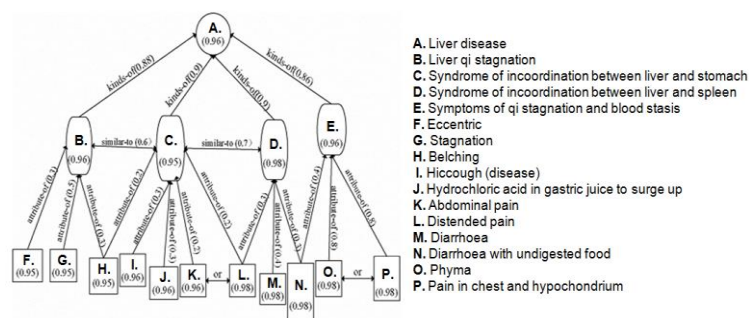


Figure 10: The structure of a fuzzy ontology on liver disease (Sun et al., 2010)

In this study, fuzzy subsets of SNOMED CT are created from fuzzy logic using the algorithms reviewed above. To generate potential fuzzy subsets, concepts must also be retrieved for the clinical notes by parsing. The parsing matching algorithm for similarity measurement between the clinical notes and SNOMED CT data, which also plays a critical role in this study, is reviewed in the section 2.8.

2.7 Similarity Measurement

The concepts in SNOMED CT that are relevant to the clinical notes are easily recognised by clinical experts, but such manual inference incurs unnecessary time and cost. To improve the retrieval of SNOMED CT data from clinical notes, an automated technique, namely, similarity measurement is adopted. Moreover, ontology has been extensively studied in the area of knowledge representation, a branch of computer science. According to Gruber (1993), “A specification of a representational vocabulary for a shard domain of discourse – definitions of classes, relations, functions and other objects – is called ontology.” By definition, ontology contains a comprehensive information and terminology in its specific fields, which is too complicated to retrieve by the technique known as ontology mapping. This study intends to retrieve the concepts from SNOMED CT that match supplied clinical notes. Several ontology mapping approaches and strong matching algorithms have emerged in recent years. Among the ontology mapping approaches are “H-match” (Castano et al., 2003), the string matching algorithms such as “Hamming distance”, defined by Hamming (1950), “Episode matching” (Das et al., 1997), “Smith Waterman Algorithm” (Smith & Waterman, 1981), “Tversky’s similarity” model (Tversky, 1977), “TF-IDF” and “Levenshtein distance”. Each of these methods are described in turn paragraph.

H-match

H-match accepts two ontologies as input and returns the mappings that identify matching concepts in both ontologies, the concept with the same or the closest intended meaning (Castano et al., 2003). H-match calculates a semantic affinity value as the linear combination of a linguistic affinity values and a contextual affinity value. For linguistic affinity evaluation, H-match relies on a lexicon of terms and terminological relationships automatically derived from the WordNet lexical system. To determine the contextual affinity value, H-match calculates the similarity from the contextual features of the ontology concepts. It adopts one of four matching models; surface, shallow, deep and intensive. Surface matching determines concept similarity from the linguistic affinity between the concept names. Shallow, deep and intensive matching also account for contextual affinity to varying extent. In shallow matching, the context comprises concept properties only. The depth of context is increased in deep and intensive matching, which semantic relations incorporate with other concepts and property values, respectively.

Hamming distance

Originally, the Hamming distance was used for detecting and correcting the errors in digital communication, which can be simply defined as the number of bits that differ between two binary vectors (Sankoff & Kruskal, 1983). The Hamming distance between two strings of equal length can be interpreted as the number of positions at which the corresponding symbols differ. Essentially, it measures the minimum number of substitutions required to interchange the two strings, or the number of errors required to transform one string into the other.

Episode matching

Episode matching identifies all minimal-length sub-strings of text A that contain episode (pattern) B as a subsequence, where two words contained in text A are of length c and episode B is of length d (Das et al., 1997). Episode matching identifies the smallest k, where k is the length of a sub-word m contained in both text A and episode B. By “minimal” is meant that no apposite sub-string of m contains B as a subsequence.

Longest common subsequence

A subsequence is a set of characters that appears in left-to-right order, but not necessarily in consecutive order in a string. A subsequence appearing in both strings is called a common subsequence. The longest common subsequence of the two strings is called the maximal length common subsequence. For example, given two strings A: “AAACCGTGAGTTATTCGTTCTA” and B: “CACCTAAGGTACCTTTGTTC”, the longest common subsequence of A and B is defined as follows:

A: AAACCGTGAGTTATTCGTTCTA

B: CACCTAAGGTACCTTTGTTC

Longest common subsequence: “ACCTAGTACTTTG”

Smith-Waterman algorithm

The Smith-Waterman algorithm, introduced by Temple F. Smith and Michael S. Waterman in 1981, a dynamic programming algorithm for that determines sequence alignment from the similarity measures of two strings. The algorithm is most commonly used in DNA or protein algorithms (Smith & Waterman, 1981). The algorithm comprises two steps:

- Calculating the score of the similarity matrix; this step is the most computationally time-intensive step.
- Seeking the optimal alignment by tracing back the similarity matrix.

An example of the Smith-Waterman algorithm is illustrated in Figure 11 (Durbin, Eddy, Krogh & Mitchison, 1998). The algorithm operates on string A: “HEAGAWGHEE” and string B: “FAWHEAE”. The Smith -Waterman algorithm generate the following result:

A: "HEAGAWGHEE"

B: "FAWHEAE"

AWGHE

AW – HE

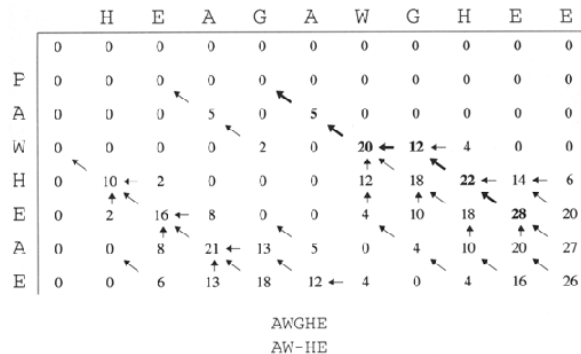


Figure 11: An example of the Smith-Waterman algorithm (Durbin et al., 1998)

Tversky's similarity

Tversky's similarity computes the similarity value between two objects A and B, based on their common and dissimilar characters (Tversky, 1977). It first recognizes the characters that are common to both A and B (symbolized "A and B"), those that exist in A but not in B (symbolized "A-B") and those existing in B but not in A (symbolized "B-A"). These results are then input to the equation $SIM_{AB} = \text{COMMON} / (\text{COMMON} + \text{DIFFERENT})$: $SIM_{AB} = \text{"A and B"} / (\text{"A and B"} + \text{"A-B"} + \text{"B-A"})$, which outputs the final similarity measure (Blough, 2001).

TF-IDF and Levenshtein distance

Similarity measures derived from string matching are frequently derived from "Term Frequency - Inverse Document Frequency" (TF-IDF) (Salton and Buckley, 1988). Frequently used in the text mining and information retrieval, TF-IDF measures the importance of a word in the document body. The importance of a word is assumed proportional to the frequency of the word in the document, but very frequent words are regarded as less important (Salton & Buckley, 1988). Edit distance also known as the Levenshtein distance, is compares comparing string similarity at the character level. The Levenshtein distance is named after the Russian scientist Vladimir Levenshtein, who devised the algorithm in 1965 (Shen, Radev, Patel, & Erkan, 2006). It uses insertions, deletions, or substitutions to identify the minimum cost of transforming one string into another (Wang, Lu & Zhang, 2008). The minimum transformation cost clearly relates to the similarity between two strings. For example:

- If both strings A and B contain the characters "string", the Levenshtein distance (A, B) = 0, because no transformations are needed. The strings are identical.
- If strings A and B contain the characters "string" and "strong", respectively, the Levenshtein distance (A, B) = 1, because a single substitution ("l" to "o") is sufficient to convert A to B.

In this study, the keywords of clinical notes are matched with the concepts of SNOMED CT by the similarity measurement between two strings. Two ontologies are typically compared by H-matching, which is not applied in this study, since the Hamming distance is applicable only to equal length strings. The algorithms of Episode matching, Longest common subsequence and Smith-Waterman can all identify the subsequence between two strings. However, because the matching string does not need to be consecutive, these algorithms may be inefficient and unsuitable for our purpose. The TF-IDF algorithm measures the importance of a word in the document corpus, which is not easily defined in this study because each concept is highly individual and the importance of each word cannot be calculated. Thus, the TF-IDF is similarity unsuitable for this study. In the algorithms of Levenshtein distance and Tversky's similarity, the similarity between two strings is computed from both their common and dissimilar characters. Moreover, the Levenshtein distance uses insertions, deletions, and substitutions to calculate the minimum cost of transforming one string into another enabling a more accurate similarity measure. Therefore, we adopt the Levenshtein distance as the similarity measurement in this study.

2.8 Parser

Natural language parsers are employed over numerous domains and have recently become more accurate and more widely available. Currently, they are used in such diverse domains as question answering, speech recognition, language modelling, language generation and machine translation (Ravi, Knight & Soricut, 2008). The most important question in parser selection is the accuracy of the parser on the domain of interest. To achieve the required accuracy in a parser, human annotators build parse trees for a sample set from the domain of interest, from which they compute a PARSEVAL score that indicates the inherent performance of the parser (Sampson & Babarczy, 2003). Another approach currently popular for large-scale data is to measure the performance of a parser on existing parse treebanks and presume that the accuracy measure will similarly apply to the domains of interest. This approach is inexpensive and simple, but offers no guarantee of parser performance on a new domain. It may also provide erroneously indicate the important decisions for design the of NLP systems that rely on syntactic parsers. Several common parsers that suitably avoid the accuracy problem are introduced below. Parser selection is one of the important aspects of the current study. If parser performs insufficiently well to achieve the study goals, it will generate ineffectual keywords, resulting in retrieval of incorrect and unrelated concepts. The first priority of parser selection in this study is cost. The licenses of the parsers in the following parser list can all be obtained free-of-charge for non-commercial research and instruction. The parsers are candidate, such as Stanford NLP Parser, HAPI, PET, RASP System, LoPar, Charniak Parser, Enju, ERG, Heart of Gold, OpenNLP and CorpusSearch 2. Each of these parsers has its own limitation. After weighing up the limitations and the capabilities, Stanford NLP Parser was selected for use in this study.

- The Stanford NLP Parser is a Java implementation of a probabilistic PCFG (probabilistic context-free grammar) and a dependency parser for English, German, Chinese, and Arabic.

Limitation: long loading time.

- HAPI is the abbreviated form of “**HL7 Application Programming Interface**”, an open-source, object-oriented HL7 2.x parser for Java. HL7 is a messaging specification for healthcare information systems.

Limitation: HAPI uses its own standard, which limits the available data format.

- PET is an open-source runtime system for unification based grammar processing. PET is a core component of the Deep Linguistic Processing with HPSG Initiative (DELPH-IN) open-source repository.

Limitation: It is difficult to set up and use.

- The RASP (“**R**obust **A**ccurate **S**tatistical **P**arsing”) system is a domain-independent, robust parsing system for English text. RASP is implemented as a series of modules written in C and Common Lisp and runs on UNIX (Briscoe, Carroll, & Watson, 2006).

Limitation: It can be deployed only in Linux and Macintosh/OSX.

- LoPar parses head-lexicalised probabilistic context-free grammars, and is suitable for symbolic parsing with a context-free grammar (CFG), statistical parsing with PCFGs and head-lexicalised PCFGs, training of PCFGs and HLPCFGs, Viterbi parsing, chunking, tagging and more.

Limitation: Runs only on Sun/Solaris or Intel/Linux platforms.

- Charniak Parser is a syntactic parser of natural languages (e.g. English), developed at the Brown Laboratory for Linguistic Information Processing (BLLIP). Charniak Parser is implemented in C++ and runs on UNIX.

Limitation: It returns some poor parser results.

- Enju is a syntactic parser of English text. Its benefits are a wide-coverage probabilistic HPSG grammar and an efficient parsing algorithm. Enju effectively analyses syntactic/semantic structures of English sentences and provides the user with phrase structures and predicate-argument structures.

Limitation: It operates more slowly than other high performance parsers.

- LinGO English Resource Grammar (ERG) is a broad-coverage, linguistically precise HPSG-based English grammar repository. Initially developed using the DFKI PAGE system, LKB is now the primary grammar engineering environment of ERG. ERG is semantically grounded in Minimal Recursion Semantics (MRS), a form of flat semantic representation that supports under specification.

Limitation: Some words are not found in the lexicon.

- Heart of Gold is a middleware architecture that integrates deep and shallow natural language processing components, thereby providing a uniform and flexible infrastructure for building applications that use Robust Minimal Recursion Semantics (RMRS) and/or general XML standoff annotation.

Limitation: Not really implemented in Microsoft Server.

- OpenNLP is a machine learning based toolkit for processing natural language text. It supports the most common NLP tasks, such as tokenization, sentence segmentation, part-of-speech tagging, named entity extraction, chunking, parsing, and coreference resolution.

Limitation: Full parsing has yet to be fully tested.

- CorpusSearch 2 is a Java program that supports research in corpus linguistics. By running CorpusSearch 2, the users can automatically find and count lexical and syntactic configurations of any complexity, correct systematic errors, and code the linguistic features of corpus sentences for later statistical analysis.

Limitation: Corpus compatibility is restricted.

2.9 Related approach

The literature review on the related approach will be conducted in this section – the automatic medical encoding system for SNOMED CT.

2.9.1 Automatic medical encoding for SNOMED CT

The latest release of SNOMED CT includes more than 300,000 concepts linked through over one million relationships. Because of the very extensive range and rapid evolution of these concepts, an efficient means of extracting relevant data from the entire set of concepts is not easily achieved. The automatic conversion of free text into a medical ontology seems like one of the optimal solutions which also can allow computational access to essential information currently locked within patient reports and medical notes (Patrick et al. 2007). Furthermore, a number of clinical applications which includes clinical error detection, medical decision support, clinical queries

answering/response, patient statistics generating, and bio-surveillance that would be automated if the clinical information locked in natural language clinical text could be converted into computer-amenable structured representations. To facilitate this, a long term goal is to convert entire natural language clinical documents into structured representations (Kate, 2013). Two existing examples of the automatic medical encoding system are discussed next.

- A new type of categorization and browsing tools (Figure 12a shows the output of the tool in categorization mode and Figure 12b shows the output of the tool in browsing mode) for SNOMED encoding that proposed and developed by Ruch, Gobeill, Lovis, and Geissbühler (2008) which can be used either as a search tool to browse the terminology or as a categorization tool to support automatic annotation of textual contents with SNOMED concepts. The system combines a pattern matcher, based on regular expressions of terms, and a vector space retrieval engine which uses stems as indexing terms that is a traditional tf.idf weighting schema, and with cosine as the normalization factor. To evaluate the system, they use a sample of MEDLINE which SNOMED CT categories have been restricted to Medical Subject Headings (MeSH) using the SNOMED-MeSH mapping provided by the UMLS. The results show that performances of this system are significantly improved to help coding textual contents as contrast to each single module - a precision at high ranks (P0) of more than 80% is observed.



Figure 12a: Output of the tool (categorization mode) (Ruch et al., 2008)

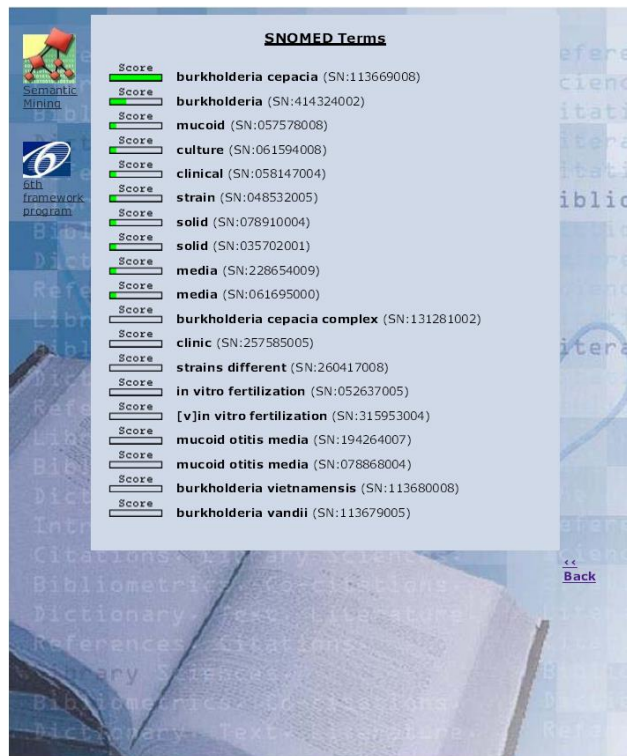


Figure 12b: Output of the tool (browsing mode) (Ruch et al., 2008)

- A system to identify medical concepts automatically from the SNOMED CT in free text in near real time that is proposed by Patrick et al. (2007). It is to code free text clinical notes to clinical terminology and create a web-service system to deliver the services. It is currently in use in a hospital environment to capture patient data response with SNOMED-CT codes in real time at the point of care. The system employed NLP methods to increase lexical concept mappings. A qualifier identifier and negation identifier (Figure 13 shows concepts, qualifiers and negations identified from the sample note from this system) have been used for recognising complex terms and negative concepts that can then create more effective information retrieval and information extraction.

A Sample Clinical Note

No neoplasm malignant^{negation} seen.

Sections confirm CRANIOPHARYNGIOMA^{concept} with small^{qualifier} fragments^{qualifier} of adjacent^{qualifier} brain tissue^{concept}.

The slides^{concept} show degenerate atypical^{qualifier} urothelial cells^{concept} occurring in sheets^{qualifier} and singly with hyperchromatic^{qualifier} enlarged^{qualifier} irregular^{qualifier} nuclei^{concept}.

Concepts, Qualifiers and Negations Identified From the Sample Note

SNOMED CT Concept	SCT Concept ID	SCT Fully Specified Name	
CRANIOPHARYNGIOMA	40009002	Craniopharyngioma (morphologic abnormality)	
	189179009	Craniopharyngioma (disorder)	
Brain tissue	256865009	Brain tissue (substance)	
Cells	4421005	Cell structure (cell structure)	
	362837007	Entire cell (cell)	
hyperchromatic	9767008	Hyperchromatism (morphologic abnormality)	
Qualifiers	SCT Concept ID	SCT Fully Specified Name	Scope of Qualification
Small	255507004	Small (qualifier value)	
	263796003	Lesser (qualifier value)	
Fragments	29140007	Fragment of (qualifier value)	
Adjacent	18769003	Juxta-posed (qualifier value)	brain tissue
Atypical	112231000	Atypical (qualifier value)	Cells
Sheets	255292000	Sheets (qualifier value)	
Enlarged	260376009	Enlarged (qualifier value)	
Irregular	49608001	Irregular (qualifier value)	
Fragments	29140007	Fragment of (qualifier value)	Tissue
Negation	Negation Phrase	Negative Term	
no neoplasm malignant	No	neoplasm malignant (86049000)	

Figure 13: Concepts, Qualifiers and Negations Identified From the Sample Note (Patrick et al., 2010)

Chapter 3: Methods

This chapter discussed the methodologies that been used to design and construct this study which are design science and experimental study. The procedure on how to conduct the experimental study for this study is also been described in experimental study section. The methods used to build system development and the techniques selected for experimental study method are discussed in the end of this chapter.

3.1 Design of Study

This section overviews the study design and execution. The two methodologies employed in this thesis are design science and experimental study. The core methodology is the design science, which contains two fundamental components, creating and evaluation the artefact. The artefact is created from a system equipped with a website interface, while the experimental study evaluates the artefact through a series of tests that generate experimental results. The concepts of these two methodologies are described here:

Design science

In the design science paradigm, the understanding and knowledge of a problem field and its solution are accomplished by constructing and applying the designed artefact. Design science is a technology-oriented applied science, in which practical problems are solved by acquiring knowledge (Chen & Purdin, 1990). Design science also attempts to serve human intentions by designing and creating objects (March & Smith, 1995). Thus, unlike the social and natural sciences, design science seeks to understand reality. March and Smith (1995) identify four types of design science outputs: constructs, models, methods and implementations. Constructs conceptualise problems within the field of interest and identify their solutions. Models specify a set of statements or suggestions linking the constructs that specify situations as problems and solution statements. A method is an algorithm or a guideline employed to execute a task, while an implementation directly operates on constructs, models and methods. The two basic activities of design science are building and evaluation. Figure 14 summarises the seven guidelines that should be implemented in some way for a complete design science study.

Guideline	Description
Guideline 1: Design as an Artifact	Design-science research must produce a viable artifact in the form of a construct, a model, a method, or an instantiation.
Guideline 2: Problem Relevance	The objective of design-science research is to develop technology-based solutions to important and relevant business problems.
Guideline 3: Design Evaluation	The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods.
Guideline 4: Research Contributions	Effective design-science research must provide clear and verifiable contributions in the areas of the design artifact, design foundations, and/or design methodologies.
Guideline 5: Research Rigor	Design-science research relies upon the application of rigorous methods in both the construction and evaluation of the design artifact.
Guideline 6: Design as a Search Process	The search for an effective artifact requires utilizing available means to reach desired ends while satisfying laws in the problem environment.
Guideline 7: Communication of Research	Design-science research must be presented effectively both to technology-oriented as well as management-oriented audiences.

Figure 14: Design of science research guidelines (Hevner, March, Park & Ram, 2004)

Experimental study

Experimental study supports the scientific method. Including the suggestion and testing of hypotheses. The first step is to develop a theory, either from observations or inferences from prior research. From this theory, an examinable statement is proposed. Key (1997) describes experimental study as “an attempt by the researcher to maintain control over all factors that may affect the result of an experiment. In doing this, the researcher attempts to determine or predict what may occur”. In general, experiment study asesses the impact of one or more independent variable on a dependent variable. A true experimental design requires an artificial environment. In an experimental study, the outcomes are recognized as the dependent variable that rely on the action of the independent variable (Howell, 1997). In accordance with the scientific method, this study identifies the problem, proposes hypotheses and realise their consequences, and constructs performs experiments to test these hypotheses. The experiments are undertaken in a controlled environment to evaluate qualities such as usability, accuracy and applicability. For this purpose, an optimal testing method is sought in this study.

As mentioned above, the propose system includes a website and operates in two parts: operational use, based on the proposed solution, and system training candidate. Figure 15 shows the outline of the operational system, in which the fuzzy sets are constructed from an information retrieval algorithm. Figure 16 is the flow chart for the parser method used in this study. During training, the membership is calculated to construct each concept of potential fuzzy sets which is a novel idea. The details of the proposed solution and its preparation for training are described next.

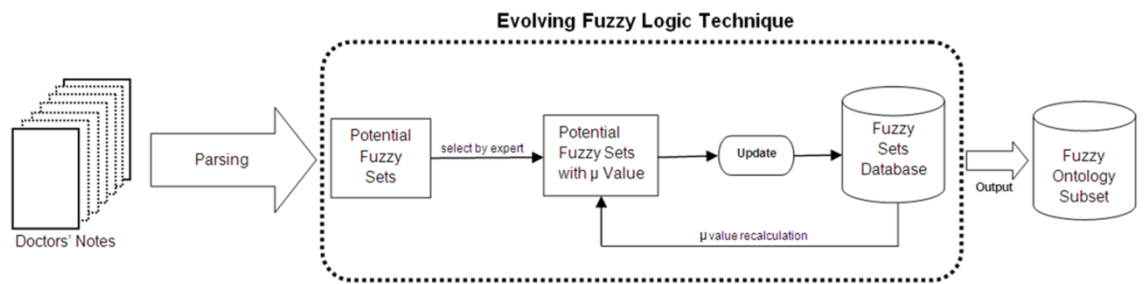


Figure 15: Outline of the operational system

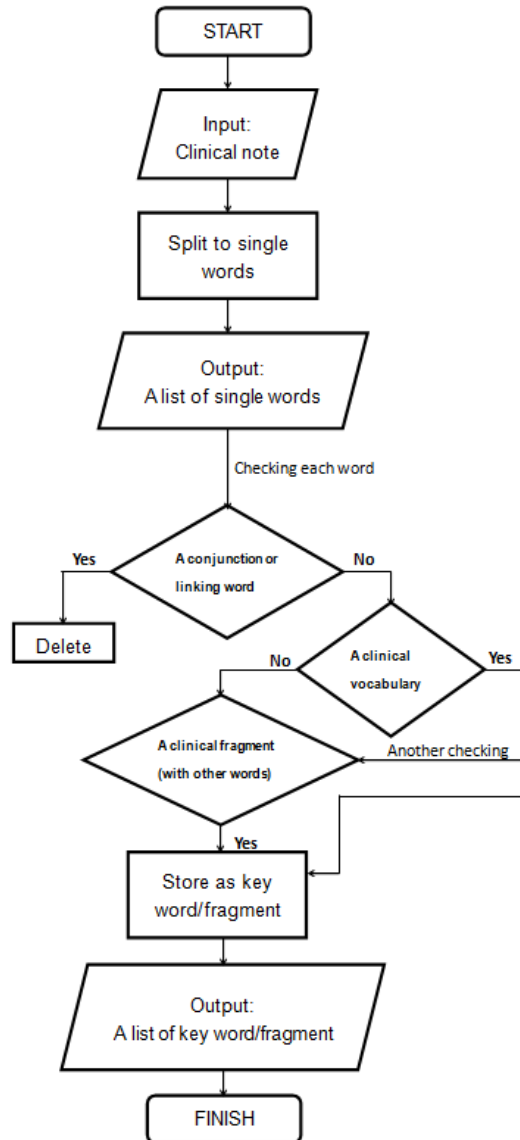


Figure 16: The method of the parser

Proposed solution

A potential fuzzy subset of SNOMED CT will be generated using each clinical note as shown in Figure 17. The first step of the process is to retrieve between the clinical note and SNOMED CT. A single clinical note is examined in each operation. Firstly, the potential keywords are generated from the parsing clinical note, and stored along with the test clinical note in the

database. Here the keyword is any single word, except for conjunctions and linking words (e.g. and, so, or). The system uses the concepts related to each keyword retrieves from SNOMED CT. To ensure efficient retrieval, the information retrieval algorithm is applied at this stage. The Levenshtein distance was chosen as the parsing matching algorithm for the similarity measurement between concepts and keywords. As mentioned in previous sections, this algorithm computes the minimum cost of string transformation, providing a clear indication of the similarity between two strings. Furthermore, we specify a standard value for determining the acceptable concept; that is a concept is accepted only if it contains the specified number of keywords. The experiment will examine several standard values and identify the optimal standard value that generates the most accurate potential concepts. Moreover, the Levenshtein distance algorithm may be extended in a second checking of each potential concept, called phrase checking, that checks the similarity between several keywords and the fragment of a potential concept substitute for single word comparison. In keyword set phrase checking, the original keyword for the potential concept, plus a specified number of other closely-related keywords in the same clinical note. Throughout the checking, the minimum cost of string transformation is also assigned to each potential concept and its keyword set. The output of this implementation is a rank of similarity of all potential concepts and its keyword set. However, the total number of keywords in a keyword set must also be varied in the experiment to identify the optimal number of keywords for precisely comparing the fragments in the free text. The limit of the total number of potential concepts to be transferred to the training section must also be determined, to prevent the training section from becoming overly cumbersome and complicated. The above process yields a membership value that states the degree of similarity given to each of the potential concepts. This degree of similarity is used to construct the fuzzy subset, in which those potential concepts with higher similarity value are selected and stored.

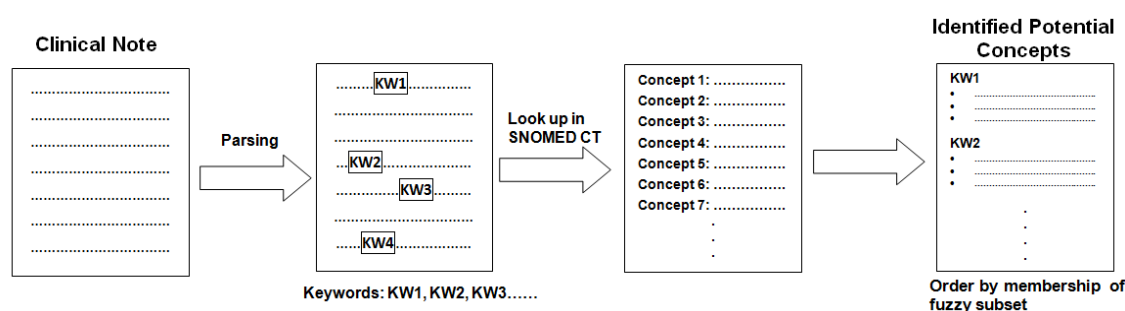


Figure 17: The procedure of the proposed solution

We now summarise the system for the proposed solution. First a fuzzy ontology model of SNOMED CT is built, in which the degree of membership is set to a standard value. Next, a corpus is constructed from anonymous and fictional clinical reports written in free text, with known coding outcomes. A system developed to present selected examples to a training group of clinical staff, who will be requested to automatically link identified keywords and phrases in the text with SNOMED CT codes. The degree of certainty of these links is recorded and forms the basis of the fuzziness of the ontology. A similar process is applied to selection of subset terms. Staffs at Auckland City Hospital have agreed to participate in this phrase. A “public”

system is then made available on the World Wide Web and advertised via professional organisations. These advertisements present both previously created and new test reports, and the preferred meanings selected by the participated users. The selected reports provide maximal coverage, confirm meanings and refine membership values between concepts. A live system that takes previously unknown reports and attempts to code them is then constructed. Finally, prior to public release, the system must be internally tested and refined. Once the system is ready for public release, the clinical stuff from the ultrasound and women's health community are invited to use the system and provide comment. During the trial, users are asked to complete a satisfaction survey and the degree of satisfaction is recorded for each returned concept mapping result. The satisfaction scores returned by users in the trial, the degree of modification of the ontology, and accuracy of pre-coded ultrasound reports are included in the evaluation outcome.

This section focus on discussing the membership calculation and using the potential concepts extracted from the operational use section. Figure 18 shows the procedure of the training section. Firstly, the system provides two lists; one listing the identified potential concepts, the other a sub-list for each of these potential concepts. The sub-list contains descriptions from SNOMED CT. In creating this sub-list, the system retrieves the details from the description section in SNOMED CT associated with the ID number of each concept. Each entity in the sub-list is accompanied by its set of descriptions. The degree of similarity of each description is also listed for experts to select. Following selection of one of the listed descriptions by an expert, the system continues to the next procedure, that of membership calculation. Membership calculation is a learning algorithm that recognises the degree of membership of concepts in the subset. It is based on a simple algorithm that updates the membership function according to the expert responses. If the hierarchical level of the concept is higher than the bottom and its membership value alters, the membership values of all child concepts are accordingly altered. This learning algorithm is also designed to rapidly reflect major changes, while presenting major value oscillations over time. It accounts for the number of assessments already conducted on the potential concept. The update of the fuzzy ontology membership value is weighted by a simple algorithm in which the new membership (μ_{New}) is determined by the old membership (μ_{Old}), and the membership of the current assessment (μ_i). The weighting is given by $\mu_{New} = (\mu_i + \mu_{Old}) / 2$

Figure 18: The procedure of the training section

Experts are requested to select one of four relational values: strongly related (1), somewhat related (0.5), partially related (0.3) and unrelated (0). In above weighting equation, this selection value is inserted as μ_i . Following calculation, if μ_{new} exceeds one it is set to one, and if it falls below zero it is set to zero. However, the membership has a cold-start problem. Therefore, the initial value is set to that noted for all concepts in the SNOMED CT, which is experimentally identified from tagged notes. The refinement of membership in this study, whereby the membership becomes increasingly accurate over time as more assessments are undertaken, is illustrated graphically in Figure 19.

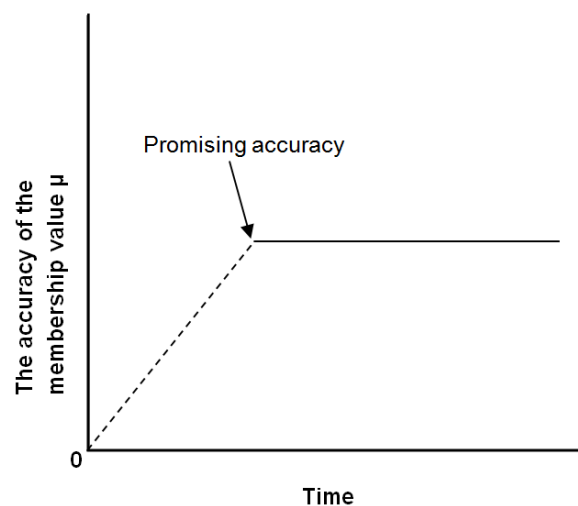


Figure 19: Temporal refinement of membership

3.2 Methods for system development

The system for this study was constructed using the following methods:

- **ASP.NET and Visual Basic.NET**

Because the experiment and the system are website-based; a group of web pages must be created. In this study, ASP.NET was used as a web application framework for building a dynamic website and a web application written in the programming language Visual Basic.NET.

- **Microsoft SQL Server**

The SNOMED CT data and the other data created in this study were handled by the Microsoft SQL server. As previously mentioned, SNOMED CT comprises a vast dataset. Microsoft SQL Server has sufficient capability for rapid handling of these data.

- **Microsoft Windows Server 2003**

The operating server used in this study is Microsoft Windows Server 2003, which can integrate ASP.NET, Visual Basic.NET and Microsoft SQL Server.

3.3 Selection of Experimental Techniques

The following techniques are used in the experimental phase of this study:

- **Parser: The Stanford NLP Parser**

Section 2.9 lists the various parsers considered in this study and their limitations. The Windows operating system precludes the use of parsers that yield poor parsing results, have limited parser lexicon, are incompletely tested or operate at prohibitively slow speed. Among the considered parsers, the Stanford NLP Parsers was selected as the parser that is free from these limitations.

- **Similarity measurement: Levenshtein distance**

SNOMED CT comprises four primary components: Concept IDs, Descriptions, Relationships, and Reference sets. Concepts IDs specify the numerical codes for each clinical term. Descriptions are the textual descriptions of the Concept IDs while relations reveal the relationships among Concept IDs with related meanings. Concepts or Descriptions are grouped by the reference sets. In this study, the similarity measurement for finding matching information between the keywords of clinical notes and SNOMED CT concepts is based on the similarity measurement between two strings. The Description component of SNOMED CT is retrieved and compared. The similarity measurement between two strings is obtained by the Levenshtein distance, based on the numbers of different and identical characters in the strings. The Levenshtein distance also accounts for insertions, deletions, and substitutions in calculating the minimum cost of transforming one string into another enabling a more accurate similarity measure. For these reasons, the Levenshtein distance is used as the similarity measurement in this study.

Chapter 4: Experiments and Results

This chapter includes all the experiments that were carried out in this study which are: the experiments in constructing the system that includes Operation use, training and server setup, and the second part of the experiment focus on the searching of keywords by using Stanford NLP Parser and retrieval of related concepts from SNOMED CT using the Levenshtein distance method to build a fuzzy subset, and then review and evaluate the subset by the clinical professionals

4.1 Prototype of the experimental system

The system is constructed using the methods that are described in Chapter 3. The results of the experimental evaluation of this system prototype are presented in this chapter:

- **Operational use**

Figure 20 shows the front page of the Website that is used for operational use section. The text box allows users to copy and paste the testing clinical notes. When the user presses the “Parse Text” button, the system generates keywords from the testing clinical notes, and uses these keywords to retrieve potential concepts from SNOMED CT. At the same time, it stores all of the database information, such as: the testing clinical notes, keywords, and the users IP address. The current system can detect several typing errors and punctuation marks such as multiple spaces between words, multiple new lines, slashes, and semicolons. Consequently, all of these unwanted items will be excluded from the keywords. Figure 21 displays an example of an entered testing clinical note (after pressing the “Parse Text” button). The right-hand-side of the text box lists the potential concepts matching the keywords. Each of these potential concepts is presented as a link that redirects the user to a new page for conducting the training section, described next.

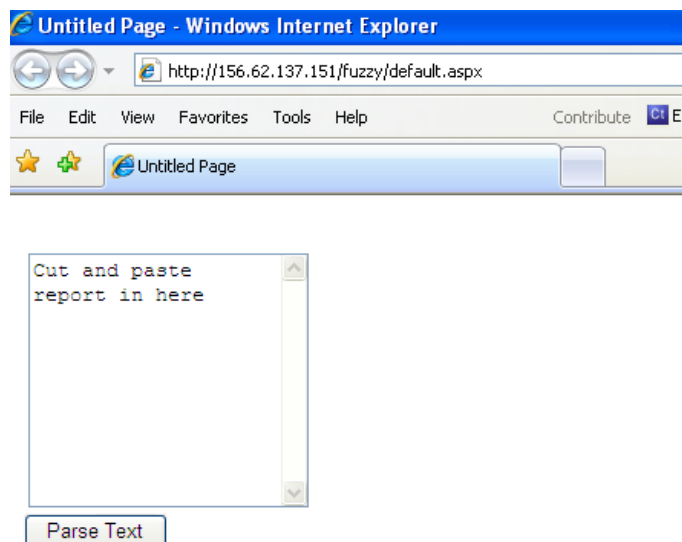


Figure 20: Front page of the system

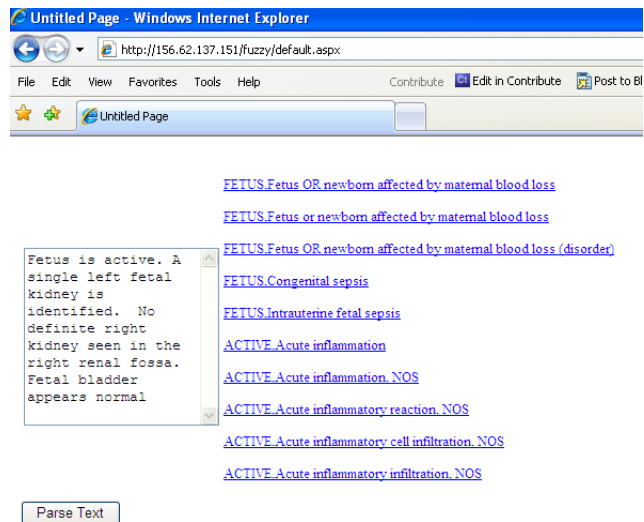


Figure 21: Example of an entered testing clinical note and its output

• Training

Figure 22 shows an example of a new page for a potential concept. The first line of text is the concept ID of the potential concept. The remaining text lists the SNOMED CT description, related to the potential concept. Each of these descriptions contains four different degrees of similarity for users to select. The selection is submitted by pressing the “Click Here to Submit” button at the top of the page. Once the user has pressed this button, the system re-calculates the membership value of the selected potential concept and stores the new membership value in the database. For example, if the old membership is 1 and the clinical professional select the “Somewhat related” in the list – the second highest from four options which means the current assessment membership is 3. The current assessment is the third assessments to evaluate this concept. We will use these three values to re-calculate the membership value for this concept and use as new membership value to replace the old membership value in the database. Once all the web pages for operational use and training have been coded, a server for running this website will be constructed.

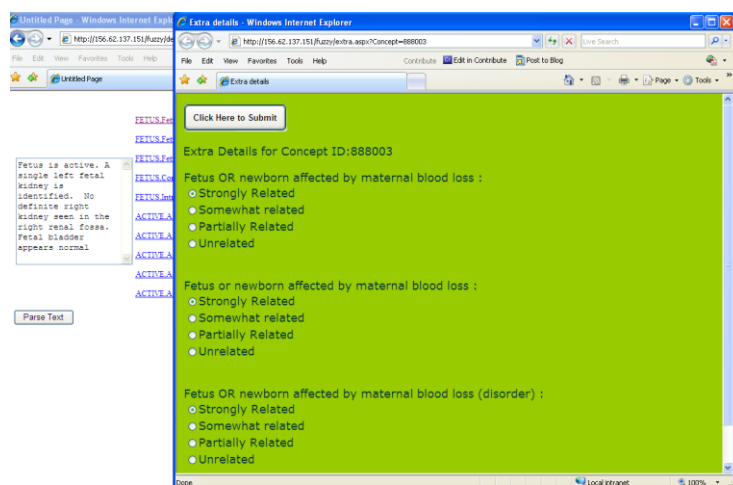


Figure 22: Example of a new page for a potential concept

Server setup

The server for running the website was named “FUZZYON1”. This system is now accessible through the URL <http://156.62.137.151/fuzzy/default.aspx>. The “Remote Desktop Web Connection” utility has also been configured, allowing the server to be remotely accessed and controlled via network. The URL for this utility is <http://156.62.137.151/tsweb>. However, for security reasons, the remote desktop is accessible only within the AUT intranet. Furthermore, the Microsoft SQL Server has been configured to store all system data (including SNOMED CT data) in the “SNOMED” database. Figure 23 displays the ERD of the three main tables of SNOMED CT; namely “Concepts”, “Descriptions” and “Relationships”. The Concepts table contains all concepts IDs in SNOMED CT associated with each clinical term. The textual descriptions of each concepts ID are stored in the Descriptions table. The Concepts and Description tables are linked together with the same field “ConceptID”. When we want to search the related concept from SNOMED CT from a keyword or a phrase, the field to search is “Term” field in Description table. The “Term” field in Description table is stored the textual description for its unique concept ID which the concept ID is store its “ConceptID” field. If any of the textual description contains this keyword or phrase, we will say this concept is the related concept to this keyword or phrase. In the SNOMED CT database, if any of the concepts has a related meaning, the relationships between the SNOMED CT concept and clinical term are stored in the Relationships table. A third web page, “sqladmin” allows users to create or delete the database tables, and is accessible only within the server. This page provides three options: “Create Extra Tables and Columns to Initialize Database” for creating an additional four tables; an option for modifying the table in the “SNOMED” database that stores the experiment result and data such as membership values, testing clinical notes and user’s IP address; and options for restoring the database to its initial state. These options are available because several testing are required prior to running the actual experiment.

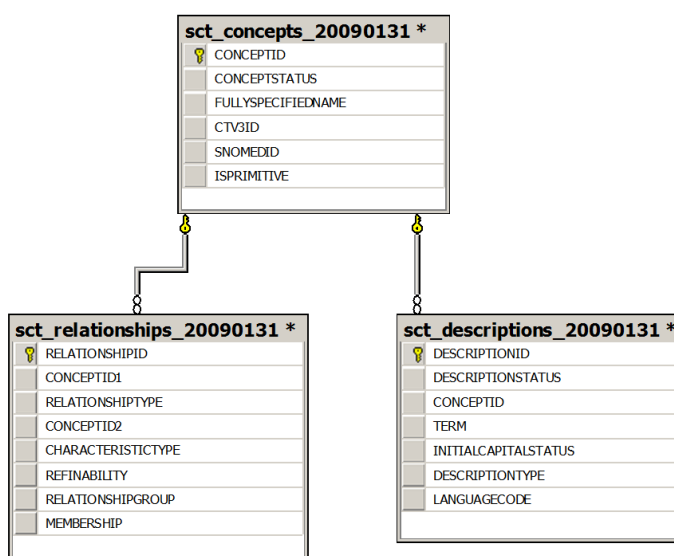


Figure 23: ERD of the three main tables of SNOMED CT

4.2 Experiments

The experiments in this study focus on the searching of keywords by the Stanford NLP Parser and retrieval of related concepts from SNOMED CT using the Levenshtein distance method. This section presents the results of four experiments, and all of them are using DP MRI data from Auckland DHB Radiology. Examples of these data are provided in Appendix 1; a small subset of this Appendix is presented in Table 1. Following the first experiment, the results are assessed by a clinical specialist from Auckland DHB Radiology. The specialist reported that the result should be condensed to contain the diagnosis information only. Based on this assessment, diagnosis information alone was generated for the second part of this experiment. The Levenshtein distance is calculated and applied for identifying the similarity between the clinical note and its related concept in third experiment. The fourth experiment uses the methods from the first two experiments to identify the relationships between the example notes and the potential fuzzy subset. The definition of “Keyword” and “Related concept” are listed in Table 2.

Clinical Information	
Recurrent adenocystic a of nasopharynx. Known bone and lung mets. Now presents with low back pain (over lumbar spine) radiating to bilateral thigh. Associated with urinary hesistency. No paraesthesia/sensory level. No leg weakness. ? Spinal cord/caud	
34 year old man who had a history of allergies and dermatitis, known to the immunology team. Presents with fevers, bilateral leg weakness, urinary retention +/- UTI. Also has a bicuspid aortic valve with a new systolic murmur, awaiting echo. In view of a	

Table 1: Example of DP MRI data

	Definition
Keyword	A word extracted from the clinical document by the parser.
Related concept	A SNOMED concept where a keyword exists in the “Term” field of the “Descriptions” table

Table 2: The definition of “Keyword” and “Related concept”

4.2.1 Experiment 1 and Result

The first experiment involves three sub-sections, each encoded by a computer program that generates the individual results. Figure 24 shows the processing steps of experiment 1. Each of the following sub-sections is detailed in the succeeding paragraphs.

- 1) Using the parser to generate the keyword(s) for each DP MRI data entry.
- 2) Using the keyword(s) to find the related concepts in SNOMED CT.
- 3) Using the keyword(s) to generate phrases from the original DP MRI data entry.

In turn, these are used to identify related concepts in SNOMED CT. The level of

relativity between each of the phrases and their identified concepts is then evaluated by clinical professionals.

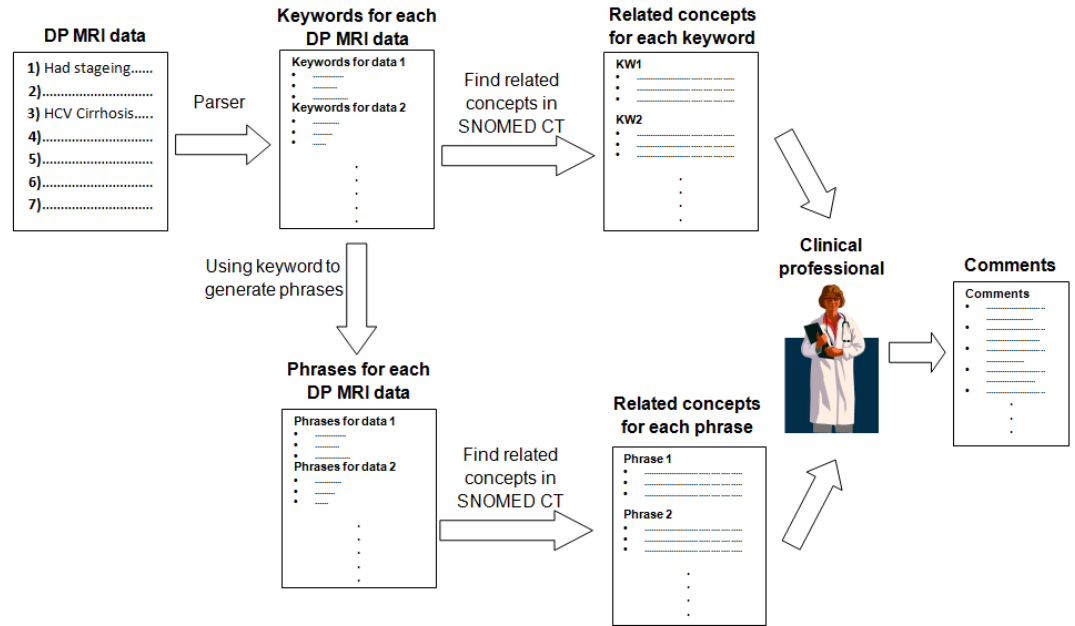


Figure 24: Flowchart of experiment 1

1) Using the parser to generate the keyword(s) for each DP MRI data.

The field “Clinical Information” in the DP MRI data are used as the textual corpus that the parser is applied in it to generate the keyword. The parser used in this sub-section is the Stanford NLP Parser. The loading time of this parser is extended because several components must first be constructed. However, the most of the high-performance parsers suffer from this limitation because they contain enormous libraries of parsing rules. Figure 25 shows the parsing result from Stanford NLP Parser for the input sentence “Active fetus with normal growth and liquor volume” which are: “fetus”, “growth”, “liquor” and “volume”. After the parsing, any word marked as “NN” (Noun), “NP” (Noun Phrase) or “NNP” (Proper Noun) is selected as a keyword. The linking words and phrases which are commonly appearing in the corpus will be filtering out by the parser as well, such as “and”, “or”, “so”, “for” and so on. There are more than 22000 records in this DP MRI dataset. In this sub-section, only the first 1100 records retrieved from the DP MRI dataset are selected, because we consider the total time spending on keywords producing for all the records will be too long and thereby reduce the computational time. Therefore, the total number records for using in this part of the experiment 1 have been reduced to 1100. A subset of the parsing result of the DP MRI data is shown in Appendix 2, a sample of which is presented in Table 3. The first and second columns of Table 3 list the MRI clinical notes and their keywords, respectively.

Tagging

Active/VB fetus/NN with/IN normal/JJ growth/NN and/CC liquor/NN volume/NN ./.

Parse

```
(ROOT
  ($
    (VP (VB Active)
      (NP (NN fetus))
      (PP (IN with)
        (NP (JJ normal) (NN growth)
          (CC and)
          (NN liquor) (NN volume)))))
    (. .)))
```

Typed dependencies

```
root(ROOT-0, Active-1)
dobj(Active-1, fetus-2)
prep(Active-1, with-3)
amod(growth-5, normal-4)
pobj(with-3, growth-5)
cc(growth-5, and-6)
nn(volume-8, liquor-7)
conj(growth-5, volume-8)
```

Typed dependencies, collapsed

```
root(ROOT-0, Active-1)
dobj(Active-1, fetus-2)
amod(growth-5, normal-4)
prep_with(Active-1, growth-5)
nn(volume-8, liquor-7)
prep_with(Active-1, volume-8)
conj_and(growth-5, volume-8)
```

Statistics

Tokens: 9
Time: 0.079 s

Figure 25: Results returned by Stanford NLP Parser to the sentence “Active fetus with normal growth and liquor volume”

Clinical Notes	Keywords
1. Recurrent adenocystic a of nasopharynx. Known bone and lung mets. Now presents with low back pain (over lumbar spine) radiating to bilateral thigh. Associated with urinary hesistency. No paraesthesia/sensory level. No leg weakness. ? Spinal cord/caud	Recurrent, nasopharynx, bone, lung, pain, lumbar, spine, thigh, Associated, hesistency, level, leg, weakness, Spinal
2. 34 year old man who had a history of allergies and dermatitis, known to the immunology team. Presents with fevers, bilateral leg weakness, urinary retention +/- UTI. Also has a bicuspid aortic valve with a new systolic murmur, awaiting echo. In view of a	year, man, history, allergies, dermatitis, team, Presents, fevers, leg, weakness, retention, VBP, UTI, valve, murmur, view

Table 3: Sample of parsing result of the DP MRI data

2) Using the keyword(s) to find the related concepts in SNOMED CT.

In this subsection, the keywords in 1) are used to generate a list of related concepts from SNOMED CT. “Related concept” means the description of the SNOMED CT concept contains the searched keyword. The textual description of each concept is stored in “Descriptions” table in SNOMED CT database and the field “Term” is stored this textual description in this table. Therefore, to obtain the related concepts for a keyword is only need to do the search “Term” field in “Descriptions” table. If the keyword appears in the “Term” field, then we can find concept ID for this “Term” in the field “ConceptID” from same table “Description”. As a result, we can say this

concept is the related concept to the keyword. Examples of keywords and the number of related concepts identified for each keyword are listed in Appendix 3. A small subset of Appendix 3 is provided in Table 4. When we find a concept contains the current searching keyword, we add one to the concept count for this keyword. Therefore, the column “Concepts count” is the total number of concepts that contain the current searching keyword (in column “Single keyword”) in SNOMED CT. Keywords with a related concept count of 0, and also duplicate keywords are deleted. Keywords with a related concept count of 0 means are less interested or common compare to other keywords. The keywords are then sorted in order of highest to lowest number of related concepts. An example of this sorting result is shown in Table 5. The graph in Figure 26, generated from the sorted list, is consistent with Zipf’s law and follows the same trends. Consequently, the plot of frequency versus rank of frequency apparently follows Zipf’s law (Newman, 2005). In Figure 26, the X-axis indicates the commonality of the identified concepts related to keywords (where keywords ordered are from the most to least common in the left-right direction). The Y-axis indicates the number of identified concepts related to each keyword. Zipf’s law states that the frequency of any word is inversely proportional to its rank in the frequency table, given some corpus of natural language utterances. That is the most frequent word will appear approximately twice as frequently as the second most frequent word in the corpus, three times as often as the third most frequent word, and so on.

Single keyword	Concept count
Recurrent	0
nasopharynx	0
bone	10659
lung	3773
pain	3444
lumbar	3180
spine	4035
thigh	1366
Associated	0
hesistency	0

Table 4: Assorted keywords with the numbers of related concepts found in SNOMED CT

Single keyword	Concept count
disorder	280539
finding	162449
organism	105754
body	91630
Product	79609
event	33164
other	31861
therapy	22512
disease	20835
accident	17901
level	17453
tumor	16460
Tumour	15085
syndrome	14230
blood	14141
skin	13880
joint	13589
•	•
•	•
•	•
•	•
•	•
ones	3
onc	3
dyplasia	3
Hippel	2
flows	2
terminus	2
Parry	2
reassess	2
Denies	2
Windsor	2
biopsies	2
dive	2
St3	2
poi	2

Table 5: Result of sorting keywords by concept count

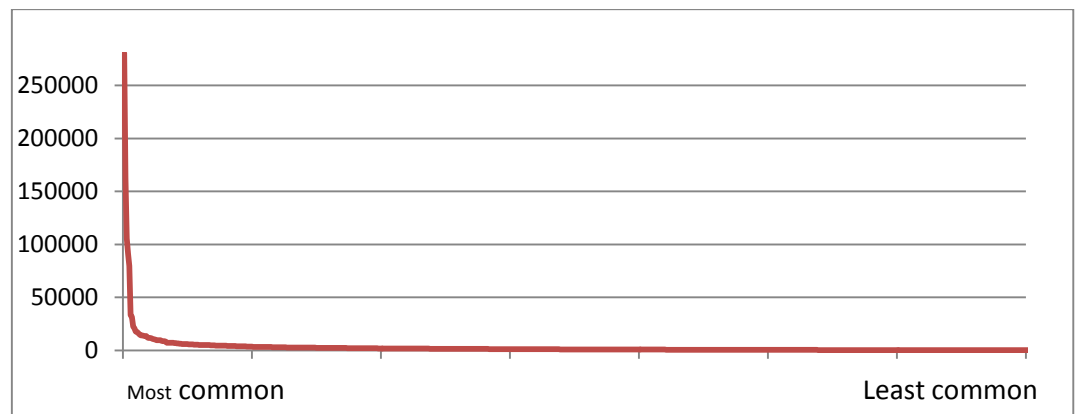


Figure 26: Table is plotted as Zipf's law

3) Using the keyword(s) to generate phrases from the original DP MRI data entry. In turn, these are used to identify related concepts in SNOMED CT. The level of relativity between each of the phrases and their identified concepts is then evaluated by clinical professionals.

This sub-experiment is conducted in four steps and the keywords means the keywords are generated in previous sub-experiment 2). First, two phrases are generated for each keyword: one combining the keyword with its neighbouring left-

side word, the other combining the keyword with its neighbouring right-side word in the clinical notes. Figure 27 shows part of a clinical report, in which the underlined word “cancer” was marked as a keyword. The two phrases generated from this keyword, following the above rules are “skin cancer” and “cancer diagnosis”. Once all the phrases have been generated, 20 phrases are randomly selected from the list. The rationale for random selection of the 20 phrases here is found minimise the total number of phrases and thereby reduce the computational time. Finally, the similarity between each sequential word and its found concept is assessed by clinical specialist. The four steps of this sub-experiment are summarised below:

- Select two sequential words for each keyword from its clinical note: keyword + left-neighbouring word; keyword + right- neighbouring word.
- Sort the list by number of found concepts (descending) and randomly select 20 sequential words from the keywords with the medium number of found concepts.
- Find the matching SNOMED concepts for each of the 20 sequential words.
- Generate a list from which clinical specialists can evaluate the similarity of the 20 sequential words and their found concepts.

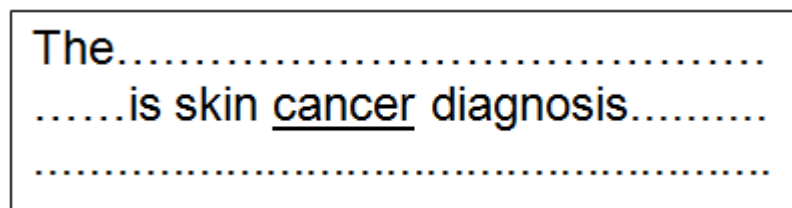


Figure 27: Example of keyword and phrases in a clinical report

The experiment results of these 20 phrases are summarised in Appendix 4. A subset of Appendix 4 is divided into two tables which individual displayed in Table 6a and Table 6b. In Table 6a, the column “Phrase” is the phrase of the keyword in column “Single Keyword” that is used to search the related concept or the clinical note contains it, and the last column “Clinical Notes Contain this Phrase” is listing the clinical notes in the DP MRI data which contain this phrase. In Table 6b, the found related concepts for each phrase are list in column “Concepts” with each of their ID in “Found Concepts ID” column. This result has also been supplied to clinical specialists to check the similarity between the 20 clinical notes and their identified concepts. The feedback provided by the clinical professionals is summarised below:

- The number of the found concepts identified from SNOMED CT has been reduced in the provided result compared to the original search result in the SNOMED CT concept search engine.
- Although number of identified concepts is reduced in the result, the variety type of returned concepts is not compromised. Since clinical professionals rely chiefly on diagnostic information, it preferable to focus on generating the diagnosis concepts only.

No	Single Keyword	Phrase	Clinical Notes Contain this Phrase
1	MRI	MRI brain	MRI brain with thin slices through mamillary bodies. Long history of alcohol abuse with ongoing cognitive deficits noted on objective testingRecent suicide attempt ?ETOH amnesic disorderPreviously booked for CT head, however, clinical director Dr Finucke
	MRI	MRI brain	Thank you. 59yo woman with met breast ca. Has nausea vomiting and CT showed suspicious base of skull mets vs meningioma. Scan recommended MRI brain. scan to determine if meningioma or base of skull mets
	MRI	MRI brain	Bilateral subdural haematomas - drained on 4/1/11. Post-op intra-ventricular haemorrhage. FAST MRI brain to assess progress ?progress
	MRI	MRI brain	stage 4 lung cancer -to start chemotherapy, recent ct brain ? brain mets advised MRI ? BRAIN METS FROM LUNG CANCER
	MRI	MRI brain	Pt w/ multiple cavernomas -- now c/o worsened UE weakness/cramping bilaterally, please perform MRI brain and spine to eval for interval progression Pt w/ multiple cavernomas -- now c/o worsened UE weakness/cramping bilaterally, please perform MRI brain an
	MRI	MRI brain	Atypical choroid plexus tumour. Resected in November 10. Follow-up MRI brain please 3 months post op (end of February/beginning March 2011) Progress following resection of brain tumour
	MRI	MRI brain	newborn spina bifida, needs MRI brain and spine today for OT tomorrow (Tuesday) hydrocephalus, chiari, syrinx...
2	perineal	perineal fistula	known Crohn's disease, on azathioprine, recent admission with vulval lump, continuing discharge, ?formation of fistula known Crohn's disease, recent admission with vulval lump, ?new perineal fistula,

Table 6a: Result of searching for phrases in SNOMED CT

No	Single Keyword	Phrase	Found Concept ID	Concept
1	MRI	MRI brain	241603006	MRI of brain with functional imaging
	MRI	MRI brain	241602001	MRI of brain with T2 mapping
2	perineal	perineal fistula	284077005	Perineal fistula (disorder)

Table 6b: Result of searching for phrases in SNOMED CT

4.2.2 Experiment 2 and Result

Given the feedback provided by professionals in experiment 1, we here focus on isolating the diagnosis concepts. The second experiment comprises four sub-sections, each encode by computer programs that generate the individual results. Figure 28 shows the processing steps in experiment 2. Each of the following sub-sections is detailed in the succeeding paragraphs.

- Selecting 100 reports from DP MRI data.
- Using the parser to generate the keyword(s) for each report.
- Using the keyword(s) to identify the related concepts in SNOMED CT and deleting duplicated keywords and keywords unrelated to any concept in SNOMED CT.
- Using the remaining keywords to generate phrases from each original DP MRI report; from these phrases concepts related to diagnosis only are extracted from SNOMED CT. The level of relationship between each of the phrases and their identified diagnostic concepts is then evaluated by clinical professionals.

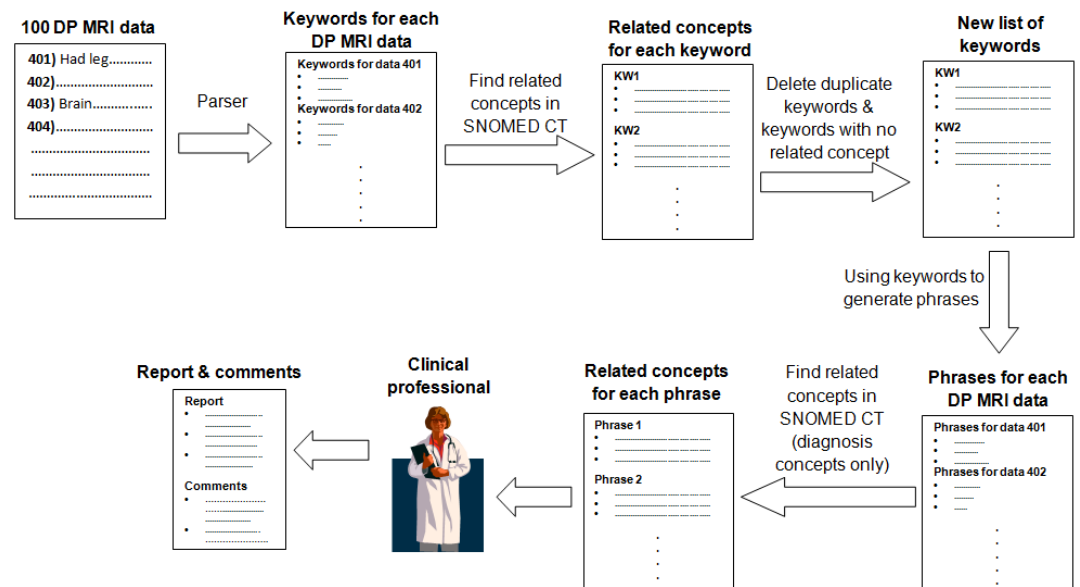


Figure 28: Flowchart of experiment 2

1) Selecting 100 reports from DP MRI data.

Reports numbered 401 through 500 were selected from the DP MRI database. As before, the number of reports was restricted to 100 to reduce processing time.

2) Using the parser to generate the keyword(s) for each report.

Again, the keyword(s) were generated using the Stanford NLP Parser. The parsing results of Reports 401-500 in the DP MRI dataset are listed in Appendix 5. A small subset of Appendix 5 is displayed in Table 7. The first and second

columns in Table 7 list the MRI clinical notes and their associated extracted keywords, respectively.

Clinical Info	Keywords
Unstaged well diff endometrial cancer - treated with definitive XRT completing 29/11/2010; Unable to have surgery due to sig medical comorbidities. Repeat MRI to check for response. Prev MRI in MMH March 2010 ?response to XRT	Unstaged, cancer, XRT, surgery, comorbidities, Repeat, MRI, response, MRI, MMH, March, response, XRT,
73 yr old chinese woman with 1 month of R thigh mass. Noticed one month ago, not noticed to be enlarging. Increased lethargy in last month, no recent fevers or weight loss. Exam: approx. 5x8cm firm mass within ? vastus medialis / adductor compartment. R i	woman, month, R, thigh, mass, month, lethargy, month, fevers, weight, loss, Exam, approx, firm, mass, medialis, adductor, compartment, R,
Stage III abdominal neuroblastoma. Evaluation post chemotherapy/surgery/autologous transplant/RT on ANBL0532 study. ?current disease status	Stage, III, neuroblastoma, Evaluation, post, study, disease, status,
Recurrent acute vestibular neuronitis Need to exclude a central cause - posterior cranial fossa lesion?	vestibular, cause, fossa, lesion,

Table 7: Subset of the parsing result of Reports number 401 – 500 in archived in the DP MRI dataset

3) Using the keyword(s) to identify related concepts in SNOMED CT and deleting duplicated keywords and keywords not related to any concept in SNOMED CT.

Here, keywords identified in 2) are used to generate a list of related concepts from SNOMED CT. Again, the related concepts means the keyword appear in the concept description, so we do the search in the “Descriptions” table in SNOMED CT database. When we find a concept contains the current searching keyword, we add one to the concept count for this keyword. Therefore, the column “Concepts count from SNOMED CT” is the total number of concepts that contain the current searching keyword (in column “Extracted single keyword”) in SNOMED CT. Keywords with a concept count of 0 and also duplicate keywords are deleted. Appendix 6 lists examples of keywords with non-zero concept count. A subset of Appendix 6 is displayed in Table 8.

Extracted single keyword	Concept count from SNOMED CT
cancer	1939
XRT	12
surgery	1776
Repeat	261
MRI	791
response	1131
MRI	791
March	55
response	1131

Table 8: Examples of the keywords with the numbers of related concepts

4) Using the remaining keywords to generate phrases from each original DP MRI report. From these phrases, concepts related to diagnosis only are

extracted from SNOMED CT. The level of relationship between each of the phrases and their identified diagnostic concepts is then evaluated by clinical professionals.

This sub-experiment is conducted in three steps. First, as in experiment 1, two phrases are generated for each remaining keyword: one combining the keyword with its left-side neighbours in the clinical notes; the other combining the keyword with its right-side neighbour. This diagnostic SNOMED concept containing the same phrases are then identified. Finally, the similarity between each sequential word and its selected concepts is assessed by clinical specialists. The three steps of this sub-experiment are summarised below:

- Select two sequential words for each keyword from its clinical note: keyword + left-neighbouring word; keyword + right-neighbouring word.
- Find the diagnostic SNOMED concepts containing the sequential words.
- Generate a list from which clinical specialists can evaluate find the similarity of the sequential words and their found concepts.

The experiment result of searching for (keyword + left-neighbouring word) is provided in Appendix7. A subset of Appendix 7 is presented in Table 9. When the single keyword with its left-neighbouring word (Left Word) has related diagnosis concept in the searching result, the description of each related concepts is list in the “Concept” column and its concept ID is shown in “Related Concepts of “SK+LW” (diagnosis only)” column. The second and third columns in the first eleven rows are all blank which means there is no related diagnosis concepts been found for them. Corresponding results for (keyword + right-neighbouring word) are presented in Appendix 8 and Table 10.

Single Keyword + Left Word	Related Concepts of "SK +LW" (diagnosis only)	Concept
endometrial cancer	Not found	Not found
definitive XRT	Not found	Not found
have surgery	Not found	Not found
comorbidities Repeat	Not found	Not found
Repeat MRI	Not found	Not found
for response	Not found	Not found
Prev MRI	Not found	Not found
MMH March	Not found	Not found
2010 response	Not found	Not found
to XRT	Not found	Not found
chinese woman	Not found	Not found
1 month	5434002	Abortion, 1 month (disorder)
R thigh	157535002	Bruise - hip &/or thigh (disorder)
R thigh	18814006	Contusion of hip AND/OR thigh (disorder)
R thigh	201892008	Osteoarthritis NOS: [pelvic region and/or thigh] or [hip] (disorder)
R thigh	203188007	Chronic osteomyelitis of the pelvic region &/or thigh (disorder)
R thigh	203251001	Periostitis without osteomyelitis, of the pelvic region and/or thigh (disorder)
R thigh	203271005	Tuberculosis of the pelvic and/or thigh bones (disorder)
R thigh	203290002	Poliomyelitis osteopathy of the pelvic region and/or thigh (disorder)
R thigh	209504009	Sprain: [hip and/or thigh] or [groin] or [hamstring] (disorder)

Table 9: Result of searching for (keyword + left-neighbouring word) in SNOMED CT

Single Keyword + Right Word	Related Concepts of "SK +RW" (diagnosis only)	Concept
cancer treated	Not found	Not found
XRT completing	Not found	Not found
surgery due	Not found	Not found
Repeat MRI	Not found	Not found
MRI to	Not found	Not found
response Prev	Not found	Not found
MRI in	Not found	Not found
response to	191008001	Immunodeficiency following hereditary defective response to Epstein-Barr virus (disorder)
response to	230434009	Seizures in response to acute event (disorder)

Table 10: The Result of searching for (keyword + right-neighbouring word) in SNOMED CT

This result shows that the number of concepts has been dramatically reduced relative to experiment 1 and that diagnostic information only has been returned. This result has also been supplied to clinical specialists to check the similarity between the clinical notes and their associated concepts. However, clinical professionals comment was not available on this experiment at the time of thesis submission, because the clinical professional was unable to complete the assessment in the required timeframe.

4.2.3 Experiment 3 and Result

Experiment 3 is an extension of experiment 2, and is performed on the same list of diagnostic concepts identified for each examining clinical report. Experiment 3 comprises two sub-sections, each encoded by the computer programs that generate the individual results. Figure 29 shows the processing steps of experiment 3. The sub-sections are listed below and detailed in the succeeding paragraphs.

- Compute the appearance percentage of the keyword.
- Apply the Levenshtein Distance method.

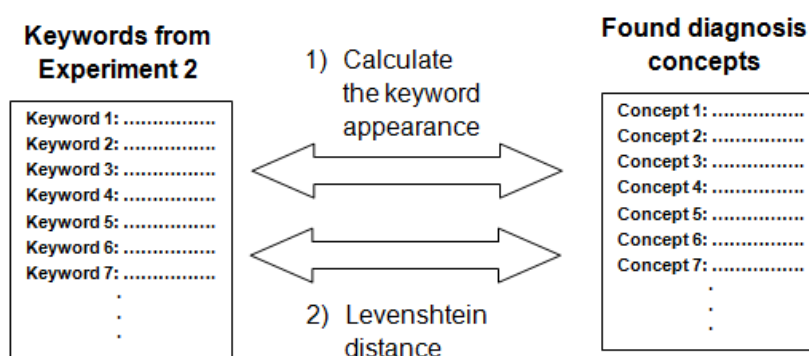


Figure 29: Flowchart of experiment 3

1) Compute the appearance percentage of the keyword.

This step assesses the number of single keywords matched to diagnostic concepts in SNOMED CT that appear in the examining clinical report. The percentage of appearance is also calculated by the following equation: “(number of single keyword appearing in diagnostic concepts and their related examining clinical report / number of single keywords appearing in diagnostic concepts) x 100%”

2) Apply the Levenshtein distance method.

In this step, the Levenshtein distance between each identified diagnostic concept and its corresponding examining clinical report is calculated. In this experiment, the Levenshtein distance is not using for calculating the edit distance between two words as its original method, but it is used to calculate the edit distance between two sentences which is the minimum number of single-word edits (insertion, deletion, substitution) required to change one sentence into the other.

The result of experiment 3 is partially displayed in Appendix 9. From this result, we infer that strategic selection of results with higher appearance percentage and lower Levenshtein distance value isolates the diagnostic concepts that best match the examining clinical report (Appendix 9). However, in this experiment, we must also consider that concepts containing fewer words in the clinical content will more likely generate a higher similarly value and lower Levenshtein distance value. This result

was not assigned to clinical professionals, because it was intended for comparison with/support to experiment 2, for which no evaluation was received.

4.2.4 Experiment 4 and Result

In this experiment, we still focus on isolating the diagnosis concepts in response to the feedback provided by professionals in experiment 1. This experiment comprises five sub-sections, each encoded by computer programs. Figure 30 shows the processing steps in experiment 4. Each of the following sub-sections is detailed in the succeeding paragraphs.

- Selecting 500 reports from DP MRI data.
- Using the parser to generate the keyword(s) for each report.
- Using the keyword(s) to identify the related concepts in SNOMED CT and deleting duplicated keywords and keywords unrelated to any concept in SNOMED CT.
- Using the remaining keywords to generate phrases from each original DP MRI report; from these phrases concepts related to diagnosis only are extracted from SNOMED CT.
- Calculating the “Support” value for each of these found diagnosis concepts.

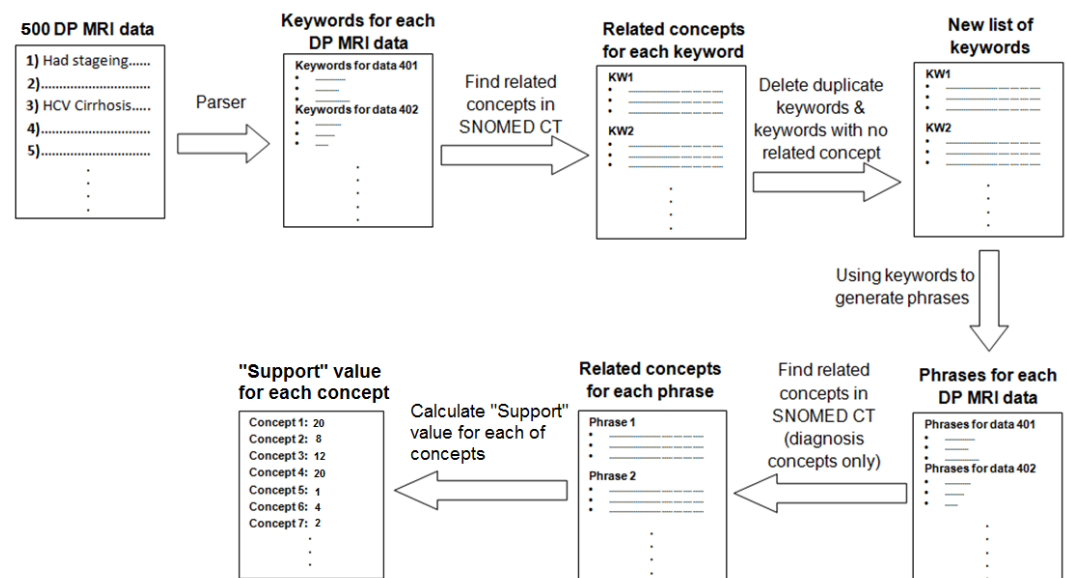


Figure 30: Flowchart of experiment 4

1) Selecting 500 reports from DP MRI data.

500 reports selection: From the DP MRI database, reports numbered 401 through 500 were selected (which is also using in experiment 2) and other 400 notes are selected randomly.

2) Using the parser to generate the keyword(s) for each report.

Again, the keyword(s) were generated using the Stanford NLP Parser. The example of parsing results can be found in Appendix 5. An example of small subset of Appendix 5 is displayed in Table 7 in page 43 in experiment 2.

3) Using the keyword(s) to identify related concepts in SNOMED CT and deleting duplicated keywords and keywords not related to any concept in SNOMED CT.

Here, keywords identified in 2) are used to generate a list of related concepts from SNOMED CT. Again, the related concepts means the keyword appear in the concept description, so we do the search in the “Descriptions” table in SNOMED CT database. When we find a concept contains the current searching keyword, we add one to the concept count for this keyword. Appendix 6 lists examples of keywords with non-zero concept count. A subset of Appendix 6 is displayed in Table 8 in page 43 in experiment 2. The column “Concepts count from SNOMED CT” is the total number of concepts that contain the current searching keyword (in column “Extracted single keyword”) in SNOMED CT. Keywords with a concept count of 0 and also duplicate keywords are deleted.

4) Using the remaining keywords to generate phrases from each original DP MRI report. From these phrases, concepts related to diagnosis only are extracted from SNOMED CT.

This sub-experiment is conducted in two steps. First, as in experiment 1, two phrases are generated for each remaining keyword: one combining the keyword with its left-side neighbours in the clinical notes; the other combining the keyword with its right-side neighbour. This diagnostic SNOMED concept containing the same phrases are then identified. Finally, the similarity between each sequential word and its selected concepts is assessed by clinical specialists. The two steps of this sub-experiment are summarised below:

- Select two sequential words for each keyword from its clinical note: keyword + left-neighbouring word; keyword + right-neighbouring word.
- Find the diagnostic SNOMED concepts containing the sequential words.

5) Calculating the “Support” value for each of these found diagnosis concepts

The “Support” value here means the total number of the found diagnosis concepts been found as the related concept from the 500 selecting reports. For example, if the concept “Broken leg” is the related concept been found in 5 of the 500 selecting reports, the “Support” value for the concept “Broken leg” is 5. The first 800 records of the experiment result (sorted from highest to lowest “Support” value) of the “Support” value for each found diagnosis concepts is provided in

Appendix 10. A subset of Appendix 10 is presented in Table 11. The Levenshtein distance has not been applied in experiment 4, because it will take a long processing time and also more work load for PC as these 500 example notes may have at least 5000 individual related concepts. Levenshtein distance been used in the previous experiment to calculate the similarity between each identified diagnostic concept and its corresponding examining clinical report and the results shows it can be used to assist in fuzzy subset construction. Therefore, Levenshtein distance can be seen as a valuable component that can be integrated into this experiment to improve the experiment results as a potential further study.

Support	Found diagnosis concepts
20	[X]Agoraphobia (& [without history of panic disorder] or [with panic disorder]) (disorder)
20	[X]Adenovirus as the cause of diseases classified to other chapters (disorder)
20	[X]Bacillus fragilis as the cause of diseases classified to other chapters (disorder)
20	[X]Clostridium perfringens as the cause of diseases classified to other chapters (disorder)
20	[X]Escherichia coli [e coli] as the cause of diseases classified to other chapters (disorder)
20	[X]Haemophilus influenzae [h influenzae] as the cause of diseases classified to other chapters (disorder)
20	[X]Klebsiella pneumoniae as the cause of diseases classified to other chapters (disorder)
20	[X]Mycoplasma pneumoniae [pplo] as the cause of diseases classified to other chapters (disorder)
20	[X]Other staphylococcus as the cause of diseases classified to other chapters (disorder)
20	[X]Other streptococcus as the cause of diseases classified to other chapters (disorder)
20	[X]Personal history of allergy to other anti-infective agents (disorder)
20	[X]Proteus (mirabilis)(morganii) as the cause of diseases classified to other chapters (disorder)
20	[X]Pseudomonas as the cause of diseases classified to other chapters (disorder)
20	[X]Staphylococcus aureus as the cause of diseases classified to other chapters (disorder)
20	[X]Streptococcus pneumoniae as the cause of diseases classified to other chapters (disorder)
20	[X]Streptococcus, group a, as the cause of diseases classified to other chapters (disorder)
20	[X]Streptococcus, group b, as the cause of diseases classified to other chapters (disorder)

Table 11: Result of the “Support” value for each found diagnosis concepts

This result shows that we can produce the “Support” value and identify the diagnosis only concepts with highest “Support” value for our examining clinical reports. Furthermore, we can then use these concepts to construct our fuzzy subset in order to obtain the higher degree of relevance and higher classification precision of fuzzy subset to this examining data field.

Chapter 5: Discussion and Conclusion

5.1 Discussion

Clinical knowledge and data are presented free-text format which is ambiguous and precludes simple data mapping. The clinical terminology SNOMED CT contains the most extensive clinical information to date. Clinical notes record various types of clinical information; such as clinical diagnosis, care, and service. Both SNOMED CT and clinical notes are expressed as free-text, and clinical professionals use SNOMED CT to assist them with clinical decision making, and to understand and further interpret their clinical notes. Clearly, mapping clinical notes to SNOMED CT data is a very difficult process, and data mining can return a massive number of concepts that are irrelevant or unhelpful to clinical professionals. Returning to section 1.3, the research questions approached in this study are:

- How can parsers reduce the number of SNOMED CT concepts that related to clinical notes?

Because both clinical notes and SNOMED CT data exist as free-text, they likely contain a number of linking words and phrases. The parser can filter out these linking words and phrases from the clinical notes, removing them from the search routine that matches clinical words/phrases to SNOMED CT concepts. The parser can also analyse sentences, thereby improving the results generated from keywords. The parser used in this study selects only “NN”, “NP” or “NNP” word types as keywords. By restricting the types of selected words to three, all other word types will be excluded as keywords after the parsing. Succinctly, the parser can select more specific keywords and eliminates the common linking words/phrases. Both of these processes reduce the number of SNOMED CT concepts related to clinical notes. In particular, the second process removes commonly used words in human language that clutter the search process.

- How do phrases and single keywords extracted from clinical notes differ in their capacity to identify related SNOMED CT concepts?

The results of experiments 1 and 2 clearly show that fewer related concepts are returned by phrase searching than by single keyword searching. In our method, phrases are composed of a keyword combined with its right or left neighbouring word. This construct enables better matching between the examining clinical note and the concepts identified in the database. Consequently, phrases extracted from clinical notes will be more similarly related to SNOMED CT concepts than single keyword.

- Can fuzzy subset be derived from examples?

As mentioned previously, the clinical professional from Auckland DHB Radiology suggested that phrases retrieved from the database comprise diagnostic information only are the most use to clinical professionals, which is based on evaluated the results from experiment 1 result. We incorporated this suggestion into the second experiment and found that the total number of related concepts was further reduced. As a result, in the system, each of these concepts has higher degree of relevance to a keyword/phrase which is used to construct the

fuzzy subset that can provide clinical professional the better use of SNOMED CT. All the keywords/phrases are generated from examples - DP MRI data from Auckland DHB Radiology. Furthermore, we can generate the diagnosis only concepts with higher appearance rate to the examining clinical reports (examples from DP MRI data) from experiment 4; which also can be used to construct the higher degree of relevance and higher classification precision fuzzy subset, which may be related to the support of each concept. This may indicate a way to derive a fuzzy subset.

Chapter 2 of this thesis presented a literature review on SNOMED CT, Fuzzy logic, Fuzzy subset, Ontology, Fuzzy ontology, and Similarity Measurement. Chapter 3 stated methods used in this study, those involved in the study design and in creating the system and constructing the experimental elements. The details and results of the three experiment attempted in this study are presented in Chapter 4. In experiment 1, a parser was used to generate keywords and phrases for each examining clinical note, whose related concepts were identified in SNOMED CT. Experiment 2 adopted the same methods as experiment 1, but the related concepts identified in SNOMED CT were restricted to diagnostic phrase only. Experiment 3 was a refinement of experiment 2, and used the same dataset to produce two measures; the percentage appearance of words in the identified concepts relative to examining clinical notes and the Levenshtein distance between the identified concept and its examining clinical notes. Due to time constraints, some sections of this study have been incompletely processed, we conclude this thesis with suggestions for further study.

5.2 Further study

The time and other limitations encountered in the present study have precluded a complete evaluation of the proposed methods. Some avenues for future work are briefly stated here:

- **Clinical parser vs. normal parser**

The parser used in this study is a standard parser that is non-specific to the clinical field. Consequently, it retains some keywords that may be less important to clinicians using a clinical parser to generate the keywords that are employed in the matching process, more specific and more relevant matching could be achieved.

- **Integrating the system into SNOMED CT**

We have created a web-based system. However, SNOMED CT has currently been integrated into a number of clinical systems, providing a rapid user-friendly service for clinical professional. Clinical professionals are not required to switch between systems to find a clinical term or perform searches. If our system could be integrated into SNOMED CT, clinical professionals could directly choose whether to continue participating in this study on SNOMED CT, or switch to the web browser and copy and paste their clinical notes into our system.

- **Crowdsourcing**

The clinical professional who evaluated the results from experiment 1 result suggested that phrases retrieved from the database comprise diagnostic information only, since this information is the most use to clinical professionals. We incorporated this suggestion into the second experiment and found that the total number of related concepts was further reduced. As previously discussed, the fuzzy logic technique reduces the size of the resulting fuzzy ontology from that of the original ontology. However, the fuzzy ontology remains difficult to use because it lacks characteristics such as annotation that could assist users in everyday use. This deficit can be improved by adopting the crowdsourcing technique. Crowdsourcing enables the rapid and inexpensive acquisition of annotations for creating a range of predictive models (Huberman, 2008). Crowdsourcing is an open call approach to problem solving or carrying out a task. However, unlike user-generated content, social networks and other popular trends, participants in a crowdsourcing ecosystem have little or no contact with each other. In particular, the work of one user is invisible to other users. Furthermore, according to von Ahn and Dabbish (2004), if a crowd of 5,000 people play an appropriately designed computer game for 24 hours a day, all images on Google (425, 000,000 images in 2005) could be labelled within 31 days. Because so many users access the internet, the online recruitment of anonymous annotators is easily gathered. However, as mentioned by Hsueh, Melville, and Sindhvani (2009), while the online recruitment of anonymous annotators is efficient and cost-effective, it may also attract non-professional advice. Such problems may be reduced by recruiting a specific professional group and requesting their area of expertise. Moreover, these online participants are not usually specifically trained for annotation, and might not invest great effort producing high quality annotations. Consequently, the obtained annotations may be inherently noisy and require additional examination or validation.

Recently, a number of successful applications that employ the crowdsourcing technique have become available. Amazon Mechanical Turk, YouTube, Wikipedia, and reCAPTCHA, which are described in turn below.

Amazon Mechanical Turk

Amazon Mechanical Turk (MTurk) is an online service that includes human intelligence as a central component. It may be described as an “artificial artificial intelligence” system that simply poses questions to participants, rather than achieves a specific task by artificial intelligence approaches (Alonso, Rose, & Stewart, 2008). Figure 31 shows an example of an MTurk task. MTurk also allows researchers to post annotation tasks called Human Intelligence Task (HITs) for a small fee. The tasks are then completed by a large number of anonymous non-expert registered workers (the registered worker in MTurk is the person formally assigned to perform the work). Thus, MTurk can annotate and collect a vast amount of linguistic data at high speed and low cost (Rashtchian, Young, Hodosh, &

Hockenmaier, 2010). Today, MTurk employs more than 200,000 registered workers from over 100 countries, and millions of tasks have been completed.

The screenshot shows the Amazon Mechanical Turk interface. At the top, there's a header with the MTurk logo and navigation tabs: 'Your Account', 'HITS', and 'Qualifications'. A search bar is present with the text 'Search for HITS containing'. Below the search bar, there's a timer 'Timer: 00:00:42 of 15 minutes' and a 'Submit' button. The main task area is titled 'Geography test' and is by 'Amazon Requester Inc.'. It shows a 'Retake Delay' and a 'Qualification Value: 0'. The task instructions are: 'Please evaluate the relevance of the following text fragment. Is the following text relevant to Andorra? Tourism, the mainstay of Andorra's tiny, well-to-do economy, accounts for more than 80% of GDP. An estimated 11.6 million tourists visit annually, attracted by Andorra's duty-free status and by its summer and winter resorts.' There are four radio button options: 'Irrelevant', 'Marginally relevant', 'Fairly relevant', and 'Highly relevant'.

Figure 31: Example of a task implemented in MTurk (Alonso et al., 2008)

YouTube

YouTube is a free video sharing website that allows anyone to upload, watch and share video clips. Most of the content on YouTube has been uploaded by individuals. According to a report compiled Reuters in 2007 (Gill, Arlittz, Li, & Mahanti, 2007), more than 300 million YouTube accounts have been registered and over 65,000 videos per day are uploaded by registered users. Another recent report reveals that more than 13 million hours of video clips were uploaded in 2010 (Scott, 2011). Currently, YouTube is localised in 25 countries, expressed in 43 languages and contains more than 140 million videos. YouTube has proven a successful online video website, in terms of both media posting, and the value added by tagging and commenting on individual clips.

Wikipedia

Wikipedia is an encyclopaedia written by the crowd. The success of Wikipedia is largely due to the collaborative writing process, where by everyone can edit any article, resulting in well organised, and easily accessible articles (Potthast, 2010). Furthermore, article readers are encouraged to improve articles by removing redundancy, correcting errors, improving the writing style, or inserting missing information. In this way, Wikipedia articles are continuously improved and updated. A recent study on Wikipedia revealed that high quality articles in Wikipedia arise from higher number of edits and different contributors (Ganjisaffar et al., 2009). According to a statistics report from Wikipedia itself, more than 3.7 million articles have been posted in the English version of Wikipedia, with over 11 million edits in June 2011. In total, Wikipedia contains 8 billion words in 19 million articles composed in about 270 languages.

reCAPTCHA

A CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) is a challenge-response test that is frequently used in web forms to recognise human users. The tests are designed to be easily solvable by humans but difficult to decipher by automated programs (Egele, Bilge, Kirda & Kruegel, 2010). CAPTCHA aims eliminate form submissions by spambots, automatic scripts that post spam content to any accessible website. reCAPTCHA is a free CAPTCHA service that helps to digitize books, newspapers and old-time radio shows. An example of reCAPTCHA is shown in Figure 32. Each reCAPTCHA test contains two words selected from scanned texts. One word is scanned from a known text in the digital library which determines whether the user entered the correct answer. The other word is an unrecognizable text scanned by Optical Character Recognition (OCR) (von Ahn, Maurer, McMillen, Abraham & Blum, 2008). The test identifies authenticated users relatively easily and quickly, provides website owners with an effective anti-spam system, and enhances the digitisation process for application providers. Currently, reCAPTCHA is digitizing the archives of The New York Times and books from Google Books. Every day, more than 200 million reCAPTCHAs are solved by humans worldwide. Each unrecognizable scanned text takes ten seconds to solve, while individual tests consume little time, the solution time of all tests undertaken globally exceeds 150,000 work hours per day (What is reCAPTCHA, n.d.).

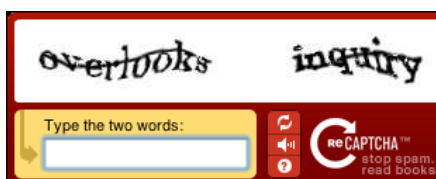


Figure 32: An example of a reCAPTCHA test (CAPTCHA example, see a live demo of a CAPTCHA, n.d.)

Based on the above discussion, it is expected that crowdsourcing will enable clinical professionals to apply their clinical insight and knowledge to the selection of higher-similarity related concepts, which would improve the concept identification system developed for this study. Crowdsourcing presents as a promising technique for improving the performance of a concept identification system. Here we discuss a means by which crowdsourcing could be implemented in the proposed system. In our system, crowdsourcing could be used to automatically obtain the membership value. Furthermore, this membership value calculation could be used to recognise the degree of membership concepts in a subset. The algorithm is based on a simple algorithm that updates the membership among function according to expert responses when the membership value changes at a higher level in the hierarchy of the ontology, an equivalent change is applied to all child concepts. The leaning algorithm is also designed to rapidly reflect major changes, but to avoid major temporal oscillations. It considers the number of assessments already conducted for a potential concept. The fuzzy

ontology membership value updated using a simple weighing that determines the new membership (μ_{New}) from the old membership (μ_{Old}), the membership value of the current assessment (μ_i), and the number of assessments that have confirmed the intended location of this concept (Q_{Hist}).

$$\mu_{New} = \mu_{Old} \pm (\sqrt{(\mu_i - \mu_{Old})^2 / Q_{Hist}})$$

For example, if the old membership is 1 and the clinical professional selects the “Somewhat related” as the relevance rating in the list to this concept for its current examining clinical note – the second highest from four options which means the current assessment membership is 3. And, this assessment is the third assessments to evaluate this concept. We will apply these three values into the equation to re-calculate the membership value (the new value after the calculation is 2) which will be recorded as the new membership value. Applying this algorithm, we expect the membership value will become increasingly accurate after repeated the membership value calculations in theoretically, and as concepts related to MRI are increasingly recognised and assigned by clinical professionals. Thus, it is expected that our system will ultimately yield highly accurate and valuable search results.

REFERENCES

- About SNOMED CT. (n.d.). Retrieved April 5, 2011, from <http://www.ihtsdo.org/snomed-ct/snomed-ct0/>
- Al shayji, S., & El Zant El Kadhi, N. (2011). Building fuzzy-logic ontology for political decision-makers. *International Journal of Mathematical Models and Methods in Applied Science*, 5(5), 991-1001.
- Alonso, O., Rose, D. E., & Stewart, B. (2008). Crowdsourcing for relevance evaluation. *SIGIR Forum*, 42(2), 9-15.
- Blough, D. S. (2001). *The perception of similarity*. Retrieved from <http://www.pigeon.psy.tufts.edu/avc/dblough/theory.htm>
- Briscoe, T., Carroll, J., & Watson, R. (2006, July 17-18). *The second release of the RASP system*. Paper presented at Proceedings of the COLING/ACL on Interactive presentation sessions, Sydney, Australia.
- Calegari, S., & Ciucci, D. (2007). Fuzzy ontology, fuzzy description logics and Fuzzy-OWL. *Applications of Fuzzy Sets Theory*, 4578, 118-126.
- CAPTCHA example, see a live demo of a CAPTCHA [Image] (n.d.). Retrieved May 20, 2011, from <http://www.captcha.net/>
- Castano, S., Ferrara, A., & Montanelli, S. (2003). *H-Match: An algorithm for dynamically matching ontologies in peer-based systems*. Paper presented at Proceedings of the First International VLDB Workshop on Semantic Web and Databases, Berlin, Germany.
- Castillo, O., Melin, P., Kacprzyk, J., & Pedrycz, W. (2007, November 2-4). *Type-2 fuzzy logic: Theory and applications*. Paper presented at the 2007 IEEE International Conference on Granular Computing, California, USA.
- Ceusters, W. M., Spackman, K. A., & Smith, B. (2007). Would SNOMED CT benefit from Realism-Based Ontology Evolution? *Proceedings of the 2007 American Medical Informatics Association Annual Symposium, Biomedical and Health Informatics: From Foundations to Applications to Policy*, 105-109.
- Chiang, M. F., Hwang, J. C., Yu, A. C., Casper, D. S., Cimino, J. J., & Starren, J. (2006). Reliability of SNOMED-CT coding by three physicians using two terminology browsers. *AMIA Annual Symposium Proceedings, 2006*, 131-135.
- Cornet, R., & de Keizer, N. (2008). Forty years of SNOMED: a literature review. *BMC Medical Informatics and Decision Making*, 8, S2. Retrieved from <http://www.biomedcentral.com/1472-6947/8/S1/S2>
- Das, G., Fleisher, R., Gasieniek, L., Gunopulos, D., & Kijlarkijlaninen, J. (1997, June 30-July 02). *Episode matching*. Paper presented at Proceedings of the 8th Annual Symposium on Combinatorial Pattern Matching, Aarhus, Denmark
- Durbin, R., Eddy, S. R., Krogh, A., & Mitchison, G. (1998). *Biological sequence analysis: Probabilistic models of proteins and nucleic acids*. Cambridge: Cambridge University Press.
- Egele, M., Bilge, L., Kirda, E., & Kruegel, C. (2010, March 22-26). *CAPTCHA smuggling: hijacking web browsing sessions to create CAPTCHA farms*. Paper presented at Proceedings of the 2010 ACM Symposium on Applied Computing, Sierre, Switzerland.
- Fensel, D., Lausen, H., Polleres, A., de Bruijn, J., Stollberg, M., Roman, D., & Domingue, J. (2006). *Enabling semantic web services: The web service modeling ontology*. New York: Springer.
- Ganjisaffar, Y., Javanmardi, S., & Lopes, C. (2009). *Leveraging crowdsourcing heuristics to improve search in Wikipedia*. Paper presented at Proceedings of the 5th International Symposium on Wikis and Open Collaboration, Florida, USA.

- Gill, P., Arlitt, M., Li, Z., & Mahanti, A. (2007). *Youtube traffic characterization: a view from the edge*. Paper presented at Proceedings of the 7th ACM SIGCOMM conference on Internet measurement, California, USA.
- Gruber, T. R. (1993). A translation approach to portable ontologies. *Knowledge Acquisition*, 5(2), 199-220.
- Hamming, R. W. (1950). Error detecting and error correcting codes. *Bell System Technical Journal*, 29(2), 147-160.
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75-105.
- Hong, S. K., & Langari, R. (1999). Fuzzy modeling and control of a nonlinear magnetic bearing system. *Journal of Intelligent and Fuzzy Systems*, 7(4), 335-346.
- Howell, D. C. (1997). *Statistical Methods for Psychology* (4th ed.). Boston: Duxbury Press.
- Hsueh, P. Y., Melville, P., & Sindhvani, V. (2009). *Data quality from crowdsourcing: a study of annotation selection criteria*. Paper presented at Proceedings of the NAACL HLT 2009 Workshop on Active Learning for Natural Language Processing, Boulder, USA.
- Huberman, B. A. (2008). Crowdsourcing and Attention. *Computer*, 41(11), 103-105.
- International Health Terminology Standards Development Organisation. (n.d.). *SNOMED CT*. Retrieved April 11, 2011, from https://sl.inforoute.ca/content/disppage.asp?cw_page=snomedct_e
- Kate, R. J. (2013, March 19). *Towards converting clinical phrases into SNOMED CT expressions*. Paper presented at Proceedings of the IWCS 2013 Workshop on Computational Semantics in Clinical Text, Potsdam, Germany.
- Key, J. P. (1997). *Research design in occupational education*. Retrieved October 5, 2011, from <http://www.okstate.edu/ag/agedcm4h/academic/aged5980a/5980/newpage2.htm>
- Lee, C. H., Wang, M. H., & Hagra, H. (2010). A Type-2 fuzzy ontology and its application to personal diabetic-diet recommendation. *IEEE Transactions on Fuzzy Systems*, 18(2), 374-395.
- Leite, M. A. A., & Ricarte, I. L. M. (2008). *Document retrieval using fuzzy related geographic ontologies*. Paper presented at Proceeding of the 2nd international workshop on Geographic information retrieval, California, USA.
- March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support systems*, 15(5), 251-266.
- Mittal, A., Parkash, K., & Mittal, H. (2010). Software cost estimation using fuzzy logic. *ACM SIGSOFT Software Engineering Notes*, 35(1).
- Mittal, H., & Bhatia, P. (2009). Software maintainability assessment based on fuzzy logic technique. *ACM SIGSOFT Software Engineering Notes*, 34(3).
- Needleman, S., & Wunsch, C. (1970). A general method applicable to the search for similarities in the amino acid sequences of two proteins. *Journal of Molecular Biology*, 48(3), 443-453.
- Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5), 323-351.
- Nunamaker, J., Chen, M., & Purdin, T. D. M. (1990). Systems development in information systems research. *Journal of Management Information Systems*, 7(3), 89-106.
- Parry, D. T. (2006). Evaluation of a fuzzy ontology based medical information system. *International journal of Health Information Systems and Informatics*, 1(1), 40-49.

- Parry, D. T., & Tsai, T. C. (2012). Crowdsourcing techniques to create a fuzzy subset of SNOMED CT for semantic tagging of medical documents. *Soft Computing*, 16(7), 1119-1127.
- Patrick, J., Wang, Y., & Budd, P. (2007). *An automated system for conversion of clinical notes into SNOMED Clinical Terminology*. Paper presented at Proceedings of Australasian Workshop on Health Knowledge Management and Discovery, Ballarat, Australia.
- Potthast, M. (2010). *Crowdsourcing a wikipedia vandalism corpus*. Paper presented at Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval, Geneva, Switzerland.
- Rashtchian, C., Young, P., Hodosh, M., & Hockenmaier, J. (2010). *Collecting image annotations using Amazon's Mechanical Turk*. Paper presented at Proceedings of the 2010 NAACL HLT Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk, California, USA.
- Ravi, S., Knight, K., & Soricut, R. (2008). *Automatic prediction of parser accuracy*. Paper presented at Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing, Honolulu, USA.
- Ruch, P., Gobeill, J., Lovis, C., & Geissbühler, A. (2008). Automatic medical encoding with SNOMED categories. *BMC medical informatics and decision making*, 8(Suppl 1), S6.
- Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513-523.
- SANKOFF, D., & KRUSKAL, J. (1983). *Time warps, string edits, and macromolecules: The theory and practice of sequence comparison*. Boston: Addison-Wesley.
- Screenshot. (n.d.). Retrieved Sep 10, 2013, from <https://itunes.apple.com/us/app/snomed-ct-lite/id448278890?mt=8>
- Scott, J. (2011). *25 Jawdropping YouTube video facts, figures, & statistics*. Retrieved from <http://www.reelseo.com/youtube-statistics/>
- Seising, R. (2007). *The Fuzzification of Systems: The Genesis of Fuzzy Set Theory and its Initial Applications - Developments up to the 1970s (Studies in Fuzziness and Soft Computing)*. New York: Springer.
- Shen, S., Radev, D. R., Patel, A., & Erkan, G. (2006). Adding syntax to dynamic programming for aligning comparable texts for the generation of paraphrases. *2006 Proceedings of the COLING/ACL on Main conference poster sessions*, 747-754.
- Shortliffe, E. H., & Cimino, J. J. (2006). *Biomedical Informatics: Computer applications in health care and biomedicine*. New York: Springer.
- Sampson, G. & Babarczy, A. (2003). *Limits to annotation precision*. Paper presented at 2003 10th conference of the European Chapter of the Association for Computational Linguistics, Budapest, Hungary.
- Smith, T. F., & Waterman, M. S. (1981). Identification of common molecular sub-sequences. *Journal of Molecular Biology*, 147(1), 195-197.
- Straccia, U. (2006). A Fuzzy Description Logic for the Semantic Web. In *Capturing Intelligence: Fuzzy Logic and the Semantic Web*, 73-90.
- Sun, Y., Zhang, D., & Cheng, L. (2010). Fuzzy ontology constructing and its application in traditional Chinese medicine. *2010 IEEE International Conference on Intelligent Computing and Intelligent Systems*, 2, 592-595.

- Tho, Q. T., Hui, S. C., Fong, A. C. M., & Cao, T. H. (2006). Automatic fuzzy ontology generation for semantic Web. *IEEE Transactions on Knowledge and Data Engineering*, 18(6), 842-856.
- Tversky, A. (1977). Features of similarity, *Psychological Review*, 84(4), 327-352.
- von Ahn, L., & Dabbish, L. (2004, April 24-29). *Labeling images with a computer game*. Paper presented at Proceedings of the SIGCHI Conference on Human factors in computing systems, Vienna, Austria.
- von Ahn, L., Maurer, B., McMillen, C., Abraham, D., & Blum, M. (2008). *reCAPTCHA: Human-based character recognition via web security measures*. Retrieved from www.cs.cmu.edu/~biglou/reCAPTCHA_Science.pdf
- Wang, C., Lu, J., & Zhang, G. (2008). *An ontology data matching method for web information integration*. Paper presented at Proceedings of the 10th International Conference on Information Integration and Web-based Applications & Services, Linz, Austria.
- What is reCAPTCHA. (n.d.). Retrieved May 20, 2011, from <http://www.google.com/recaptcha/learnmore>
- Widiantoro, D. H., & Yen, J. (2001). *Using fuzzy ontology for query refinement in a personalized abstract search engine*. Paper presented at Proceedings Joint 9th IFSA World Congress and 20th NAFIPS International Conference, Vancouver, Canada
- Zadeh, L. A. (1965). Fuzzy Sets. *Journal of Information and Control*, 8(3), 338-353.
- Zadeh, L. A. (1994). Fuzzy logic: issues, contentions and perspectives. *1994 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 6, 183.
- Zadeh, L. A. (2008a). *Fuzzy logic – A new look* [PDF document]. Retrieved from <http://sinbad2.ujaen.es/sinbad2/files/Conferencias/Zadeh%20presentation.pdf>
- Zadeh, L. A. (2008b, May 19-22). *Is there a need for fuzzy logic?* Paper presented at 2008 Annual Meeting of the North American Fuzzy Information Processing Society, New York, USA.
- Zadeh, L. A. (2009, June 14-17). *From fuzzy logic to extended fuzzy logic - A first step*. Paper presented at 2009 Annual Meeting of the North American Fuzzy Information Processing Society, Cincinnati, USA.

APPENDICES

	Page
Appendix 1: An example of DP MRI data	i
Appendix 2: Result of parsing the DP MRI data	i
Appendix 3: Single keywords and the number of related concepts identified in SNOMED CT	ii
Appendix 4: Experiment search results of 20 phrases	iii
Appendix 5: Results of parsing reports 401 – 500 in the DP MRI dataset	ix
Appendix 6: Examples of keywords and the number of related concept identified in SNOMED CT	xiv
Appendix 7: Examples of experimental searching for (keyword + left-neighbourhood word)...	xxv
Appendix 8: Examples of experimental searching for (keyword + 1 right-neighbourhood word)	xxxiv
Appendix 9: A portion of the results of experiment 3	xl
Appendix 10: A portion (the first 800 records) of the results of experiment 4.....	xlvi

Appendix 1: An example of DP MRI data

Clinical Information
Recurrent adenocystic a of nasopharynx. Known bone and lung mets. Now presents with low back pain (over lumbar spine) radiating to bilateral thigh. Associated with urinary hesistency. No paraesthesia/sensory level. No leg weakness. ? Spinal cord/caud
34 year old man who had a history of allergies and dermatitis, known to the immunology team. Presents with fevers, bilateral leg weakness, urinary retention +/- UTI. Also has a bicuspid aortic valve with a new systolic murmur, awaiting echo. In view of a
Communicating hydrocephalus. For VP shunt on 21/12/10. For 12/12/10 pleasePost op VP shunt Fast MRI.
Suprasellar tumour. For resection on 23/12/10. Post op full MRI + STEALTH For 24/12/10 pleaseMRI post resectionSTEALTH in case further surgery required
Left unilateral tinnitus. r/o vestibular schwannoma
asymmetrical SNHL and tinnitus right ear. r/o vestibular schwannoma
Worsening claudication for 5 years. Bilateral calf claudication at 50 mts. Type 2 DM on tablets. ABI 1.0 right 0.8 left. O/E no tibial pulses. Renal impairment wit creatinine 170.Previous MI and cardiac stenting. Assessment peripheral vascular disease.
Underwent exploratory laminectomy and detethering of cord at C2 level. Review in 6 weeks please. MRI to ensure no further compression of cord.
Periosteal reaction Right ankle. ? Psoriatic arthritis. Clinically no ankle swelling but some MTPsynovitis ? in keeping with Psoriatic arthritis
left leg claudication ? iliac artery occlusion ? status of left leg arteries

Appendix 2: Result of parsing the DP MRI data

Clinical Notes	Keywords
1. Recurrent adenocystic a of nasopharynx. Known bone and lung mets. Now presents with low back pain (over lumbar spine) radiating to bilateral thigh. Associated with urinary hesistency. No paraesthesia/sensory level. No leg weakness. ? Spinal cord/caud	Recurrent, nasopharynx, bone, lung, pain, lumbar, spine, thigh, Associated, hesistency, level, leg, weakness, Spinal
2. 34 year old man who had a history of allergies and dermatitis, known to the immunology team. Presents with fevers, bilateral leg weakness, urinary retention +/- UTI. Also has a bicuspid aortic valve with a new systolic murmur, awaiting echo. In view of a	year, man, history, allergies, dermatitis, team, Presents, fevers, leg, weakness, retention, VBP, UTI, valve, murmur, view
3. Communicating hydrocephalus. For VP shunt on 21/12/10. For 12/12/10 pleasePost op VP shunt Fast MRI.	Communicating, VP, op, VP, Fast, MRI
4. Suprasellar tumour. For resection on 23/12/10. Post op full MRI + STEALTH For 24/12/10 pleaseMRI post resectionSTEALTH in case further surgery required	tumor, resection, Post, MRI, +, STEALTH, pleaseMRI, resectionSTEALTH, case, surgery
5. Left unilateral tinnitus. r/o vestibular schwannoma	tinnitus, vestibular, schwannoma
6. asymmetrical SNHL and tinnitus right ear. r/o vestibular schwannoma	asymmetrical, SNHL, ear, vestibular, schwannoma
7. Worsening claudication for 5 years. Bilateral calf claudication at 50 mts. Type 2 DM on tablets. ABI 1.0 right 0.8 left. O/E no tibial pulses. Renal impairment wit creatinine 170.Previous MI and cardiac stenting. Assessment peripheral vascular disease.	claudication, years, Bilateral, claudication, mts, Type, tablets, ABI, pulses, impairment, wit, stenting, Assessment, vascular, disease
8. Underwent exploratory laminectomy and detethering of cord at C2 level. Review in 6 weeks please. MRI to ensure no further compression of cord.	laminectomy, detethering, cord, level, Review, weeks, MRI, compression, cord
9. Periosteal reaction Right ankle. ? Psoriatic arthritis. Clinically no ankle swelling but some MTPsynovitis ? in keeping with Psoriatic arthritis	reaction, arthritis, MTPsynovitis, arthritis
10. left leg claudication ? iliac artery occlusion ? status of left leg arteries	leg, claudication, artery, occlusion, status, leg, arteries

Appendix 3: Single keywords and the number of related concepts identified in SNOMED CT

Single keyword	Concept count
Recurrent	0
nasopharynx	0
bone	10659
lung	3773
pain	3444
lumbar	3180
spine	4035
thigh	1366
Associated	0
hesistency	0
level	17453
leg	3403
weakness	268
Spinal	5137
year	763
man	247
history	5994
allergies	0
dermatitis	0
team	189
Presents	0
fevers	33
leg	3403
weakness	268
retention	0
VBP	0
UTI	52
valve	3402
murmur	436
view	67
Communicating	0
Fast	401
MRI	791
tumor	16460
resection	0
Post	757
MRI	791
STEALTH	0
pleaseMRI	0
resectionSTEALTH	0
case	619
surgery	1776
tinnitus	0

Appendix 4: Experiment search results of 20 phrases

No	Single Keyword	Phrase	Found Concept ID	Concept	Clinical Notes Contain this Phrase
1	MRI	MRI brain	241603006	MRI of brain with functional imaging	MRI brain with thin slices through mamillary bodies. Long history of alcohol abuse with ongoing cognitive deficits noted on objective testingRecent suicide attempt ?ETOH amnesic disorderPreviously booked for CT head, however, clinical director Dr Finucke
			241602001	MRI of brain with T2 mapping	Thank you. 59yo woman with met breast ca. Has nausea vomiting and CT showed suspicious base of skull mets vs meningioma. Scan recommended MRI brain. scan to determine if meningioma or base of skull mets
					Bilateral subdural haematomas - drained on 4/1/11. Post-op intra-ventricular haemorrhage. FAST MRI brain to assess progress ?progress
					stage 4 lung cancer -to start chemotherapy, recent ct brain ? brain mets advised MRI ? BRAIN METS FROM LUNG CANCER
					Pt w/ multiple cavernomas -- now c/o worsened UE weakness/cramping bilaterally, please perform MRI brain and spine to eval for interval progression Pt w/ multiple cavernomas -- now c/o worsened UE weakness/cramping bilaterally, please perform MRI brain an
					Atypical choroid plexus tumour. Resected in November 10. Follow-up MRI brain please 3 months post op (end of February/beginning March 2011) Progress following resection of brain tumour
					newborn spina bifida, needs MRI brain and spine today for OT tomorrow (Tuesday) hydrocephalus, chiari, syrinx...
2	perineal	perineal fistula	284077005	Perineal fistula (disorder)	known Crohn's disease, on azathioprine, recent admission with vulval lump, continuing discharge, ?formation of fistula known Crohn's disease, recent admission with vulval lump, ?new perineal fistula,
3	oxygen	home oxygen	170621009	Home oxygen supply stopped (situation)	14 weeks old baby (ex prem 31 weeks, currently 2 months corrected age) with history of apnoea and bradycardia of newborn, currently on home oxygen. Plan is to trial wean off on next admission exclude central cause (brain stem/cerebral)
			268512000	Home oxygen supply (finding)	
			275915008	Home oxygen supply started (finding)	
			445966008	Assessment for home oxygen therapy	
			426990007	Home oxygen therapy (procedure)	
			34847001	Home oxygen therapy education (procedure)	
4	post	post operative	394935005	Combined post-operative chemotherapy and radiotherapy	Participant in ReBICS study (Reducing Brain Injury in Cardiac Surgery). Unwell last week for a routine post operative MRI Brain. Needs MRI Brain (DWI) to rule out peri operative stroke. ? perioperative stroke
			65590007	Diagnostic radiography with contrast media by injection, positive contrast, post-operative (procedure)	
			48995004	Fluoroscopy, post-operative	

			17348004	Operative procedure on posterior segment of eye (procedure)	
			49730009	Post-operative follow-up, anesthesia (procedure)	
			403690007	Extensive post-operative bruising	
5	post	post chemo	237787007	Post-chemotherapy ovarian failure (disorder)	high risk locally advanced breast cancer. 1 year post chemo. 1/12 increasing headaches, ataxia, rhombeg positive, getting worse. CT 1/12 ago Tiny high density foci (left cerebellum and tentorium bilaterally) .MRI recommended. (was requested 3 weeks ago, refer
		post chemotherapy	236813004	Post-chemotherapy testicular hypofunction (disorder)	Stage III abdominal neuroblastoma. Evaluation post chemotherapy/surgery/autologous transplant/RT on ANBL0532 study. ?current disease status
			262502001	Post-chemotherapy (qualifier value)	
6	fibrosis	cystic fibrosis	444260001	Cystic fibrosis 20 common mutation test	47y with cystic fibrosis. Diagnosed as adult. Previous left mastoidectomy. Known left sinus disease. Previous sinus surgery. Complains of severe persistent left-sided daily headache causing dizziness and collapse. Mastoid cavity clean and dry. ? Central c
			443530000	Cystic fibrosis 29 common mutation test	
			44489000	Cystic fibrosis carrier detection (procedure)	
			69103000	Cystic fibrosis carrier detection, amniotic fluid (procedure)	
			41636009	Cystic fibrosis carrier detection, blood (procedure)	
			117254004	Cystic fibrosis respiratory culture (procedure)	
			171191008	Cystic fibrosis screening (procedure)	
			314080004	Cystic fibrosis screening test (procedure)	
			86964003	Cystic fibrosis sweat test (procedure)	
			67799006	Cystic fibrosis, prenatal detection (procedure)	
			428273001	Cystic fibrosis not suspected	
7	vessels	retinal vessels	247125005	Anastomotic retinal vessels (finding)	Family history of sudden death from intracranial aneurysm, and tortuous retinal vessels. Apparently autosomal dominant - this woman has abnormal retinal vessels and anterior segment dysgenesis, lens subluxation suggesting a connective tissue disorder - ? C
					?COL4A1 mutation, autosomal dominant family history of tortuous retinal vessels and intracranial aneurysms, mother affected-this patient has retinal tortuosity ?intracranial aneurysm or vascular abnormalities, strong family history consistent with COL4A1
8	Palsy	Bell's Palsy	161483005	History of - Bell's palsy (situation)	41 year old, presented with acute Bell's Palsy, documented in ED 8 days ago, started on Prednisone and Aciclovir. - Presented to ophthalmology follow up today. - On examination, Decreased function of L) 5th nerve (with weak masseter, V1, V2, V3 numbness and
			193093009	Bell's palsy (disorder)	Profound hearing loss, for cochlear implantation Incomplete recovery from left 'Bell's' palsy 3 years ago ?lesion along left facial nerve?

9	Palsy	cerebral Palsy	278512001	Ataxic cerebral palsy (disorder)	8yo girl with cerebral palsy, sustained R.elbow medial condyle # and had ORIF 11/11/10. K wires removed 20/12/10. At clinic F/U, Xrays suggested osteomyelitis distal R.humerus. Febrile, redness, swelling, purulent discharge from R.elbow. CRP22 ? osteo
			192957004	Ataxic infantile cerebral palsy (disorder)	Will need GA 4 year old girl with persistent toe-walking and expressive language delay. Increased tone left lower leg and clonus at ankle. Will need GA. ? mild cerebral palsy ? other central cause
			75019001	Athetoid cerebral palsy (disorder)	
			128188000	Cerebral palsy (disorder)	
			230774004	Cerebral palsy with spastic tetraparesis (disorder)	
			307756005	Cerebral palsy, not congenital or infantile, acute (disorder)	
			275466008	Congenital cerebral palsy (disorder)	
			275467004	Congenital spastic cerebral palsy (disorder)	
			230780007	Dyskinetic cerebral palsy (disorder)	
			230781006	Dystonic/rigid cerebral palsy (disorder)	
10	Palsy	Palsy LMN	164106005	On examination - cranial nerve 7-palsy-LMN (finding)	Pt s/p head injury diving into pool w/ R facial n palsy (LMN) and RUE weakness worse C7/8 -- no obvious injury on CT scan. Needs MRI brain and Cspine to eval for etiology. Pt s/p head injury diving into pool w/ R facial n palsy (LMN) and RUE weakness wor
11	new	new onset	233821000	New onset angina (disorder)	New onset epilepsy, possibly temporal lobe Is there a focal lesion that is the cause for the epilepsy
12	pneumonia	aspiration pneumonia	422588002	Aspiration pneumonia	14 year old girl. 2 VF arrest in the community. Most recent episode complicated by aspiration pneumonia. Normal QT intervals on ECG and normal exercise test. No family history of sudden death. ? ARVC, ? other structural abnormalities
			83608006	Aspiration pneumonia due to inhalation of milk (disorder)	
			42004004	Aspiration pneumonia due to inhalation of vomitus (disorder)	
			72854003	Aspiration pneumonia due to near drowning (disorder)	
			75426006	Aspiration pneumonia due to regurgitated food (disorder)	
			40786001	Aspiration pneumonia due to regurgitated gastric secretions (disorder)	
			44549008	Aspiration pneumonia resulting from a procedure (disorder)	
			276695003	Neonatal aspiration pneumonia (disorder)	
			438764004	Postoperative aspiration pneumonia	
			430969000	Recurrent aspiration pneumonia	
13	acuity	visual acuity	400908006	Central, steady, maintained visual acuity assessment (procedure)	12 yo boy been having migraines assoc with vomiting averaging 2 episodes a week for last 9 years. Over past 5 years has had worsening central vision. Examination showed reduced visual

					acuity associated with bilateral optic disc hyperaemia and macular oede
			441841000	Pinhole visual acuity test (procedure)	
			16830007	Visual acuity testing (procedure)	
			419775003	Best corrected visual acuity	
			425141002	Brightness acuity testing visual acuity	
			397536007	Corrected visual acuity	
			397534005	Corrected visual acuity - left eye	
			397535006	Corrected visual acuity - right eye	
			251742009	Distance visual acuity - binocular (observable entity)	
			251739003	Distance visual acuity (observable entity)	
14	stone	duct stone	384647004	Bile duct stone removal	biliary colic and cholecystectomy 2008-9cholecystectomy was almost a year after ERCPepisode biliary colic / cholangitis 12/2010 ? retained stone ? retained bile duct stone
			235548005	Chemodissolution of bile duct stone (procedure)	
			30585001	Common duct stone analysis (procedure)	
			65313007	Percutaneous extraction of common duct stones (procedure)	
			277051004	Percutaneous fragmentation of bile duct stone (procedure)	
			235923000	Retained bile duct stone (disorder)	
15	Atypical	Atypical meningioma	128914005	Atypical meningioma (morphologic abnormality)	Atypical meningioma. Widespread sagittal convexity with parafalcine meningioma. Previously had multiple debulking surgeries. Recently received palliative radiotherapy to the brain, 36 Gy in 12 fractions completed 24/08/2010. Needs follow up scan in May
16	Atypical	Atypical choroid	128904001	Atypical choroid plexus papilloma (morphologic abnormality)	Atypical choroid plexus tumour. Resected in November 10. Follow-up MRI brain please 3 months post op (end of February/beginning March 2011) Progress following resection of brain tumour
17	Fixed	Fixed flexion	299234001	Hip joint fixed flexion deformity (finding)	Right hip pain x 6/12. No trauma. Normal inflammatory markers. Reduced ROM with pain in right hip. Fixed flexion deformity. Xrays - ? asymmetry in SIJ + ?osteopaenia in right hip. MRI within 1/52 if possible please. Discussed with child/grandfather - abl
			281492001	Acquired fixed flexion deformity of the hip (disorder)	
			240230000	Acquired fixed flexion deformity of the elbow (disorder)	
			281493006	Acquired fixed flexion deformity of the knee (disorder)	
			281479004	Acquired fixed flexion deformity of the wrist (disorder)	
			299419009	Ankle joint fixed flexion deformity (finding)	
			281478007	Acquired fixed flexion deformity of the shoulder (disorder)	

			26803600 5	Acquired fixed flexion deformity of finger (disorder)	
18	Relative	Relative afferent	24701500 2	Relative afferent pupil defect (finding)	32 year old man with complete visual loss in one eye and loss of temporal hemifield other eye. Relative afferent pupillary defect. Bilateral cupped discs. Also symptoms of hyperthyroidism. ? Pituitary lesion? ? Pituitary tumour
			23212200 3	Relative afferent pupillary defect (disorder)	
19	dose	high dose	44808900 4	Medication dose too high	Adenocarcinoma with solitary left brain met.nov 2010. lung primary in situ 3 nov drainage and biopsy of malignant cyst. dec reoperation for infected subdural collection. planning high dose radiotherapy. baseline scan for radiotherapy planning
			41624600 6	Prolonged high dose use of cannabis	Known case of ADEM with acute deterioration in seizure activity. MRI on 28/01/11: lesions consistent with ADEM, unable to be clear whether represent old lesions or new ones. Started on high dose Methylprednisolone. Intubated, ventilated. any progress/reso
			17066200 3	Using inhaled steroids - high dose (finding)	12 year old girl with Takayasu's disease. Now presenting with chest, back pain and severe headaches. Raised inflammatory markers ? On high dose steroids . relapse / deterioration of disease ? relapse / worsening arteritis
			31983005	High dose immunologic tolerance	12 yr old girl with idiopathic pulmonary haemosiderosis with recurrent presentations with relapses, 3 admissions in 3 months, on high dose steroids consistently for past 3 months. New issue in past week of intermittent severe abdominal pain in RUQ and ep
			22886700 5	High dose rate brachytherapy	
			25219200 3	High dose dexamethasone suppression test (procedure)	
			26182900 3	High dose (qualifier value)	
			44375100 0	Measurement of platelet aggregation using high dose ristocetin	
			44858300 6	High dose opiate anaesthesia	
			41727300 2	Diphtheria (high dose)/ tetanus/pertussis/inactivated poliomyelitis vaccine 0.5mL prefilled syringe	
20	atresia	pulmonary atresia	25362500 4	Pulmonary atresia with absent pulmonary artery (disorder)	22 year old. Congenitally corrected transposition of the great arteries with VSD and pulmonary atresia Right modified BT shunt, now disconnected Lateral tunnel Fontan operation 1993 Mr Kerr Routine Cardiac MRI surveillance
			25362400 0	Pulmonary atresia with confluent pulmonary arteries (disorder)	
			44957500 6	Acquired pulmonary atresia	
			25359400 0	Muscular pulmonary atresia (disorder)	
			44856400 4	Pulmonary atresia	
			44783800 3	Pulmonary atresia and ventricular septal defect with aorta from left ventricle	

			44783900 6	Pulmonary atresia and ventricular septal defect with aorta from right ventricle	
			25359000 9	Pulmonary atresia with intact ventricular septum (disorder)	
			25359100 8	Pulmonary atresia with ventricular septal defect (disorder)	
			44815500 3	Pulmonary atresia with ventricular septal defect of Fallot type	
			25330300 1	Solitary aortic trunk with pulmonary atresia (disorder)	
			25351300 5	Tetralogy of Fallot with pulmonary atresia (disorder)	
			39572200 1	Repair of pulmonary atresia	
			38464800 9	Repair of pulmonary atresia with ventricular septal defect	

Appendix 5: Results of parsing reports 401 – 500 in the DP MRI dataset

Clinical Info	Keywords
Unstaged well diff endometrial cancer - treated with definitive XRT completing 29/11/2010; Unable to have surgery due to sig medical comorbidities. Repeat MRI to check for response. Prev MRI in MMH March 2010 ?response to XRT	Unstaged, cancer, XRT, surgery, comorbidities, Repeat, MRI, response, MRI, MMH, March, response, XRT,
73 yr old chinese woman with 1 month of R thigh mass. Noticed one month ago, not noticed to be enlarging. Increased lethargy in last month, no recent fevers or weight loss. Exam: approx. 5x8cm firm mass within ? vastus medialis / adductor compartment. R i	woman, month, R, thigh, mass, month, lethargy, month, fevers, weight, loss, Exam, approx, firm, mass, medialis, adductor, compartment, R,
Stage III abdominal neuroblastoma. Evaluation post chemotherapy/surgery/autologous transplant/RT on ANBL0532 study. ?current disease status	Stage, III, neuroblastoma, Evaluation, post, study, disease, status,
Recurrent acute vestibular neuronitis Need to exclude a central cause - posterior cranial fossa lesion?	vestibular, cause, fossa, lesion,
Large suprasellar JPA. Multiple debulkings. Completed 10 months chemotherapy. ?current status at end of planned chemotherapy	suprasellar, Multiple, months, status, end, chemotherapy,
metastatic melanoma. recent MRI spine to investigate bilateral leg weakness and paraesthesia. no spinal cord compression but concerns re spinal stenosis especially at cervical spine. seen by neurology team clinical problem likely related to cervical can	metastatic, MRI, spine, leg, weakness, paraesthesia, cord, compression, concerns, stenosis, spine, neurology, team, problem,
Haemophagocytic lymphohistiocytosis (HLH). Previously widespread disease with liver, spleen, lung and widespread skin involvement with multiple subcutaneous nodules. ?any evidence of previously documented disease	Haemophagocytic, HLH, disease, liver, spleen, lung, skin, involvement, subcutaneous, evidence, disease,
03/08/2006 Right frontal craniotomy and acrylic cranioplasty for grade 1 frontal meningioma.11/10/2007 Right craniotomy for right frontal convexity meningioma and medial wing sphenoid meningioma.24/05/2010 Right craniotomy for recurrent sph	craniotomy, cranioplasty, grade, meningioma, convexity, meningioma, wing, meningioma, sph,
Retrobulbar optic neuritis, right sided hemiparesis. Slow to recover ? New lesions to suggest MS	Retrobulbar, neuritis, hemiparesis, New, MS,
Local advanced lower rectal adenocarcinoma. Completed combined chemo/RT end of December. Awaiting surgery. Needs CT staging chest abdo pelvis + MRI pelvis to assess treatment response please. ? treatment response assessment	Local, adenocarcinoma, Completed, end, December, surgery, Needs, chest, abdo, pelvis, MRI, pelvis, treatment, response, treatment, response, assessment,
Pineal parenchymal tumour of intermediate differentiation. Debulked. Radiotherapy. Completed chemotherapy January 2011. ?current status of residual tumour	Pineal, tumor, differentiation, Completed, January, status, tumor,
18 week old baby with flexor spasms, developmental regression and modified hypsarrhythmia = infantile spasms. Needs MRI head as part of work up please ? underlying brain pathology. There was parental report of some abnormal ventricles found on antenatal	week, baby, flexor, spasms, regression, spasms, Needs, MRI, part, work, brain, report, ventricles, antenatal,
HCV Cirrhosis. Regenerative nodules, for alternating MRI/US screening. USS Oct '10. AFP normal. Please arrange MR liver in approx April. ?HCC ?complications of cirrhosis	HCV, Regenerative, US, screening, USS, Oct, MR, liver, approx, April, complications, cirrhosis,
Pt w/ recurrent meningioma for operation, needs current MRI +/- w/ stealth query change/preop planning	Pt, w, meningioma, operation, MRI, JJ, w, NN,
SCC Anus, completed chemoradiotherapy Assess response after treatment please	SCC, Anus, Assess, response, treatment,
Left lateral ventricular tumour - resected 19/11/20. Ventricular blood clot - evacuated 23/11/10. VP shunt placement on 14/12/10. Shunt failure. For shunt revision. Fast MRI brain pre-op to check ventricle size Fast MRI pre-op (shunt revision)	ventricular, tumor, Ventricular, blood, VP, placement, failure, revision, MRI, brain, pre-op, size, Fast, MRI, pre-op, revision,
Community VF arrest. CAD but uncertain whether this is the cause of arrest as not necessarily obstructive. Essentially normal echo. Creat 64. ?acute MI. ?myocardial scarring. ?ARVC.	Community, VF, arrest, CAD, cause, arrest, myocardial,
Homocystinuria with Systemic lupus erythematosus. Chronic steroids for 3 years. Onset L sided thigh pain and lower leg. clinically reduced ROM L hip. ? AVN secondary to chronic steroids ? erosions. ? other hip pathology for cause L thigh pain	Homocystinuria, Systemic, erythematosus, steroids, years, Onset, L, thigh, pain, leg, ROM, L, hip, steroids, hip, pathology, cause, L, thigh, pain,
66 year old woman presented acutely with cholecystitis and perforated gallbladder. Percutaneous cholecystostomy inserted 10/12/10 and discharged 16/12/10. Represented with RUQ pain, pus discharge from cholecystostomy and raised inflammatory markers and der	year, woman, cholecystitis, gallbladder, Percutaneous, RUQ, pain, pus, discharge, cholecystostomy, markers, der,
70 y.o. man with R. leg claudication. Has had 2 x SFA angioplasties with no symptomatic improvement.MRA to reassess vasculature ?reassess R. SFA stenoses in light of 2 x angioplasties with no symptomatic improvement	man, R, leg, claudication, SFA, improvement, vasculature, R, SFA, light, improvement,

Background: 73yo woman with background of locally advanced rectal cancer known to oncology and being reviewed for chemo-radiation. Currently: Presented dizzy with vacant spell and collapsed but no LOC. Denies slurred speech/visual change prior to collapse	Background, woman, background, cancer, oncology, chemo-radiation, Presented, spell, LOC, Denies, change, collapse,
Sudden onset right thigh & leg pain 5 month ago, radiates on medial aspect, has very prominent right leg VV, claudication symptoms right leg for 5 months: on examination prominent right femoral & popliteal pulse, absent pedal pulses ABPI 0.57 right & 0.53 l	Sudden, onset, right, thigh, leg, pain, month, radiates, aspect, right, leg, VV, claudication, symptoms, leg, months, examination, femoral, pulse, pulses, ABPI, right, l,
myelomamid spine pain Collapse T10 ?additional spine changes/cord compression	spine, spine, compression,
Pt w/ Left cavernous angioma, needs Rpt MRI/MRA to evaluate for interval change Pt w/ Left cavernous angioma, needs Rpt MRI/MRA to evaluate for interval change	Pt, w, angioma, Rpt, MRA, change, Pt, w, angioma, Rpt, MRA, change,
Pt w/ Cav mal, needs MRI brain to assess for interval change Pt w/ Cav mal, needs MRI brain to assess for interval change	Pt, w, Cav, mal, MRI, brain, change, Pt, w, Cav, mal, MRI, brain, change,
Metastatic medulloblastoma (M2). Residual tumour post operatively. Completed RT 11/01/11. Assessment prior to commencing chemotherapy. ?current evidence of disease	Metastatic, medulloblastoma, M2, tumor, post, Completed, Assessment, evidence, disease,
Bilateral renal lesions-m right wilms tumour- left nephrogenic rest. Treated as per Protocol AREN 0534 - not on study. Last chemotherapy 12/01/11 Disease surveillance	lesions-m, tumor, rest, Protocol, AREN, 0534, study, chemotherapy, Disease, surveillance,
investigated in sept 2010 for dizziness & LOC, noted to have builateral carotid bruit & had duplex done at NSH showing bilateral ICA stenoses of 80-99% but EDV do not match upto that criteria, status of each ica artery stenoses	sept, dizziness, LOC, bruit, duplex, NSH, ICA, stenoses, %, EDV, upto, criteria, status, ica, artery,
55 year old lady from Afghanistan with radicular lumbar pain. 3-4 year hx of increasingly severe back pain, radiating down left leg-->toes with associated numbness; On exam, decreased sensation S1, normal power and tone. On xray, arthrosis of L4/5, L5/S1.	year, lady, Afghanistan, radicular, pain, year, hx, pain, leg, toes, numbness, exam, sensation, S1, power, tone, xray, arthrosis, 5,
Pituitary adenoma resection 21/12/10 followed by post op bilateral anterior choroid artery infarcts. Ongoing poor/no recovery. GCS E-4, V-trachy, M6; no movement right side. MRI to assess for ?evolution of infarct, other pathology as clinical picture not i	adenoma, resection, post, op, anterior, artery, infarcts, recovery, GCS, E-4, V-trachy, M6, movement, side, MRI, evolution, infarct, pathology, picture,
15 min appt normal shunted hydrocephalus, needs fast MRI in conjunction with his clinic visit 19/5/2011 status of ventricles	min, hydrocephalus, MRI, conjunction, clinic, status, ventricles,
S/P RESECTION OF COLLOID CYST, needs fast MRI brain in conjunction with visit 10/3/2011 ventricular size	P, RESECTION, COLLOID, CYST, MRI, brain, conjunction, visit, ventricular, size,
FIGO Stage IB1 Ca cervix - completed definitive chemo/XRT 29/11/2010 ?response to treatment	FIGO, Stage, IB1, cervix, XRT, response, treatment,
normal shunted hydrocephalus, needs fast brain MRI in conjunction with clinic visit in february/march ventricular size	hydrocephalus, brain, MRI, conjunction, clinic, ventricular, size,
normal shunt hydrocephalus, needs well child fast brain MRI ventricular size	hydrocephalus, child, brain, MRI, ventricular, size,
29 year old female, presented with bilateral optic nerve head swelling. No history of renal impairment. ? Space Occupying Lesion? Venous Sinus Thrombosis/Stenosis? Contraindication for LP	year, female, nerve, history, impairment, Space, Occupying, Lesion, Venous, Stenosis, LP,
s/p resection of pituitary adenoma, needs MRI brain + contrast, pituitary sequencing. For f/u appt in february, please coordinate with appt. progression/residual pituitary tumour?	p, resection, adenoma, MRI, brain, +, contrast, appt, February, appt, tumor,
1 year history of palpitations and syncope preceding aborted sudden cardiac death, with further malignant arrhythmias in hospital. Prolonged QT interval. Is there any evidence of ischaemia in the territory of the regional wall motion abnormalities seen on	year, history, palpitations, death, arrhythmias, hospital, Prolonged, evidence, territory, wall, motion, abnormalities,
L asym SNHL, tinnitus, imbalance, increasing severe headaches 1/12, off work, on analgesics URGENT?SOL, ?CPA lesion	L, asym, SNHL, tinnitus, imbalance, headaches, URGENT, CPA, lesion,
14 weeks old baby (ex prem 31 weeks, currently 2 months corrected age) with history of apnoea and bradycardia of newborn, currently on home oxygen. Plan is to trial wean off on next admission exclude central cause (brain stem/cerebral)	weeks, baby, prem, weeks, months, age, history, apnoea, bradycardia, home, oxygen, Plan, trial, admission, cause, brain, cerebral,
R asym SNHL ?CPA, lesion	R, CPA, lesion,
Right cervical HD with intracranial extension. Transformed into Diffuse large Bell Cell Lymphoma. Treated with auto transplant Sept 2010. PET scan negative at end of therapy. Routine follow up with MRI Routine follow	HD, extension, Diffuse, Bell, Cell, Lymphoma, auto, transplant, Sept, PET, end,

up of right cervical, mediastinal ma	therapy, MRI, Routine, ma,
Adult onset complex partial seizures with secondary generalization on 5 occasions. Normal exam with no epilepsy risk factors. EEG awaited. structural pathology e.g. MTS or cavernoma?	onset, seizures, generalization, occasions, exam, epilepsy, factors, EEG, pathology, g, MTS, cavernoma,
bilat fem pop bypasses, iliac stents. Presents with several weeks of left leg claudication, now 2-3 days of constant left foot pain, loss of pulses. O/E - cap refill preserved, audible but sluggish PT pulse on doppler only. Vascular service request MRA ?	fem, bypasses, stents, Presents, weeks, leg, claudication, days, foot, pain, loss, pulses, E, cap, refill, PT, pulse, doppler, Vascular, service, request,
SEARCH CLINICAL TRIAL - UNIVERSITY MRI - CYCLE 2 - DAY 1 - CHEST/ABDOMEN/PELVIS - EXACT DATE REQUIRED 04/02/2010 MRI - CHEST/ABDOMEN/PELVIS	SEARCH, CLINICAL, TRIAL, UNIVERSITY, MRI, CYCLE, DAY, PELVIS, EXACT, DATE, REQUIRED, MRI, PELVIS,
Clinical evidence of Vaginal recurrence of previous endometrial carcinoma-biopsy proven Extent of pelvic disease	evidence, Vaginal, recurrence, carcinoma, Extent, disease,
Liver Transplant Assessment 1st February to 4th February HCVHCC ?Bony Mets	Liver, Assessment, February, February, HCVHCC, Bony, Mets,
R asym SNHL ? CPA lesion	R, CPA, lesion,
52yo non-english speaking man (Mandarin only) stood on a nail on 31st of December, took abx he had at home and eventually saw GP on Jan 6th when pain worsening. Given tetanus shot and augmentin for 7 days, and diclofenac. Revisted doctor on 10th of Jan	speaking, man, nail, 31st, December, abx, home, GP, Jan, pain, tetanus, shot, augmentin, days, doctor, 10th, Jan,
Delayed implant reconstruction of left breast following mastectomy for cancer. USS in October suggested an implant rupture and this correlates with clinical finding I the implant ruptured because if it has it needs replacing	reconstruction, breast, cancer, USS, October, rupture, correlates, finding, implant,
25+3 -> 31+6 hydrocephalus -> shunt placed 6/1/11Head uss post-op showed large areas of necrotic brain ? large subependymal nodules. No signs of infection.? cause Could we please organise for MRI head tomorrow prior to transfer back to Middlemore Hospital	uss, post-op, areas, brain, subependymal, signs, infection, cause, organise, MRI, head, tomorrow, Middlemore, Hospital,
R temporal mass excised 30.3.10 carcinoma from unknown primary. CT staging March and July 10 (NSH) - growth of RUL nodule and new R adrenal mass in July - likely mets. Now the adrenal mass disappeared in October 2010. Needs repeat surveillance scan in 4 w	R, mass, carcinoma, primary, CT, March, July, NSH, growth, RUL, R, adrenal, mass, July, mets, mass, October, Needs, surveillance, w,
Pulmonary TB on treatment since May 2010Back pain and previous abnormal thoracic spine views with ?paraspinal collection. Has previously declined to attend to definitive immaging but now agrees to attend. Remains on TB treatment ?TB vertebral osteomyeliti	TB, treatment, May, pain, spine, collection, immaging, Remains, TB, treatment, TB, osteomyeliti,
2.5 yo female, tethered cord, needs MRI lumbar spine no contrast, under GA for pre-op assessment. Please perform prior to the february 10th clinic. cord tethering	yo, female, cord, MRI, lumbar, spine, contrast, GA, assessment, February, clinic, cord, tethering,
Metastatic prostate cancer to pelvic lymph nodes and bones. Has significant c-spine and base of skull discomfort. The pain radiates to the R. neck posterior skull and in V2/3 trigeminal distribution. ?? base of skull or c-spine involvement.Please book for	prostate, cancer, lymph, nodes, bones, c-spine, base, skull, discomfort, pain, R, neck, posterior, skull, distribution, base, skull, involvement, book,
OLT 1989CRF & cardiomyopathyCr 203 Known panceratic cystLast imaged Oct 2009Recommended 6/12 f/u to assess growth	OLT, 1989CRF, cardiomyopathyCr, cystLast, Oct, u, growth,
6 mo female with tethered cord on newborn MRI. Please perform MRI L-spine no GA, no contrast prior to March 2011 for surgical planning cord tethering	mo, cord, MRI, MRI, L-spine, GA, contrast, March, planning, cord, tethering,
Hx of the right leg intermittent claudication for a few weeks. haevy smoker, diabetic. No pedal pulses on the right side. Weak doppler sound for PTA. Discussed with Mr Bouchier, needs urgent MRA as an outpatient and review in clinic with the result PVD,	Hx, leg, claudication, weeks, smoker, pulses, side, Weak, PTA, Discussed, Mr, Bouchier, MRA, outpatient, review, clinic, result, PVD,
30 years ago, had tibial plateau fracture, 3 years later plate removed.Niggling pain since with occasional locking.Today, at beach, wiping feet, knee locked and presented to ED.o/e: no instability 20-110 degree ROM with soft blockingNo knee effusion.x-ray	years, plateau, fracture, years, plate, pain, Today, beach, feet, knee, ED, e, instability, degree, blockingNo, knee, x-ray,
diffuse lymphadenopathy noted, ?lymphoma. occipital lesion with some bleeding into lesion, needs operative treatmentMRI STEALTH please to guide operative management MRI STEALTH protocol to guide operative management	lymphadenopathy, lesion, treatmentMRI, STEALTH, management, MRI, STEALTH, protocol, management,
57 yo (Cantonese Only) lady with 2/7 hx of left sided hip pain radiating to groin and left abdomen. Temp 37.8 BP 96/55, WCC 13.1, CRP 201.CT abdo provisional report - low density change deep within the gluteus medius muscle, with likely infective or inf	yo, lady, hx, left, hip, pain, groin, abdomen, Temp, BP, WCC, CT, report, density, change, medius, muscle, infective, inf,
6/12 with acute flacid paralysis ? causeGrowing enterovirus from throat & rectal swabFlacid paralysis noticed afetr LP. Truncal & limg hypotoniaCentral pain & blindnessReviwed by neurologist suggest Spinal	paralysis, enterovirus, throat, paralysis, afetr, LP, Truncal, limg, pain, blindnessReviwed,

MRI & also reimaging brain ? Spinla bruise/ haema	neurologist, Spinal, MRI, brain, Spinla, bruise, haema,
Pt w/ meningioma, needs MRI +/- for ? change Pt w/ meningioma, needs MRI +/- for ? change	Pt, w, meningioma, MRI, +, change, Pt, w, meningioma, MRI, +,
Lump over upper back(lower cervical and upper thoracic spine) since 15 but gradually enlarging over the past year, associated with pain over last 6-7 months. Had USS of lump in Dec 2010 but MRI recommended as lump was too large to determine its relations	Lump, back, spine, year, pain, months, USS, lump, Dec, MRI, lump, relations,
History of L idiopathic apical orbit syndrome. Prev MRI orbit showed swelling of sup rectus, med rectus, sup oblique, Lat rectus with optic disc compression. Presents with reduced VA and colour vision left eye with worsening proptosis. RAPD LE. D/w Neuror	History, L, orbit, syndrome, Prev, MRI, orbit, sup, rectus, med, rectus, rectus, disc, compression, Presents, VA, color, vision, eye, proptosis, RAPD, w, Neuror,
20 year old hair dresser, 1- 2 month history of left hip weakness and intermittent pain, recently exacerbated after striding into a shallow pool and landing heavily on both feet - felt a jarring in her back and increasing left hip pain. MRI of her spine	year, hair, dresser, month, history, hip, weakness, pain, pool, feet, back, hip, pain, MRI, spine,
Mixed germ cell tumour PA nodes BEP chemotherapy liver lesion ?nature liver specific MRI please ?liver lesion involved by cancer	germ, cell, tumor, PA, nodes, liver, lesion, nature, liver, MRI, liver, lesion, cancer,
85yo man with L)sided weakness, CT head done 13/1/11 - ?glioma as per Dr. Moriarty. From Masterton hospital, CT head from Oct 2010 at initial presentation on PACS. MRI head for further investigation of ?glioma. Cr 88. High grade glioma as definitive diagn	man, L, weakness, CT, head, glioma, Dr, Moriarty, Masterton, hospital, CT, Oct, presentation, PACS, MRI, investigation, grade, diagn,
Found collapsed on bathroom floor yesterday 5pm ??duration. Initial reports show GCS 14 and mild R) sided weakness. Review this morning - GCS 5. Bilateral upgoing plantars. Increased tone bilaterally - L>R. ??brainstem infarct vs metabolic disorder.	Found, bathroom, floor, yesterday, duration, reports, GCS, R, weakness, Review, morning, GCS, upgoing, tone, bilaterally, L, R, ??, brainstem, disorder,
To re-image left breast following mammography for surveillance Further imaging following mammography	breast, surveillance, imaging,
Transient episode of right arm weakness and slurred speech yesterday - resolved. Background of untreated type II diabetes. Newly diagnosed hypertension (230/130) and fast AF. Also has impaired LV systolic function on ECHO but no thrombus. Likely TIA but	episode, arm, weakness, speech, yesterday, Background, type, II, diabetes, hypertension, AF, LV, function, ECHO, thrombus,
3 month h/o multiple joint/bone pain and malaise.multiple antibiotics iv.fluctuating inflammatory markers. any lesions ?malignancy	month, pain, malaise, antibiotics, markers, malignancy,
patient presented with left leg cellulitis.On IV flucloz for > a week.Still spiking temperature and has large area of erythmatous lesion on lower leg left side.? osteomyelitis ? Osteomyelitis	patient, leg, IV, week, temperature, area, erythmatous, leg, side, Osteomyelitis,
Had staging CT for what turned out to be benign palatal lesion under ENT. Incidental finding of 20mm lesion segment 7 liver -subcapsular with some ring enhancement. Other liver lesions look cystic. felt subcapsular lesion difficult to visualize on uss in	CT, lesion, ENT, finding, lesion, segment, liver, subcapsular, ring, enhancement, liver, lesions, subcapsular, uss,
6 yr boy with 1st generalised tonic-clonic seizure on 29 Nov. No obvious focal seizure. No signs of raised intracranial pressure. Normal neuro exam. EEG showed right temporal focus. ?any brain dysplasia or tumour in right temporal region (focal abnorm	yr, boy, seizure, Nov, seizure, signs, pressure, exam, EEG, focus, brain, dysplasia, tumor, region, abnorm,
Nasopharyngeal carcinoma left posterior nose. Completed chemo/radiotherapy on 1/10/10. Residual disease still present. Case was discussed at the head and neck radiology meeting on the 14/1/11 Delianate residual tumour. ?intracranial spread on left side.	carcinoma, posterior, nose, disease, Case, head, neck, meeting, tumor, intracranial, side,
Desmoplastic small round cell tumour of abdo, mets to liver. Debulked 95% 31st august. Remaining pelvic component irradiated completed 14th Jan 2011. Last chemotherapy sept 2010. Disease status six weeks post radiation.	round, cell, tumor, abdo, liver, %, August, component, Jan, chemotherapy, Disease, status, weeks, radiation,
Average risk medulloblastoma. post resection and RT. Midway through chemotherapy. Recent overshunting and shunt revised 14/01/11. ?evidence of progressive disease	risk, medulloblastoma, resection, RT, Midway, chemotherapy, overshunting, shunt, evidence, disease,
38 year old woman intermittent episodes of vertigo, clumsiness and sensory symptoms in the right hand ? Evidence of demyelination	year, woman, episodes, vertigo, clumsiness, symptoms, hand, Evidence, demyelination,
R T3N3M0 SCC of left tonsil. For combined chemo-radiotherapy treatment. ? extent of disease.For radiotherapy planning.	R, SCC, tonsil, treatment, extent, disease, planning,
SEARCH - ELIGIBILITY SCAN - CHEST/ABDOMEN/PELVISEXACT DATE REQUIRED 02/02/2011 SEARCH TRIAL - ADVANCED HCC - ELIGIBILITY MRI - CHEST/ABDOMEN/PELVIS	SEARCH, ELIGIBILITY, SCAN, PELVISEXACT, DATE, REQUIRED, SEARCH, TRIAL, ADVANCED, HCC, ELIGIBILITY, MRI, PELVIS,
Known metastatic prostate cancer with widespread bone disease. Has severe pain in L hip and occasionally radiating down L. leg. Neurological	prostate, cancer, bone, disease, pain, L, hip, L, leg, exam, pain,

exam grossly normal, but somewhat limited by pain. ? nerve root impingement that could be amenable to radiotherapy.	nerve, root, radiotherapy,
several months of neck and right arm pain with clinical findings suggestive of right C6 radiculopathy, with EMG showing mild reinnervation of right C5/6 muscles. C6 radiculopathy?	months, neck, arm, pain, findings, C6, EMG, reinnervation, muscles,
Frequent runs of asymptomatic ventricular tachycardia, rate up to 213bpm, monomorphic, but occurs during exercise as well as at rest. First noted under anaesthesia for tonsillectomy. No family history of arrhythmia or sudden death. Unable at this time to	runs, ventricular, tachycardia, exercise, rest, First, anaesthesia, tonsillectomy, family, history, arrhythmia, death, time,
T1N1 NPC. For definitive radiotherapy +/- chemo. ? extent of diseaseFor radiotherapy planning purpose.	T1N1, radiotherapy, +, chemo, extent, diseaseFor, purpose,
T4N2M0 NPC. For definitive radiotherapy alone. ? extent of disease.For radiotherapy planning.	T4N2M0, radiotherapy, extent, disease, planning,
19 year old female with new onset fecal incontinence on background recent T12-L2 fusion Aug 2010 for L1 burst #. O/E reduced anal tone with empty rectum; sensation preserved. No LL weakness or sensory changes. To exclude cauda equina compression	year, onset, incontinence, background, T12-L2, fusion, Aug, L1, burst, NN, E, tone, rectum, sensation, weakness, changes, equina, compression,
Patient has 3 month history of gradual decline. T2 hyperintensity on MRI scan. Please do MRI t2 stealth with and without contrast before theatre Stealth MRI, patient for needle biopsy	Patient, month, history, decline, T2, MRI, MRI, stealth, contrast, theater, Stealth, MRI, patient, needle, biopsy,
ECG complete heart block, ?infiltrative cardiomyopathy	heart, block, cardiomyopathy,
Severe hypertension with evidence of end organ damage, mildly impaired renal dysfunction ? Renal artery stenosis	hypertension, evidence, end, organ, damage, dysfunction, artery,
Mr Jones 58 years old with staph bacterial endocarditis: hx presented to MMH 2nd Jan 2011 with fevers malaise night sweats chills rigors and decreased appetite. CRP on admission 320. WCC 8.7. Deranged LFTs. Initially thought to have encephalitis so st	Mr, Jones, years, endocarditis, hx, Jan, fevers, malaise, night, rigors, appetite, admission, Deranged, encephalitis,
40 y/o presents with 1/52 worsening lumbar back pain on a history of chronic lumbar back pain. Spondylosis L4/L5 on previous imaging. On waitlist for fusion. Multiple high volume post-void residual volumes, 390ml /430ml. MSU clear, bloods NAD. Reques	o, lumbar, pain, history, lumbar, pain, Spondylosis, imaging, fusion, volume, volumes, 430ml, MSU, bloods, Reques,
RUQ pain, previous cholecystectomy.Deranged LFT's.USS no stone RUQ pain, previous cholecystectomy.Deranged LFT's.USS - no cholidocholithiasisplease assess the biliary system.	RUQ, pain, cholecystectomy, LFT, USS, stone, pain, cholecystectomy, LFT, USS, cholidocholithiasisplease, system,
Presented with Shingles like lesion over the left ear, neck and scalp 6 days ago with sudden onset of lower motor neurone seventh nerve palsy ont he left face 4 days ago. Patient subsequently developed loss of vision on the left eye to counting finger 2 d	Shingles, lesion, ear, neck, scalp, days, onset, motor, neurone, nerve, palsy, ont, face, days, Patient, loss, vision, eye, finger,
44y old with metastatic NSCLC with previously demonstrated spinal leptomeningeal disease recently treated with radiotherapy. Presented yesterday with confusion, headache, anorexia and hypertension (230/120). Unable to co-operate with exam but is relative	NSCLC, disease, radiotherapy, yesterday, confusion, headache, anorexia, hypertension, exam,
8 year old patient admitted with tenderness left ankle.Sudden onset. No history of trauma. Unable to weight bear. No history of viral infection. O/E left ankle hot, swollen. Decreased ROM. Tender anteriorly. Pain on dorsiflexion and inversion.ESR 138 CRP	year, patient, tenderness, ankle, onset, history, trauma, weight, bear, history, infection, ankle, Decreased, Tender, Pain, dorsiflexion, inversion, ESR, CRP,
Bicuspid aortic valve. Coarctation of the aorta. Staph. aureus endarteritis completing antibiotics 17/01/2010. ? severity of coarctation ? evidence for endarteritis	valve, Coarctation, aorta, aureus, antibiotics, severity, coarctation, evidence, endarteritis,
54 yr old lady on PD (ESRF) admitted with R) 5th toe infection and positive blood culture(gram positive cocci) - o/e febrile and small necrotic tissue underneath and purulent fluid on the 5th toe with surrounding cellulitis. No peripheral pulses except fe	yr, lady, PD, ESRF, R, toe, infection, blood, culture, gram, cocci, febrile, tissue, purulent, toe, cellulitis, pulses, fe,
Right leg redness. Absent distal pulses in both feet. Absent right popliteal. Bilateral femoral pulses palpable..? Calf claudication in the past. Exclude lower limb ischemia. Bilateral leg ishaemia, Bilateral absent distal pulses. Redness of right leg ? i	leg, redness, distal, feet, popliteal, femoral, palpable, claudication, past, limb, ischemia, leg, ishemia, pulses, Redness, leg,
Thank you. 36yo woman with GBM. Acute periorbital swelling with proptosis. ? orbital cellulitis ? periorbital cellulitis. ? cavernous sinus thrombosis from bacteraemia from multiple abscess from thigh. ? thrombosis ? clot ? abscess	woman, GBM, Acute, proptosis, orbital, periorbital, sinus, bacteremia, abscess, thigh, abscess,
R hemifacial spasm, referred to neurosurgery Rule out CPA or other posteruior fossa lesions casing hemifacial spasm, ?vascular loop	R, spasm, Rule, CPA, posteruior, fossa, lesions, spasm, vascular, loop,

Appendix 6: Examples of keywords and the number of related concept identified in SNOMED CT

Extracted single keyword	Concept count from SNOMED CT
cancer	1939
XRT	12
surgery	1776
Repeat	261
MRI	791
response	1131
MRI	791
March	55
response	1131
XRT	12
woman	180
month	550
thigh	1366
mass	1381
month	550
lethargy	62
month	550
fevers	33
weight	1216
loss	3256
Exam	1797
firm	58
mass	1381
medialis	88
adductor	234
Stage	3058
III	4332
post	757
study	1901
disease	20835
status	2713
cause	869
fossa	904
lesion	7169
Multiple	4007
months	377
status	2713
end	905
MRI	791
spine	4035
leg	3403
weakness	268

cord	4451
concerns	6
spine	4035
team	189
problem	2268
HLH	8
disease	20835
liver	3303
spleen	1045
lung	3773
skin	13880
evidence	196
disease	20835
grade	1303
wing	86
neuritis	305
New	707
Local	2601
end	905
surgery	1776
Needs	274
chest	2202
abdo	78
pelvis	1846
MRI	791
pelvis	1846
response	1131
response	1131
Pineal	165
tumor	16460
January	26
status	2713
tumor	16460
week	558
baby	1206
flexor	931
spasms	44
spasms	44
Needs	274
MRI	791
part	4025
work	823
brain	2814
report	968
HCV	40
USS	35

Oct	46
liver	3303
MRI	791
SCC	108
Anus	841
Assess	733
response	1131
tumor	16460
blood	14141
failure	2284
revision	1677
MRI	791
brain	2814
size	1302
Fast	401
MRI	791
revision	1677
arrest	442
CAD	36
cause	869
arrest	442
Systemic	1095
steroids	148
years	351
Onset	1010
thigh	1366
pain	3444
leg	3403
ROM	19
hip	3100
steroids	148
hip	3100
cause	869
thigh	1366
pain	3444
year	763
woman	180
RUQ	3
pain	3444
pus	134
markers	38
der	56
man	247
leg	3403
SFA	36
SFA	36

light	1110
woman	180
cancer	1939
oncology	141
spell	10
LOC	31
Denies	2
change	1054
collapse	551
Sudden	186
onset	1010
right	5836
thigh	1366
leg	3403
pain	3444
month	550
aspect	625
right	5836
leg	3403
symptoms	1471
leg	3403
months	377
femoral	1925
pulse	1101
pulses	171
ABPI	13
right	5836
spine	4035
spine	4035
angioma	198
MRA	63
change	1054
angioma	198
MRA	63
change	1054
Cav	30
mal	150
MRI	791
brain	2814
change	1054
Cav	30
mal	150
MRI	791
brain	2814
change	1054
post	757

evidence	196
disease	20835
rest	286
Protocol	113
study	1901
Disease	20835
sept	9
LOC	31
bruit	150
duplex	95
ICA	21
stenoses	6
upto	28
criteria	33
status	2713
ica	21
artery	11877
year	763
lady	12
pain	3444
year	763
pain	3444
leg	3403
toes	334
numbness	54
exam	1797
power	547
tone	293
xray	2474
adenoma	838
post	757
anterior	5322
artery	11877
infarcts	13
recovery	155
GCS	26
movement	2177
MRI	791
infarct	358
picture	49
min	175
MRI	791
clinic	1410
status	2713
COLLOID	168
CYST	2721

MRI	791
brain	2814
visit	824
size	1302
FIGO	596
Stage	3058
IB1	9
cervix	2145
XRT	12
response	1131
brain	2814
MRI	791
clinic	1410
size	1302
child	3401
brain	2814
MRI	791
size	1302
year	763
female	3161
nerve	9719
history	5994
Space	1580
Lesion	7169
Venous	2139
Stenosis	1501
adenoma	838
MRI	791
brain	2814
contrast	1609
appt	18
appt	18
year	763
history	5994
death	1263
hospital	1587
evidence	196
wall	2959
motion	1379
SNHL	15
tinnitus	91
URGENT	233
CPA	25
lesion	7169
weeks	659

baby	1206
prem	28
weeks	659
months	377
age	807
history	5994
apnoea	302
home	2010
oxygen	760
Plan	297
trial	210
cause	869
brain	2814
cerebral	2930
CPA	25
lesion	7169
Diffuse	1297
Bell	30
Cell	11558
Lymphoma	1751
auto	4
Sept	9
PET	288
end	905
therapy	22512
MRI	791
Routine	225
onset	1010
seizures	234
exam	1797
epilepsy	648
factors	294
EEG	286
MTS	4
fem	41
stents	7
weeks	659
leg	3403
days	337
foot	5394
pain	3444
loss	3256
pulses	171
cap	315
refill	140
pulse	1101

doppler	269
Vascular	2428
service	2974
request	159
SEARCH	10
CLINICAL	1731
TRIAL	210
MRI	791
CYCLE	1227
DAY	1358
PELVIS	1846
DATE	288
REQUIRED	100
MRI	791
PELVIS	1846
evidence	196
Vaginal	2737
Extent	118
disease	20835
Liver	3303
Bony	471
CPA	25
lesion	7169
speaking	54
man	247
nail	1473
home	2010
Jan	10
pain	3444
tetanus	906
shot	126
days	337
doctor	307
Jan	10
breast	4473
cancer	1939
rupture	1658
finding	162449
implant	1459
uss	35
areas	64
brain	2814
signs	511
cause	869
organise	74
MRI	791

head	4758
tomorrow	3
Hospital	1587
mass	1381
primary	6731
March	55
July	26
growth	1337
RUL	8
adrenal	1304
mass	1381
July	26
mass	1381
Needs	274
May	72
pain	3444
spine	4035
Remains	14
female	3161
cord	4451
MRI	791
lumbar	3180
spine	4035
contrast	1609
clinic	1410
cord	4451
prostate	1419
cancer	1939
lymph	4350
nodes	2263
bones	922
base	1087
skull	1659
pain	3444
neck	4516
skull	1659
base	1087
skull	1659
book	38
Oct	46
growth	1337
cord	4451
MRI	791
MRI	791
contrast	1609
March	55

planning	409
cord	4451
leg	3403
weeks	659
smoker	288
pulses	171
Weak	197
PTA	95
MRA	63
review	348
clinic	1410
result	560
PVD	33
years	351
plateau	72
fracture	9664
years	351
plate	544
pain	3444
Today	5
beach	35
feet	307
knee	3340
degree	1609
knee	3340
lesion	7169
MRI	791
protocol	113
lady	12
left	6300
hip	3100
pain	3444
groin	321
abdomen	1371
Temp	180
WCC	6
report	968
density	450
change	1054
medius	55
muscle	8462
inf	164
throat	913
Truncal	173
pain	3444
Spinal	5137

MRI	791
brain	2814
bruise	363
MRI	791
change	1054
MRI	791
Lump	702
back	1481
spine	4035
year	763
pain	3444
months	377
USS	35
lump	702
Dec	9
MRI	791
lump	702
History	5994
orbit	813
syndrome	14230
MRI	791
orbit	813
sup	81
rectus	522
med	304
rectus	522
rectus	522
disc	2027
vision	1029
eye	6152
RAPD	6
year	763
hair	1410
dresser	40
month	550
history	5994
hip	3100
weakness	268
pain	3444
pool	227
feet	307
back	1481
hip	3100
pain	3444
spine	4035

Appendix 7: Examples of experimental searching for (keyword + left-neighbourhood word)

Single Keyword + Left Word	Related Concepts of "SK +LW" (diagnosis only)	Concept
endometrial cancer		
definitive XRT		
have surgery		
comorbidities Repeat		
Repeat MRI		
for response		
Prev MRI		
MMH March		
2010 response		
to XRT		
chinese woman		
1 month		
	5434002	Abortion, 1 month (disorder)
R thigh		
	157535002	Bruise - hip &/or thigh (disorder)
	18814006	Contusion of hip AND/OR thigh (disorder)
	201892008	Osteoarthritis NOS: [pelvic region and/or thigh] or [hip] (disorder)
	203188007	Chronic osteomyelitis of the pelvic region &/or thigh (disorder)
	203251001	Periostitis without osteomyelitis, of the pelvic region and/or thigh (disorder)
	203271005	Tuberculosis of the pelvic and/or thigh bones (disorder)
	203290002	Poliomyelitis osteopathy of the pelvic region and/or thigh (disorder)
	209504009	Sprain: [hip and/or thigh] or [groin] or [hamstring] (disorder)
	209511008	Other thigh sprain (disorder)
	210661006	Open wound of hip and/or thigh (disorder)
	210676003	Degloving injury, hip or thigh (disorder)
	210679005	Severe multi tissue damage hip or thigh (disorder)
	210680008	Massive multi tissue damage hip or thigh (disorder)
	210772007	Multiple open wounds of hip and/or thigh (disorder)
	211383006	Multiple superficial injuries of hip and/or thigh (disorder)
	212440003	Other thigh injuries (disorder)
	268016009	Acute osteomyelitis of the pelvic region and/or thigh (disorder)
	268018005	Chronic osteomyelitis of the pelvic region and/or thigh (disorder)
	269377008	Bruise - hip &/or thigh (disorder)
	287068009	Pathological fracture - pelvis and/or thigh (disorder)
	287076006	Fracture malunion - pelvis and/or thigh (disorder)
	35250003	Crushing injury of hip AND/OR thigh (disorder)
	50079005	Injury of hip AND/OR thigh (disorder)
thigh mass		
one month		
Increased lethargy		

last month		
recent fevers		
or weight		
weight loss		
	71479007	AIDS with abnormal weight loss (disorder)
loss Exam		
5x8cm firm		
firm mass		
vastus medialis		
medialis adductor		
Stage III		
	236415000	Membranous glomerulonephritis - stage III (disorder)
	312703006	Full thickness macular hole stage III (disorder)
Evaluation post		
ANBL0532 study		
current disease		
	58184002	Recurrent disease (disorder)
disease status		
central cause		
cranial fossa		
fossa lesion		
JPA Multiple		
10 months		
current status		
at end		
recent MRI		
MRI spine		
bilateral leg		
leg weakness		
spinal cord		
	10310006	Open fracture of C1-C4 level with incomplete spinal cord lesion (disorder)
	10575009	Open fracture of vertebral column with spinal cord injury (disorder)
	110150000	Spinal cord concussion (disorder)
	111497000	Arterial thrombosis of spinal cord (disorder)
	111622009	Closed fracture of third cervical vertebra without spinal cord injury (disorder)
	111623004	Closed fracture of C1-C4 level with other specified spinal cord injury (disorder)
	111624005	Closed fracture of C5-C7 level with other specified spinal cord injury (disorder)
	111625006	Open fracture of C1-C4 level with other specified spinal cord injury (disorder)
	111626007	Open fracture of C5-C7 level with other specified spinal cord injury (disorder)
	111627003	Closed fracture of T1-T6 level with other specified spinal cord injury (disorder)
	111628008	Closed fracture of T7-T12 level with other specified spinal cord injury (disorder)
	111629000	Open fracture of T1-T6 level with other specified spinal cord injury (disorder)

	111630005	Open fracture of T7-T12 level with other specified spinal cord injury (disorder)
	111632002	Closed fracture of sacrum and coccyx with other spinal cord injury (disorder)
	111635000	Open fracture of sacrum and coccyx with other spinal cord injury (disorder)
	111698000	Injury at C1-C4 level with other specified spinal cord injury and without bone injury (disorder)
	111699008	Injury at C5-C7 level with other specified spinal cord injury and without bone injury (disorder)
	11230003	Laceration of spinal cord due to birth trauma (disorder)
	11413003	Incomplete spinal cord lesion at T1-T6 level without bone injury (disorder)
	11523005	Lumbar spinal cord injury without bone injury (disorder)
	11782000	Closed fracture of second cervical vertebra without spinal cord injury (disorder)
	11807002	Injury at C5-C7 level with spinal cord injury AND without bone injury (disorder)
	125609005	Open fracture of cervical region with spinal cord injury (disorder)
	126962006	Neoplasm of spinal cord (disorder)
	13456001	Closed fracture of T1-T6 level with incomplete spinal cord lesion (disorder)
	13498003	Closed fracture of multiple cervical vertebrae without spinal cord injury (disorder)
	14584002	Closed fracture of sacrum AND/OR coccyx with spinal cord injury (disorder)
	154571009	(Secondary Ca brain/spinal cord) or (bone secondary Ca) or (brain secondary Ca) (disorder)
	155017004	Other diseases of spinal cord (disorder)
	15701003	Open fracture of second cervical vertebra without spinal cord injury (disorder)
	157094006	Spine or spinal cord injury due to birth trauma (disorder)
	157183009	Fracture of vertebra without spinal cord lesion (disorder)
	157185002	Fracture of thoracic spine without spinal cord lesion (disorder)
	157188000	Fracture of vertebra without spinal cord lesion NOS (disorder)
	157189008	Fracture of spine with spinal cord lesion (disorder)
	157194008	Fracture of spine with spinal cord lesion NOS (disorder)
	15724005	Fracture of vertebral column without spinal cord injury (disorder)
	157642005	Spinal cord (& [no fracture]) &/or nerve root injury (disorder)
	157643000	Spinal cord lesion - no # (disorder)
	157647004	Spinal cord/nerve root injury NOS (disorder)
	16300007	Spinal cord abscess (disorder)
	16581008	Injury of spine AND/OR spinal cord as birth trauma (disorder)
	16695002	Open fracture of T1-T6 level with spinal cord injury (disorder)
	1734006	Fracture of vertebral column with spinal cord injury (disorder)
	17463000	Closed fracture of lumbar vertebra without spinal cord injury (disorder)
	188462001	Secondary malignant neoplasm of brain and spinal cord (disorder)
	188465004	Secondary malignant neoplasm of brain or spinal cord NOS (disorder)
	189017000	Lipoma of spinal cord (disorder)
	189166000	Benign tumor of spinal cord (disorder)
	189484008	Neoplasm of uncertain behavior of brain and/or spinal cord (disorder)
	189486005	Neoplasm of uncertain behavior of spinal cord (disorder)
	189489003	Neoplasm of uncertain behavior of brain or spinal cord NOS (disorder)

	18960007	Closed fracture of T7-T12 level with incomplete spinal cord lesion (disorder)
	19090007	Injury at C1-C4 level with complete lesion of spinal cord AND without bone injury (disorder)
	192893003	Other diseases of spinal cord (disorder)
	192897002	Myelopathy due to acute infarction of spinal cord (disorder)
	192898007	Myelopathy due to arterial thrombosis of spinal cord (disorder)
	192899004	Myelopathy due to edema of spinal cord (disorder)
	192913003	Myelopathy NOS (& [spinal cord compression NOS]) (disorder)
	192927008	Multiple sclerosis of the spinal cord (disorder)
	193080004	(Cord compression) or (spinal cord compression) (disorder)
	194483002	[X]Subacute combined degeneration of the spinal cord in diseases classified elsewhere (disorder)
	194556000	[X]Other specified diseases of spinal cord (disorder)
	20377001	Open fracture of C5-C7 level with incomplete spinal cord lesion (disorder)
	204076003	Other specified spinal cord anomalies (disorder)
	204077007	(Other specified anomalies of spinal cord) or (neuroenteric cyst) (disorder)
	204078002	(Atelomyelia) or (myelatelasia) or (myelodysplasia of spinal cord) (disorder)
	204081007	Spinal cord hypoplasia (disorder)
	204082000	Other specified spinal cord anomalies NOS (disorder)
	204083005	Spinal cord anomalies NOS (disorder)
	204095005	Congenital spinal cord anomaly (disorder)
	205852002	[X]Other specified congenital malformations of spinal cord (disorder)
	206220004	Spine or spinal cord injury due to birth trauma (disorder)
	206223002	Spinal cord laceration due to birth trauma (disorder)
	206224008	Spinal cord rupture due to birth trauma (disorder)
	206225009	Spine or spinal cord injury due to birth trauma NOS (disorder)
	207885009	(Fracture without spinal cord lesion: [vertebra] or [transverse process of spine]) or (fracture of spine without mention of spinal cord injury) (disorder)
	207886005	Closed fracture: [cervical spine] or [atlas without spinal cord lesion] or [axis without spinal cord lesion] (disorder)
	207888006	Closed fracture: [atlas] or [first cervical vertebra without mention of spinal cord lesion] (disorder)
	207889003	Closed fracture: [axis] or [second cervical vertebra without mention of spinal cord lesion] (disorder)
	207890007	Closed fracture of third cervical vertebra (& [without mention of spinal cord lesion]) (disorder)
	207891006	Closed fracture of fourth cervical vertebra (& [without mention of spinal cord lesion]) (disorder)
	207892004	Closed fracture of fifth cervical vertebra (& [without mention of spinal cord lesion]) (disorder)
	207893009	Closed fracture of sixth cervical vertebra (& [without mention of spinal cord lesion]) (disorder)
	207894003	Closed fracture of seventh cervical vertebra (& [without mention of spinal cord lesion]) (disorder)
	207912004	Open fracture: [cervical spine] or [atlas without spinal cord lesion] or [axis without spinal cord lesion] (disorder)
	207914003	Open fracture: [atlas] or [open fracture first cervical vertebra without mention of spinal cord lesion] (disorder)
	207915002	Open fracture: [axis] or [second cervical vertebra without mention of spinal cord lesion] (disorder)
	207916001	Open fracture of third cervical vertebra (& [without mention of spinal cord lesion]) (disorder)
	207917005	Open fracture of fourth cervical vertebra (& [without mention of spinal cord lesion]) (disorder)
	207918000	Open fracture of fifth cervical vertebra (& [without mention of spinal cord lesion]) (disorder)

	207919008	Open fracture of sixth cervical vertebra (& [without mention of spinal cord lesion]) (disorder)
	207920002	Open fracture of seventh cervical vertebra (& [without mention of spinal cord lesion]) (disorder)
	207996002	Fracture of spine without mention of spinal cord lesion NOS (disorder)
	207997006	Fracture of spine with spinal cord lesion (& [transverse process]) (disorder)
	208011006	Closed fracture of cervical spine with spinal cord lesion NOS (disorder)
	208012004	Open fracture of cervical spine with spinal cord lesion (disorder)
	208025007	Open fracture of cervical spine with spinal cord lesion NOS (disorder)
	208026008	Closed fracture of thoracic spine with spinal cord lesion (disorder)
	208039007	Closed fracture of thoracic spine with spinal cord lesion NOS (disorder)
	208040009	Open fracture of thoracic spine with spinal cord lesion (disorder)
	208053007	Open fracture of thoracic spine with spinal cord lesion NOS (disorder)
	208054001	Closed fracture of lumbar spine with spinal cord lesion (disorder)
	208061002	Open fracture of lumbar spine with spinal cord lesion (disorder)
	208069000	Closed fracture of sacrum with spinal cord lesion (disorder)
	208070004	Closed fracture of sacrum with unspecified spinal cord lesion (disorder)
	208073002	Closed fracture of sacrum with other spinal cord injury (disorder)
	208075009	Closed fracture of sacrum with spinal cord lesion NOS (disorder)
	208076005	Open fracture of sacrum with spinal cord lesion (disorder)
	208077001	Open fracture of sacrum with unspecified spinal cord lesion (disorder)
	208080000	Open fracture of sacrum with other spinal cord injury (disorder)
	208081001	Open fracture of sacrum with spinal cord lesion NOS (disorder)
	208082008	Closed fracture of coccyx with spinal cord lesion (disorder)
	208083003	Closed fracture of coccyx with unspecified spinal cord lesion (disorder)
	208086006	Closed fracture of coccyx with other spinal cord injury (disorder)
	208087002	Closed fracture of coccyx with spinal cord lesion NOS (disorder)
	208088007	Open fracture of coccyx with spinal cord lesion (disorder)
	208089004	Open fracture of coccyx with unspecified spinal cord lesion (disorder)
	208092000	Open fracture of coccyx with other spinal cord injury (disorder)
	208093005	Open fracture of coccyx with spinal cord lesion NOS (disorder)
	208094004	Closed fracture of spine with spinal cord lesion unspecified (disorder)
	208095003	Open fracture of spine with spinal cord lesion unspecified (disorder)
	208096002	Fracture of spine with spinal cord lesion NOS (disorder)
	21167003	Tuberculoma of spinal cord (disorder)
	212140008	(Injuries: [nerve and/or spinal cord] or [cranial nerve]) or (neuroma - traumatic) (disorder)
	212154000	Spinal cord injury without evidence of spinal bone injury (disorder)
	212194003	Spinal cord injury of multiple sites without spinal bone injury (disorder)
	212195002	Spinal cord injury without spinal bone injury NOS (disorder)
	212355006	Injury of nerves and spinal cord at neck level (disorder)
	212356007	Concussion and edema of cervical spinal cord (disorder)

	212357003	Injury of nerves and spinal cord at thorax level (disorder)
	212358008	Concussion and edema of thoracic spinal cord (disorder)
	212359000	Injury of nerves and lumbar spinal cord at abdomen, lower back and pelvis level (disorder)
	212360005	Concussion and edema of lumbar spinal cord (disorder)
	212368003	Nerve and spinal cord injury NOS (disorder)
	213383004	Injuries of brain and cranial nerves with injuries of nerves and spinal cord at neck level (disorder)
	213426008	[X]Other and unspecified injuries of cervical spinal cord (disorder)
	213444000	[X]Other and unspecified injuries of thoracic spinal cord (disorder)
	213462009	[X]Other injury of lumbar spinal cord (disorder)
	213591009	[X]Injuries of nerves and spinal cord of other multiple body regions (disorder)
	21573009	Closed fracture of seventh cervical vertebra without spinal cord injury (disorder)
	22973003	Closed fracture of vertebral column with spinal cord injury (disorder)
	230141006	Spinal cord compression NOS (disorder)
	230739000	Spinal cord stroke (disorder)
	230741004	Venous infarction of spinal cord (disorder)
	23283004	Open fracture of T1-T6 level with incomplete spinal cord lesion (disorder)
	241992009	Spinal cord decompression injury (disorder)
	24392008	Injury at C1-C4 level with spinal cord injury AND without bone injury (disorder)
	253124001	Myelodysplasia of the spinal cord (disorder)
	25331007	Injury at T7-T12 level with spinal cord injury AND without bone injury (disorder)
	254946004	Glial tumor of spinal cord (disorder)
	254947008	Glioma of spinal cord (disorder)
	254948003	Astrocytoma of spinal cord (disorder)
	254949006	Ependymoma of spinal cord (disorder)
	254950006	Oligodendroglioma of spinal cord (disorder)
	254951005	Mixed glial tumor of spinal cord (disorder)
	254953008	Neuronal and mixed neuronal - glial tumor of spinal cord (disorder)
	254954002	Embryonal tumor of spinal cord (disorder)
	262696004	Contusion of spinal cord (disorder)
	262705002	Laceration of spinal cord (disorder)
	262710003	Transection of spinal cord (disorder)
	262715008	Injury of spinal cord vasculature (disorder)
	262718005	Traumatic spinal cord hemorrhage (disorder)
	263178004	Fracture of transverse process of spine without spinal cord lesion (disorder)
	263179007	Fracture of transverse process of spine with spinal cord lesion (disorder)
	263181009	Closed fracture of atlas without spinal cord lesion (disorder)
	263182002	Closed fracture of axis without spinal cord lesion (disorder)
	263183007	Closed fracture of third cervical vertebra without mention of spinal cord lesion (disorder)
	263184001	Closed fracture of fourth cervical vertebra without mention of spinal cord lesion (disorder)
	263185000	Closed fracture of fifth cervical vertebra without mention of spinal cord lesion (disorder)
	263186004	Closed fracture of sixth cervical vertebra without mention of spinal cord lesion (disorder)
	263187008	Closed fracture of seventh cervical vertebra without mention of spinal cord lesion (disorder)

	263188003	Open fracture of atlas without spinal cord lesion (disorder)
	263189006	Open fracture of axis without spinal cord lesion (disorder)
	26738009	Spinal cord injury without spinal bone injury (disorder)
	268155007	Other specified anomalies of spinal cord (disorder)
	269061001	Fracture of vertebra without spinal cord lesion (disorder)
	269078005	Fracture of spine with spinal cord lesion (disorder)
	269251003	Nerve and spinal cord injuries (disorder)
	269315007	Fracture of vertebra without spinal cord lesion NOS (disorder)
	269400007	Spinal cord (& [no fracture]) &/or nerve root injury (disorder)
	269401006	Spinal cord lesion - no # (disorder)
	269621001	(Secondary Ca brain/spinal cord) or (bone secondary Ca) or (brain secondary Ca) (disorder)
	274154002	Fracture of thoracic spine without spinal cord lesion (disorder)
	276827001	Malignant glioma of spinal cord (disorder)
	276829003	Glioblastoma multiforme of spinal cord (disorder)
	277326006	Spinal cord rupture (disorder)
	277338002	Cystic dermoid choristoma of spinal cord (disorder)
	282784007	Spinal cord syndrome (disorder)
	282807009	Spinal cord compression (disorder)
	28425007	Closed fracture of vertebral column without spinal cord injury (disorder)
	287154000	Spinal cord/nerve root injury (disorder)
	287155004	Spinal cord/nerve root injury NOS (disorder)
	28753006	Closed fracture of first cervical vertebra without spinal cord injury (disorder)
	30943005	Open fracture of sacrum AND/OR coccyx without spinal cord injury (disorder)
	31062001	Open fracture of thoracic vertebra without spinal cord injury (disorder)
	314950006	Migration of spinal cord stimulator (disorder)
	31882001	Closed fracture of T1-T6 level with spinal cord injury (disorder)
	3446000	Open fracture of T7-T12 level with spinal cord injury (disorder)
	34996005	Open fracture of fifth cervical vertebra without spinal cord injury (disorder)
	363475005	Malignant tumor of spinal cord (disorder)
	36838006	Spinal cord injury of thoracic region without bone injury (disorder)
	370987005	Anaplastic astrocytoma of spinal cord (disorder)
	371029002	Ischemic disorder of spinal cord (disorder)
	372310001	Paralysis due to lesion of spinal cord (disorder)
	38170003	Open fracture of vertebral column without spinal cord injury (disorder)
	39881005	Injury at C5-C7 level with complete lesion of spinal cord AND without bone injury (disorder)
	402743009	Spinal cord defect affecting skin (disorder)
	405754008	Cervical spinal cord injury (disorder)
	405755009	Anterior cervical spinal cord injury, without injury to spinal bone, C1-4 (disorder)
	405756005	Anterior cervical spinal cord injury, without spinal bone injury, C5-7 (disorder)
	405759003	Cervical spinal cord injury without spinal bone injury (disorder)
	405762000	Incomplete cervical spinal cord injury, unspecified, without spinal bone injury, C1-4 (disorder)
	405763005	Incomplete cervical spinal cord injury, unspecified,

		without spinal bone injury, C5-7 (disorder)
	405764004	Posterior cervical spinal cord injury without spinal bone injury, C5-7 (disorder)
	405765003	Posterior cervical spinal cord injury, without spinal injury, C1-4 (disorder)
	405766002	Unspecified cervical spinal cord injury, without spinal bone injury, C1-4 (disorder)
	405767006	Unspecified cervical spinal cord injury, without spinal bone injury, C5-7 (disorder)
	40795009	Injury at T1-T6 level with complete lesion of spinal cord AND without bone injury (disorder)
	41078002	Open fracture of lumbar vertebra without spinal cord injury (disorder)
	416903004	Tuberculoma of spinal cord (disorder)
	417484006	Tuberculous abscess of spinal cord (disorder)
	423182004	Lipoma of terminal spinal cord (disorder)
	423361002	Lipoma of dorsal spinal cord (disorder)
	424334007	Malignant tumor of spinal cord, intramedullary (disorder)
	424549003	Malignant tumor of spinal cord, extramedullary (disorder)
	4260009	Sacral spinal cord injury without bone injury (disorder)
	426470009	Lipoma of terminal spinal cord (disorder)
	426510004	Schwannoma of spinal cord (disorder)
	428514001	Closed fracture of sacrum without spinal cord injury (disorder)
	430399004	Infection of spinal cord due to Histoplasma (disorder)
	430404002	Abscess of medulla of spinal cord due to Histoplasma (disorder)
	431520004	Inflammation of spinal cord due to toxin (disorder)
	432249006	Infarction of spinal cord (disorder)
	44434003	Closed fracture of thoracic spine with spinal cord injury (disorder)
	46549005	Multiple spinal cord injuries without spinal bone injury (disorder)
	46995009	Closed fracture of C5-C7 level with spinal cord injury (disorder)
	48522003	Spinal cord disorder (disorder)
	48956000	Open fracture of lumbar vertebra with spinal cord injury (disorder)
	49747005	Closed fracture of lumbar vertebra with spinal cord injury (disorder)
	52622006	Closed fracture of fourth cervical vertebra without spinal cord injury (disorder)
	53810004	Injury at T1-T6 level with spinal cord injury AND without bone injury (disorder)
	54957006	Open fracture of seventh cervical vertebra without spinal cord injury (disorder)
	55983005	Injury at T7-T12 level with complete lesion of spinal cord AND without bone injury (disorder)
	57360002	Spinal cord injury to cervical region without bone injury (disorder)
	5842009	Spinal cord dysplasia (disorder)
	58919004	Open fracture of cervical vertebra without spinal cord injury (disorder)
	60576007	Subacute combined degeneration of spinal cord (disorder)
	62297002	Open fracture of third cervical vertebra without spinal cord injury (disorder)
	63081009	Acute infarction of spinal cord (disorder)
	63903007	Incomplete spinal cord lesion at T7-T12 level without bone injury (disorder)
	64469005	Open fracture of sixth cervical vertebra without spinal cord injury (disorder)
	65491009	Open fracture of sacrum AND/OR coccyx with spinal cord injury (disorder)
	65605001	Edema of spinal cord (disorder)

	66158003	Closed fracture of C5-C7 level with incomplete spinal cord lesion (disorder)
	67975003	Fibrocartilagenous emboli of spinal cord (disorder)
	71286001	Spinal cord compression (disorder)
	72346008	Open fracture of multiple cervical vertebrae without spinal cord injury (disorder)
	72513001	Closed fracture of T7-T12 level with spinal cord injury (disorder)
	76435008	Open fracture of C1-C4 level with spinal cord injury (disorder)
	78085003	Open fracture of T7-T12 level with incomplete spinal cord lesion (disorder)
	78211006	Open fracture of thoracic spine with spinal cord injury (disorder)
	78687007	Closed fracture of cervical vertebra without spinal cord injury (disorder)
	80980003	Open fracture of C5-C7 level with spinal cord injury (disorder)
	81042008	Congenital anomaly of spinal cord (disorder)
	81425005	Closed fracture of sacrum AND/OR coccyx without spinal cord injury (disorder)
	81442004	Closed fracture of cervical region with spinal cord injury (disorder)
	81501004	Open fracture of first cervical vertebra without spinal cord injury (disorder)
	81642009	Late effect of spinal cord injury (disorder)
	8303001	Closed fracture of sixth cervical vertebra without spinal cord injury (disorder)
	83716008	Traumatic spinal cord compression (disorder)
	83991006	Incomplete spinal cord lesion at C5-C7 level without bone injury (disorder)
	87804006	Closed fracture of fifth cervical vertebra without spinal cord injury (disorder)
	8840000	Closed fracture of C1-C4 level with spinal cord injury (disorder)
	8939001	Incomplete spinal cord lesion at C1-C4 level without bone injury (disorder)
	89825003	Closed fracture of thoracic vertebra without spinal cord injury (disorder)
	9048001	Open fracture of fourth cervical vertebra without spinal cord injury (disorder)
	90584004	Spinal cord injury (disorder)
	92405007	Benign neoplasm of spinal cord (disorder)
	9341001	Late effect of fracture of spine AND/OR trunk without spinal cord lesion (disorder)
	94068003	Primary malignant neoplasm of spinal cord (disorder)
	94600009	Secondary malignant neoplasm of spinal cord (disorder)
	95108005	Neoplasm of uncertain behavior of spinal cord (disorder)
	95649006	Demyelinating disease of the spinal cord (disorder)

Appendix 8: Examples of experimental searching for (keyword + 1 right-neighbourhood word)

Single Keyword + Right Word	Related Concepts of "SK +RW" (diagnosis only)	Concept
cancer treated		
XRT completing		
surgery due		
Repeat MRI		
MRI to		
response Prev		
MRI in		
response to		
	191008001	Immunodeficiency following hereditary defective response to Epstein-Barr virus (disorder)
	230434009	Seizures in response to acute event (disorder)
woman with		
month of		
thigh mass		
mass Noticed		
month ago		
lethargy in		
month no		
fevers or		
weight loss		
	71479007	AIDS with abnormal weight loss (disorder)
loss Exam		
Exam approx		
firm mass		
mass within		
medialis adductor		
adductor compartment		
Stage III		
	236415000	Membranous glomerulonephritis - stage III (disorder)
	312703006	Full thickness macular hole stage III (disorder)
III abdominal		
post chemotherapy		
study current		
disease status		
cause posterior		
fossa lesion		
Multiple debulkings		

months chemotherapy		
status at		
	205051008	Hip: [unilateral congenital subluxation] or [unstable] or [prelaxation] or [predislocation status at birth] or [congenital instability of joint] (disorder)
end of		
	111638003	Open fracture of sternal end of clavicle (disorder)
	111644004	Closed fracture of upper end of lower leg (disorder)
	111645003	Open fracture of upper end of tibia (disorder)
	12588000	Closed traumatic dislocation of distal end of ulna (disorder)
	127286005	Fracture of upper end of humerus (disorder)
	13019000	Open posterior dislocation of proximal end of tibia (disorder)
	14741001	Closed posterior dislocation of proximal end of tibia (disorder)
	157207002	Fracture of proximal end of humerus (disorder)
	157209004	Fracture of distal end of humerus (disorder)
	157212001	Fracture of proximal end of radius and ulna (disorder)
	157215004	Fracture of proximal end of radius (disorder)
	157217007	Fracture of distal end of radius (disorder)
	157218002	Fracture of proximal end of ulna (disorder)
	157220004	Fracture of distal end of ulna (disorder)
	157235003	Fracture of distal end of femur (disorder)
	157241005	Fracture of proximal end of tibia and fibula (disorder)
	1658003	Closed fracture of acromial end of clavicle (disorder)
	17222009	Closed fracture of distal end of radius (disorder)
	17359008	Fracture of lower end of humerus (disorder)
	19259001	Closed fracture of upper end of radius AND ulna (disorder)
	20026002	Open dislocation of proximal end of metacarpal bone of wrist (disorder)
	20354001	Open fracture of lower end of humerus (disorder)
	20433007	Fracture of upper end of tibia (disorder)
	20736003	Closed posterior dislocation of distal end of femur (disorder)
	208288001	Fracture of proximal end of humerus (disorder)
	208290000	Fracture of distal end of humerus (disorder)
	208350000	Fracture of proximal end of ulna (disorder)
	208351001	Fracture of proximal end of radius (disorder)
	208355005	Fracture of distal end of radius (disorder)
	208356006	Fracture of distal end of radius and ulna (disorder)
	208594005	Fracture of distal end of femur (disorder)
	208640008	Fracture of proximal end of tibia (disorder)
	208641007	Fracture of distal end of tibia and fibula (disorder)

	208656003	Fracture of distal end of tibia and fibula (disorder)
	21867001	Fracture of upper end of fibula (disorder)
	2202000	Open posterior dislocation of distal end of femur (disorder)
	22682006	Open lateral dislocation of proximal end of tibia (disorder)
	2295008	Closed fracture of upper end of forearm (disorder)
	23813004	Open fracture of proximal end of radius (disorder)
	23900009	Closed fracture of upper end of tibia (disorder)
	249356002	Posterior end of inferior turbinate hypertrophy (disorder)
	25529004	Open fracture of distal end of radius (disorder)
	25586009	Open dislocation of distal end of fibula (disorder)
	263192005	Fracture of distal end of humerus (disorder)
	263195007	Fracture of proximal end of radius (disorder)
	263199001	Fracture of distal end of radius (disorder)
	263203001	Fracture of proximal end of ulna (disorder)
	263205008	Fracture of distal end of ulna (disorder)
	263206009	Fracture of proximal end of radius and ulna (disorder)
	263208005	Fracture of distal end of radius and ulna (disorder)
	263225007	Fracture of proximal end of femur (disorder)
	263232003	Fracture of distal end of femur (disorder)
	263242001	Fracture of distal end of fibula (disorder)
	263243006	Fracture of distal end of tibia and fibula (disorder)
	27152008	Hyperemesis gravidarum before end of 22 week gestation with carbohydrate depletion (disorder)
	278537006	Fracture of distal end of tibia (disorder)
	33041006	Closed fracture of proximal end of ulna (disorder)
	33141009	Open fracture of upper end of forearm (disorder)
	33192001	Closed fracture of lower end of radius AND ulna (disorder)
	33370009	Hyperemesis gravidarum before end of 22 week gestation with electrolyte imbalance (disorder)
	35442005	Closed fracture of lower end of forearm (disorder)
	367415002	Fracture of lower end of humerus (disorder)
	41359009	Closed anterior dislocation of proximal end of tibia (disorder)
	42636007	Closed fracture of upper end of humerus (disorder)
	43765002	Closed multiple fractures of lower end of humerus (disorder)
	44333005	Open medial dislocation of proximal end of tibia (disorder)
	45455009	Multiple open fractures of lower end of femur (disorder)
	46357007	Open multiple fractures of upper end of radius (disorder)
	48561006	Closed fracture of sternal end of clavicle (disorder)

	50397009	Closed fracture of distal end of ulna (disorder)
	51879004	Closed medial dislocation of proximal end of tibia (disorder)
	52784000	Fracture of acromial end of clavicle (disorder)
	56642004	Fracture of sternal end of clavicle (disorder)
	5895007	Closed multiple fractures of upper end of radius (disorder)
	59556008	Open fracture of upper end of fibula (disorder)
	59639009	Closed fracture of upper end of fibula (disorder)
	60180008	Closed fracture of lower end of femur (disorder)
	63456003	Open anterior dislocation of proximal end of tibia (disorder)
	63540005	Open anterior dislocation of distal end of femur (disorder)
	64213003	Closed anterior dislocation of distal end of femur (disorder)
	649002	Open fracture of distal end of ulna (disorder)
	65775005	Open fracture of upper end of radius AND ulna (disorder)
	6584000	Closed fracture of lower end of humerus (disorder)
	69166006	Fracture of upper end of lower leg (disorder)
	69830008	Multiple closed fractures of lower end of femur (disorder)
	70381003	Open dislocation of distal end of ulna (disorder)
	71119005	Fetal blood loss from cut end of co-twin's cord (disorder)
	71139009	Closed fracture of proximal end of radius (disorder)
	72435006	Open fracture of acromial end of clavicle (disorder)
	73387003	Closed traumatic dislocation of proximal end of metacarpal bone of wrist (disorder)
	7341005	Closed multiple fractures of upper end of ulna (disorder)
	75580000	Open dislocation of distal end of tibia (disorder)
	78490005	Open fracture of proximal end of ulna (disorder)
	78849002	Closed traumatic dislocation of distal end of fibula (disorder)
	78980006	Open fracture of lower end of forearm (disorder)
	79763005	Closed traumatic dislocation of distal end of tibia (disorder)
	84595002	Open multiple fractures of upper end of ulna (disorder)
	86844009	Open fracture of upper end of humerus (disorder)
	87621000	Hyperemesis gravidarum before end of 22 week gestation with dehydration (disorder)
	88116004	Open fracture of lower end of radius AND ulna (disorder)
	89635006	Open fracture of lower end of femur (disorder)
	9344009	Open fracture of upper end of lower leg (disorder)
	9787009	Closed lateral dislocation of proximal end of tibia (disorder)
MRI spine		
spine to		

leg weakness		
weakness and		
cord compression		
	192913003	Myelopathy NOS (& [spinal cord compression NOS]) (disorder)
	193080004	(Cord compression) or (spinal cord compression) (disorder)
	206088003	Fetus or neonate affected by other cord compression (disorder)
	206093000	Fetus or neonate affected by cord compression NOS (disorder)
	230140007	Cord compression NOS (disorder)
	230141006	Spinal cord compression NOS (disorder)
	237308002	Umbilical cord compression (disorder)
	282807009	Spinal cord compression (disorder)
	71286001	Spinal cord compression (disorder)
	83716008	Traumatic spinal cord compression (disorder)
concerns re		
spine seen		
team clinical		
problem likely		
HLH Previously		
disease with		
	134422007	Alzheimer's disease with early onset (disorder)
	134423002	Alzheimer's disease with late onset (disorder)
	155452000	(Lymphatic disease) or (vein disease) or (vein &/or lymphatic disease with CVS NOS) (disorder)
	186813004	Acute Chagas' disease with heart involvement (disorder)
	186814005	Chagas' disease with other organ involvement (disorder)
	186815006	Chagas' disease with digestive system involvement (disorder)
	186816007	Chagas' disease with nervous system involvement (disorder)
	186817003	Chagas' disease without mention of organ involvement (disorder)
	192161009	Dementia in Alzheimer's disease with early onset (disorder)
	192162002	Dementia in Alzheimer's disease with late onset (disorder)
	192802002	Dementia in Alzheimer's disease with early onset (disorder)
	192803007	Dementia in Alzheimer's disease with late onset (disorder)
	194762007	Malignant hypertensive heart disease without congestive cardiac failure (disorder)
	194763002	Malignant hypertensive heart disease with congestive cardiac failure (disorder)
	194766005	Benign hypertensive heart disease without congestive cardiac failure (disorder)
	194767001	Benign hypertensive heart disease with congestive cardiac failure (disorder)
	194774006	Hypertensive renal disease with renal failure (disorder)
	194779001	Hypertensive heart and renal disease with (congestive) heart failure (disorder)

	194780003	Hypertensive heart and renal disease with renal failure (disorder)
	194781004	Hypertensive heart and renal disease with both (congestive) heart failure and renal failure (disorder)
	196001008	Chronic obstructive pulmonary disease with acute lower respiratory infection (disorder)
	196002001	Chronic obstructive pulmonary disease with acute exacerbation, unspecified (disorder)
	19604005	Triglyceride storage disease with ichthyosis (disorder)
	196133001	Lung disease with systemic sclerosis (disorder)
	196134007	Lung disease with diseases EC (disorder)
	196136009	Lung disease with polymyositis (disorder)
	196137000	Lung disease with Sjögren's disease (disorder)
	196138005	Lung disease with systemic lupus erythematosus (disorder)
	196139002	Lung disease with syphilis (disorder)
	196140000	Lung disease with diseases EC NOS (disorder)
	196606004	Gastro-oesophageal reflux disease with ulceration (disorder)

Appendix 9: A portion of the results of experiment 3

Related Concept (diagnosis)	Clinical Report	Word Found	Word Found / Total Keywords	G%	Levenshtein Distance
Spondylolisthesis, grade 1 (disorder)	03/08/2006 Right frontal craniotomy and acrylic cranioplasty for grade 1 frontal meningioma.11/10/2007 Right craniotomy for right frontal convexity meningioma and medial wing sphenoid meningioma.24/05/2010 Right craniotomy for recurrent sph	2	2/3	67%	35
Cervical intraepithelial neoplasia grade 1 (disorder)		2	2/5	40%	35
Vaginal intraepithelial neoplasia grade 1 (disorder)		2	2/5	40%	35
Vulval intraepithelial neoplasia grade 1 (disorder)		2	2/5	40%	35
Intraventricular (nontraumatic) hemorrhage, grade 1, of fetus and newborn (disorder)		3	3/9	33%	34
Chronic rejection of renal transplant - grade 1 (disorder)		2	2/7	29%	35
Vaginal intraepithelial neoplasia grade 1 (disorder)		2	2/5	40%	35
Vulval intraepithelial neoplasia grade 1 (disorder)		2	2/5	40%	35
Cervical intraepithelial neoplasia grade 1 (disorder)		2	2/5	40%	35
Sudden cardiac death (disorder)	1 year history of palpitations and syncope preceding aborted sudden cardiac death, with further malignant arrhythmias in hospital. Prolonged QT interval. Is there any evidence of ischaemia in the territory of the regional wall motion abnormalities seen on	2	2/3	67%	36
Sudden cardiac death, so described (disorder)		2	2/5	40%	36
Onset of psoriasis in infancy (<1 year) (disorder)		4	4/7	57%	36
Agoraphobia without history of panic disorder with limited symptom attacks (disorder)		3	3/10	30%	35
Agoraphobia without history of panic disorder (disorder)		2	2/6	33%	36
Agoraphobia without history of panic disorder without limited symptom attacks (disorder)		2	2/10	20%	36
[X] Agoraphobia (& [without history of panic disorder] or [with panic disorder]) (disorder)		3	3/11	27%	36
Myocardial infarction: [old] or [healed] or [personal history of] (disorder)		2	2/9	22%	37
Labor and delivery complicated by biochemical evidence of fetal stress (disorder)		3	3/10	30%	35
Spinal cord injury without evidence of spinal bone injury (disorder)		2	2/9	22%	36
Cervical cord injury without evidence of spinal bone injury (disorder)		2	2/9	22%	36
[X]Personal history of allergy to other anti-infective agents (disorder)		2	2/10	20%	36

Abortion, 2 months (disorder)	14 weeks old baby (ex prem 31 weeks, currently 2 months corrected age) with history of apnoea and bradycardia of newborn, currently on home oxygen. Plan is to trial wean off on next admission exclude central cause (brain stem/cerebral)	2	2/3	67%	38
Agoraphobia without history of panic disorder with limited symptom attacks (disorder)		3	3/10	30%	38
Brain stem laceration with open intracranial wound AND loss of consciousness (disorder)		3	3/11	27%	38
Brain stem laceration with open intracranial wound AND prolonged loss of consciousness (more than 24 hours) AND return to pre-existing conscious level (disorder)		4	4/23	17%	38
Brain stem laceration with open intracranial wound AND brief loss of consciousness (less than one hour) (disorder)		3	3/16	19%	39
Brain stem contusion without open intracranial wound AND with prolonged loss of consciousness (more than 24 hours) AND return to pre-existing conscious level (disorder)		4	4/24	17%	37
Brain stem laceration without open intracranial wound AND with loss of consciousness (disorder)		3	3/12	25%	38
Brain stem laceration with open intracranial wound (disorder)		2	2/7	29%	39
Brain stem contusion with open intracranial wound AND prolonged loss of consciousness (more than 24 hours) AND return to pre-existing conscious level (disorder)		4	4/23	17%	38
Brain stem laceration without open intracranial wound AND with concussion (disorder)		2	2/10	20%	39
Brain stem multiple sclerosis (disorder)		1	1/4	25%	40
Brain stem laceration without open intracranial wound AND with moderate loss of consciousness (1-24 hours) (disorder)		3	3/16	19%	38
Brain stem laceration with open intracranial wound AND no loss of consciousness (disorder)		3	3/12	25%	39
Brain stem laceration without open intracranial wound (disorder)		1	1/7	14%	40
Brain stem contusion with open intracranial wound AND moderate loss of consciousness (1-24 hours) (disorder)		3	3/15	20%	39
Brain stem compression (disorder)		1	1/3	33%	39
Brain stem contusion without open intracranial wound AND with loss of consciousness (disorder)		3	3/12	25%	38
Brain stem contusion without open intracranial wound AND with no loss of consciousness (disorder)		3	3/13	23%	38
Brain stem contusion without open intracranial wound AND with moderate loss of consciousness (1-24 hours) (disorder)		3	3/16	19%	38
Brain stem laceration with open intracranial wound AND moderate loss of consciousness (1-24 hours) (disorder)		3	3/15	20%	39
Brain stem laceration with open intracranial wound AND concussion (disorder)		2	2/9	22%	39
Brain stem contusion with open intracranial wound AND prolonged loss of consciousness (more than 24 hours) without return to pre-existing conscious level (disorder)		4	4/23	17%	38
Brain stem vertigo (disorder)		1	1/3	33%	39
Brain stem contusion without open intracranial wound (disorder)		1	1/7	14%	40
Brain stem contusion with open intracranial wound AND no loss of consciousness (disorder)		3	3/12	25%	39
Brain stem contusion without open intracranial wound AND with brief loss of consciousness (less than one hour) (disorder)		3	3/17	18%	38
Brain stem contusion with open intracranial wound AND brief loss of consciousness (less than one hour) (disorder)		3	3/16	19%	39
Brain stem laceration without open intracranial wound AND with brief loss of consciousness (less than one hour) (disorder)		3	3/17	18%	38
Agoraphobia without history of panic disorder (disorder)		2	2/6	33%	38
Brain stem contusion without open intracranial wound AND with prolonged loss of consciousness (more than 24 hours)		4	4/24	17%	37

without return to pre-existing conscious level (disorder)					
Brain stem herniation (disorder)		1	1/3	33%	39
Brain stem laceration with open intracranial wound AND prolonged loss of consciousness (more than 24 hours) without return to pre-existing conscious level (disorder)		4	4/23	17%	38
Brain stem contusion with open intracranial wound AND loss of consciousness (disorder)		3	3/11	27%	38
Brain stem laceration without open intracranial wound AND with prolonged loss of consciousness (more than 24 hours) without return to pre-existing conscious level (disorder)		4	4/24	17%	37
Brain stem laceration without open intracranial wound AND with prolonged loss of consciousness (more than 24 hours) AND return to pre-existing conscious level (disorder)		4	4/24	17%	37
Brain stem contusion with open intracranial wound AND concussion (disorder)		2	2/9	22%	39
Brain stem laceration without open intracranial wound AND with no loss of consciousness (disorder)		3	3/13	23%	38
Brain stem contusion with open intracranial wound (disorder)		2	2/7	29%	39
Brain stem contusion without open intracranial wound AND with concussion (disorder)		2	2/10	20%	39
Agoraphobia without history of panic disorder without limited symptom attacks (disorder)		2	2/10	20%	38
Benign neoplasm of brain stem (disorder)		3	3/5	60%	37
Primary malignant neoplasm of brain stem (disorder)		3	3/6	50%	37
Secondary malignant neoplasm of brain stem (disorder)		3	3/6	50%	37
Neoplasm of uncertain behavior of brain stem (disorder)		4	4/7	57%	36
Brain stem hemorrhage (disorder)		1	1/3	33%	39
Brain stem ischemia (disorder)		1	1/3	33%	39
Brain stem infarction (disorder)		1	1/3	33%	39
Disorder of brain stem (disorder)		3	3/4	75%	37
Brain stem contusion without open intracranial wound and with unspecified state of consciousness (disorder)		4	4/13	31%	38
Brain stem contusion with open intracranial wound and unspecified state of consciousness (disorder)		4	4/12	33%	38
Brain stem laceration without open intracranial wound and with unspecified state of consciousness (disorder)		4	4/13	31%	38
Brain stem laceration with open intracranial wound and unspecified state of consciousness (disorder)		4	4/12	33%	38
Neoplasm of brain stem (disorder)		3	3/4	75%	37
Brain stem contusion (disorder)		1	1/3	33%	39
Brain stem laceration (disorder)		1	1/3	33%	39
Antepartum haemorrhage with coagulation defect (disorder)		1	1/5	20%	39
[X] Agoraphobia (& [without history of panic disorder] or [with panic disorder]) (disorder)		3	3/11	27%	38
(Brain stem compression) or (herniation of brain stem) (disorder)		4	4/8	50%	37
Myocardial infarction: [old] or [healed] or [personal history of] (disorder)		3	3/9	33%	37
Antepartum hemorrhage with coagulation defect unspecified (disorder)		1	1/6	17%	39
Antepartum hemorrhage with coagulation defect - delivered (disorder)		1	1/6	17%	39
Antepartum hemorrhage with coagulation defect - not delivered (disorder)		1	1/7	14%	39

Antepartum hemorrhage with coagulation defect NOS (disorder)		1	1/6	17%	39
Antepartum hemorrhage with trauma (disorder)		1	1/4	25%	39
Antepartum hemorrhage with trauma unspecified (disorder)		1	1/5	20%	39
Antepartum hemorrhage with trauma - delivered (disorder)		1	1/5	20%	39
Antepartum hemorrhage with trauma - not delivered (disorder)		1	1/6	17%	39
Antepartum hemorrhage with trauma NOS (disorder)		1	1/5	20%	39
Antepartum hemorrhage with uterine leiomyoma (disorder)		1	1/5	20%	39
Antepartum hemorrhage with uterine leiomyoma unspecified (disorder)		1	1/6	17%	39
Antepartum hemorrhage with uterine leiomyoma - delivered (disorder)		1	1/6	17%	39
Antepartum hemorrhage with uterine leiomyoma - not delivered (disorder)		1	1/7	14%	39
Antepartum hemorrhage with uterine leiomyoma NOS (disorder)		1	1/6	17%	39
Fetus with drug damage with antenatal problem (disorder)		2	2/7	29%	39
Fetus with radiation damage with antenatal problem (disorder)		2	2/7	29%	39
Fetal-maternal hemorrhage with antenatal problem (disorder)		1	1/6	17%	39
Prolonged first stage with antenatal problem (disorder)		1	1/6	17%	39
Prolonged second stage with antenatal problem (disorder)		1	1/6	17%	39
Other obstetric pelvic organ damage with postnatal problem (disorder)		1	1/8	12%	39
Third-stage postpartum hemorrhage with postnatal problem (disorder)		1	1/7	14%	39
Other immediate postpartum hemorrhage with postnatal problem (disorder)		1	1/7	14%	39
Secondary postpartum hemorrhage with postnatal problem (disorder)		1	1/6	17%	39
Retained placenta with no hemorrhage with postnatal problem (disorder)		2	2/8	25%	39
Retained products with no hemorrhage with postnatal problem (disorder)		2	2/8	25%	39
Intrapartum hemorrhage with coagulation defect (disorder)		1	1/5	20%	39
Following injury: [other cerebral haemorrhage without mention of open intracranial wound] or [intracranial haematoma] or [cerebral compression] (disorder)		3	3/17	18%	39
Antepartum haemorrhage with uterine leiomyoma (disorder)		1	1/5	20%	39
Antepartum hemorrhage with hypofibrinogenemia (disorder)		1	1/4	25%	39
Antepartum hemorrhage with hyperfibrinolysis (disorder)		1	1/4	25%	39
Antepartum hemorrhage with afibrinogenemia (disorder)		1	1/4	25%	39

Antepartum hemorrhage with coagulation defect (disorder)		1	1/5	20%	39
Perinatal subependymal hemorrhage with intraventricular extension (disorder)		1	1/6	17%	39
Perinatal subependymal hemorrhage with intraventricular and intracerebral extension (disorder)		2	2/8	25%	38
[X]Personal history of allergy to other anti-infective agents (disorder)		3	3/10	30%	37
[X]Other specified disturbances of cerebral status of newborn (disorder)	15 min appt normal shunted hydrocephalus, needs fast MRI in conjunction with his clinic visit 19/5/2011 status of ventricles	3	3/9	33%	19
Predislocation status of hip at birth (disorder)		2	2/6	33%	21
Arthritis of temporomandibular joint as part of polyarthritis (disorder)	18 week old baby with flexor spasms, developmental regression and modified hypsarrhythmia = infantile spasms. Needs MRI head as part of work up please ? underlying brain pathology. There was parental report of some abnormal ventricles found on antenatal	4	4/8	50%	34
Infantile spasms NOS (disorder)		1	1/3	33%	36
Enamel defect as part of syndrome (disorder)		3	3/6	50%	34
Dentine defect as part of syndrome (disorder)		3	3/6	50%	34
Refractory infantile spasms (disorder)		2	2/3	67%	35
Congenital absence of part of brain (disorder)		4	4/6	67%	34
Congenital hypoplasia of part of brain (disorder)		4	4/6	67%	34
Fracture of interligamentous part of clavicle (disorder)		3	3/6	50%	35
Arthritis of temporomandibular joint as part of polyarthritis (disorder)		4	4/8	50%	34
Malignant neoplasm of fixed part of tongue NOS (disorder)		3	3/8	38%	35
Malignant tumor of cervical part of esophagus (disorder)		3	3/7	43%	35
Malignant tumor of thoracic part of esophagus (disorder)		3	3/7	43%	35
Malignant tumor of abdominal part of esophagus (disorder)		3	3/7	43%	35
Malignant neoplasm of other specified part of esophagus (disorder)		3	3/8	38%	35
Malignant neoplasm of skin of other and unspecified part of face NOS (disorder)		5	5/12	42%	34
Malignant neoplasm of central part of female breast (disorder)		3	3/8	38%	35
Malignant neoplasm of other part of brain NOS (disorder)		4	4/8	50%	34
[X]Malignant neoplasm, overlapping lesion of brain and other part of central nervous system (disorder)		5	5/14	36%	34
Malignant neoplasm of other specified part of nervous system (disorder)		3	3/9	33%	35
Secondary malignant neoplasm of other part of nervous system (disorder)		3	3/9	33%	35

Benign neoplasm of other part of nervous system (disorder)		3	3/8	38%	35
[X]Malignant neoplasm, overlapping lesion of brain and other part of central nervous system (disorder)		5	5/14	36%	34
(Cerebellar hypoplasia) or (reduction deformities of brain NOS) or ([agenesis] or [aplasia] or [hypoplasia]) of part of brain NEC) (disorder)		6	6/19	32%	34
Congenital undergrowth of proximal part of limb (& lower limb) (disorder)		3	3/9	33%	35
Congenital undergrowth of distal part of limb (& lower limb) (disorder)		3	3/9	33%	35
Other deletions of part of a chromosome (disorder)		3	3/7	43%	35
Other deletion of part of a chromosome NOS (disorder)		3	3/8	38%	35

Appendix 10: A portion (the first 800 records) of the results of experiment 4

Support	Found diagnosis concepts
20	[X] Agoraphobia (& [without history of panic disorder] or [with panic disorder]) (disorder)
20	[X]Adenovirus as the cause of diseases classified to other chapters (disorder)
20	[X]Bacillus fragilis as the cause of diseases classified to other chapters (disorder)
20	[X]Clostridium perfringens as the cause of diseases classified to other chapters (disorder)
20	[X]Escherichia coli [e coli] as the cause of diseases classified to other chapters (disorder)
20	[X]Haemophilus influenzae [h influenzae] as the cause of diseases classified to other chapters (disorder)
20	[X]Klebsiella pneumoniae as the cause of diseases classified to other chapters (disorder)
20	[X]Mycoplasma pneumoniae [pplo] as the cause of diseases classified to other chapters (disorder)
20	[X]Other staphylococcus as the cause of diseases classified to other chapters (disorder)
20	[X]Other streptococcus as the cause of diseases classified to other chapters (disorder)
20	[X]Personal history of allergy to other anti-infective agents (disorder)
20	[X]Proteus (mirabilis)(morganii) as the cause of diseases classified to other chapters (disorder)
20	[X]Pseudomonas as the cause of diseases classified to other chapters (disorder)
20	[X]Staphylococcus aureus as the cause of diseases classified to other chapters (disorder)
20	[X]Streptococcus pneumoniae as the cause of diseases classified to other chapters (disorder)
20	[X]Streptococcus, group a, as the cause of diseases classified to other chapters (disorder)
20	[X]Streptococcus, group b, as the cause of diseases classified to other chapters (disorder)
20	[X]Streptococcus, group d, as the cause of diseases classified to other chapters (disorder)
20	[X]Unspecified staphylococcus as the cause of diseases classified to other chapters (disorder)
20	[X]Unspecified streptococcus as the cause of diseases classified to other chapters (disorder)
20	Agoraphobia without history of panic disorder (disorder)
20	Agoraphobia without history of panic disorder with limited symptom attacks (disorder)
20	Agoraphobia without history of panic disorder without limited symptom attacks (disorder)
20	Coronavirus as the cause of diseases classified to other chapters (disorder)
20	Helicobacter pylori as the cause of diseases classified to other chapters (disorder)
20	Methicillin resistant staphylococcus aureus as the cause of diseases classified to other chapters (disorder)
20	Myocardial infarction: [old] or [healed] or [personal history of] (disorder)
20	Papillomavirus as the cause of diseases classified to other chapters (disorder)
20	Parvovirus as the cause of diseases classified to other chapters (disorder)
20	Reovirus as the cause of diseases classified to other chapters (disorder)
20	Respiratory syncytial virus as the cause of diseases classified to other chapters (disorder)
20	Retrovirus as the cause of diseases classified to other chapters (disorder)
19	Fracture of distal end of femur (disorder)
19	Fracture of distal end of humerus (disorder)
19	Fracture of distal end of radius (disorder)
19	Fracture of distal end of tibia and fibula (disorder)
19	Fracture of proximal end of radius (disorder)
19	Fracture of proximal end of ulna (disorder)
18	Open fracture of base of skull with other and unspecified intracranial hemorrhage (disorder)
18	Spinal cord injury without evidence of spinal bone injury (disorder)
17	Cervical cord injury without evidence of spinal bone injury (disorder)
17	Labor and delivery complicated by biochemical evidence of fetal stress (disorder)

16	(Backache NOS) or (back pain [& low]) (disorder)
16	(Benign neoplasm: [digestive system] or [liver] or [rectum]) or (polyposis coli) or (papilloma rectum NOS) or (islet cell tumour) (disorder)
16	[X]Fracture of lower leg, part unspecified (disorder)
16	[X]Injury of unspecified blood vessel at lower leg level (disorder)
16	[X]Injury of unspecified muscle and tendon at lower leg level (disorder)
16	[X]Traumatic amputation of lower leg, level unspecified (disorder)
16	Bacterial causes of diseases classified elsewhere (disorder)
16	Bruise - knee &/or lower leg (disorder)
16	Closed fracture of base of skull with intracranial injury, with 1-24 hours loss of consciousness (disorder)
16	Closed fracture of base of skull with intracranial injury, with less than 1 hour loss of consciousness (disorder)
16	Closed fracture of base of skull with intracranial injury, with loss of consciousness of unspecified duration (disorder)
16	Closed fracture of base of skull with intracranial injury, with more than 24 hours loss of consciousness and return to pre-existing conscious level (disorder)
16	Closed fracture of base of skull with intracranial injury, with more than 24 hours loss of consciousness without return to pre-existing conscious level (disorder)
16	Closed fracture of base of skull with intracranial injury, with no loss of consciousness (disorder)
16	Closed fracture of base of skull without mention of intracranial injury, with 1-24 hours loss of consciousness (disorder)
16	Closed fracture of base of skull without mention of intracranial injury, with less than 1 hour loss of consciousness (disorder)
16	Closed fracture of base of skull without mention of intracranial injury, with loss of consciousness of unspecified duration (disorder)
16	Closed fracture of base of skull without mention of intracranial injury, with more than 24 hours loss of consciousness and return to pre-existing conscious level (disorder)
16	Closed fracture of base of skull without mention of intracranial injury, with more than 24 hours loss of consciousness without return to pre-existing conscious level (disorder)
16	Closed fracture of base of skull without mention of intracranial injury, with no loss of consciousness (disorder)
16	Closed fracture of lower leg (disorder)
16	Closed fracture of upper end of lower leg (disorder)
16	Congenital heart block (disorder)
16	Fracture malunion - lower leg (disorder)
16	Fracture of distal end of radius and ulna (disorder)
16	Fracture of distal end of ulna (disorder)
16	Fracture of lower end of humerus (disorder)
16	Fracture of lower leg (disorder)
16	Fracture of proximal end of humerus (disorder)
16	Fracture of proximal end of radius and ulna (disorder)
16	Fracture of thoracic spine without spinal cord lesion (disorder)
16	Fracture of upper end of lower leg (disorder)
16	Hodgkin's disease, lymphocytic depletion of intrapelvic lymph nodes (disorder)
16	Hodgkin's disease, lymphocytic-histiocytic predominance of intrapelvic lymph nodes (disorder)
16	Hodgkin's disease, mixed cellularity of intrapelvic lymph nodes (disorder)
16	Hodgkin's disease, nodular sclerosis of intrapelvic lymph nodes (disorder)
16	Hodgkin's granuloma of intrapelvic lymph nodes (disorder)
16	Hodgkin's paraganuloma of intrapelvic lymph nodes (disorder)
16	Hodgkin's sarcoma of intrapelvic lymph nodes (disorder)
16	Injury of lower leg (disorder)
16	Injury of spinal nerve root (disorder)
16	Leukemic reticuloendotheliosis of intrapelvic lymph nodes (disorder)
16	Lymphosarcoma of intrapelvic lymph nodes (disorder)
16	Mixed germ cell tumour (disorder)

16	Motor neurone disease (disorder)
16	Mycosis fungoides of intrapelvic lymph nodes (disorder)
16	Neuralgia/neuritis - lower leg (disorder)
16	Nodular lymphoma of intrapelvic lymph nodes (disorder)
16	Nodular lymphoma of lymph nodes of inguinal region and lower limb (disorder)
16	Open fracture of base of neck of femur (disorder)
16	Open fracture of base of skull with intracranial injury, with 1-24 hours loss of consciousness (disorder)
16	Open fracture of base of skull with intracranial injury, with less than 1 hour loss of consciousness (disorder)
16	Open fracture of base of skull with intracranial injury, with loss of consciousness of unspecified duration (disorder)
16	Open fracture of base of skull with intracranial injury, with more than 24 hours loss of consciousness and return to pre-existing conscious level (disorder)
16	Open fracture of base of skull with intracranial injury, with more than 24 hours loss of consciousness without return to pre-existing conscious level (disorder)
16	Open fracture of base of skull with intracranial injury, with no loss of consciousness (disorder)
16	Open fracture of base of skull without mention of intracranial injury, with 1-24 hours loss of consciousness (disorder)
16	Open fracture of base of skull without mention of intracranial injury, with less than 1 hour loss of consciousness (disorder)
16	Open fracture of base of skull without mention of intracranial injury, with loss of consciousness of unspecified duration (disorder)
16	Open fracture of base of skull without mention of intracranial injury, with more than 24 hours loss of consciousness and return to pre-existing conscious level (disorder)
16	Open fracture of base of skull without mention of intracranial injury, with more than 24 hours loss of consciousness without return to pre-existing conscious level (disorder)
16	Open fracture of base of skull without mention of intracranial injury, with no loss of consciousness (disorder)
16	Open fracture of lower leg (disorder)
16	Open fracture of upper end of lower leg (disorder)
16	Pathological fracture - lower leg (disorder)
16	Reticulosarcoma of intrapelvic lymph nodes (disorder)
16	Reticulosarcoma of lymph nodes of inguinal region and lower limb (disorder)
16	S��zary's disease of intrapelvic lymph nodes (disorder)
16	Spinal cord (& [no fracture]) &/or nerve root injury (disorder)
16	Spinal cord/nerve root injury NOS (disorder)
16	Types of diseases (disorder)
15	(Back pain: [lumbar spine] or [low] or [acute lumbar]) or (lumbalgia) or (lumbago) (disorder)
15	(Benign islet cell tumour) or (endocrine tumour of pancreas) (disorder)
15	[X]Enterovirus as the cause of diseases classified to other chapters (disorder)
15	[X]Other specified bacterial agents as the cause of diseases classified to other chapters (disorder)
15	[X]Other viral agents as the cause of diseases classified to other chapters (disorder)
15	Deep full thickness burn of the lower leg without loss of body part (disorder)
15	Deep full thickness burn of the lower leg, with loss of body part (disorder)
15	Deep third degree burn of lower leg with loss of body part (disorder)
15	Germ cell tumour (disorder)
14	(Benign tumour: [brain/nervous] or [brain] or [cerebral]) or (acoustic neuroma) or (pituitary adenoma) or (meningioma) (disorder)
14	(Frigidity or impotence) or (lack or loss of libido) or (psychogen.dyspareunia) or (fear of ejaculation) or (premature ejaculation) (disorder)
14	(Neuralgia/neuritis - NOS) or (nerve root pain NEC) (disorder)
14	[X] Cause of diseases classified to other chapters: [OS bacterial agents] or [coliform bacteria] (disorder)
14	[X]Dislocation of other and unspecified parts of lumbar spine and pelvis (disorder)
14	[X]Other and unspecified peripheral and cutaneous T-cell lymphomas (disorder)
14	[X]Sequelae of burn and corrosion classifiable only according to extent of body surface involved (disorder)
14	[X]Sprain and strain of other and unspecified parts of lumbar spine and pelvis (disorder)

14	Acute back pain &/or sciatica (disorder)
14	Angiocentric T-cell lymphoma (disorder)
14	Closed fracture base of skull with intracranial injury (disorder)
14	Closed fracture base of skull without intracranial injury (disorder)
14	Closed fracture of base of skull (disorder)
14	Closed fracture of base of skull with cerebral laceration AND/OR contusion (disorder)
14	Closed fracture of base of skull with intracranial hemorrhage (disorder)
14	Closed fracture of base of skull with intracranial injury (disorder)
14	Closed fracture of base of skull with intracranial injury of other and unspecified nature (disorder)
14	Closed fracture of base of skull with intracranial injury, unspecified state of consciousness (disorder)
14	Closed fracture of base of skull with intracranial injury, with concussion, unspecified (disorder)
14	Closed fracture of base of skull with other and unspecified intracranial hemorrhage (disorder)
14	Closed fracture of base of skull with subarachnoid, subdural AND/OR extradural hemorrhage (disorder)
14	Closed fracture of base of skull without intracranial injury (disorder)
14	Closed fracture of base of skull without mention of intracranial injury, unspecified state of consciousness (disorder)
14	Closed fracture of base of skull without mention of intracranial injury, with concussion, unspecified (disorder)
14	Closed fracture of skull NOS with intracranial injury, with 1-24 hours loss of consciousness (disorder)
14	Closed fracture of skull NOS with intracranial injury, with less than 1 hour loss of consciousness (disorder)
14	Closed fracture of skull NOS with intracranial injury, with loss of consciousness of unspecified duration (disorder)
14	Closed fracture of skull NOS with intracranial injury, with more than 24 hours loss of consciousness and return to pre-existing conscious level (disorder)
14	Closed fracture of skull NOS with intracranial injury, with more than 24 hours loss of consciousness without return to pre-existing conscious level (disorder)
14	Closed fracture of skull NOS with intracranial injury, with no loss of consciousness (disorder)
14	Closed fracture of skull NOS without mention of intracranial injury, with 1-24 hours loss of consciousness (disorder)
14	Closed fracture of skull NOS without mention of intracranial injury, with less than 1 hour loss of consciousness (disorder)
14	Closed fracture of skull NOS without mention of intracranial injury, with loss of consciousness of unspecified duration (disorder)
14	Closed fracture of skull NOS without mention of intracranial injury, with more than 24 hours loss of consciousness with return to pre-existing conscious level (disorder)
14	Closed fracture of skull NOS without mention of intracranial injury, with more than 24 hours loss of consciousness without return to pre-existing conscious level (disorder)
14	Closed fracture of skull NOS without mention of intracranial injury, with no loss of consciousness (disorder)
14	Closed fracture of vault of skull with intracranial injury, with 1-24 hours loss of consciousness (disorder)
14	Closed fracture of vault of skull with intracranial injury, with less than 1 hour loss of consciousness (disorder)
14	Closed fracture of vault of skull with intracranial injury, with loss of consciousness of unspecified duration (disorder)
14	Closed fracture of vault of skull with intracranial injury, with more than 24 hours loss of consciousness and return to pre-existing conscious level (disorder)
14	Closed fracture of vault of skull with intracranial injury, with more than 24 hours loss of consciousness without return to pre-existing conscious level (disorder)
14	Closed fracture of vault of skull with intracranial injury, with no loss of consciousness (disorder)
14	Closed fracture of vault of skull without mention of intracranial injury, with 1-24 hours loss of consciousness (disorder)
14	Closed fracture of vault of skull without mention of intracranial injury, with less than 1 hour loss of consciousness (disorder)
14	Closed fracture of vault of skull without mention of intracranial injury, with loss of consciousness of unspecified duration (disorder)
14	Closed fracture of vault of skull without mention of intracranial injury, with more than 24 hours loss of consciousness and return to pre-existing conscious level (disorder)
14	Closed fracture of vault of skull without mention of intracranial injury, with more than 24 hours loss of consciousness without return to pre-existing conscious level (disorder)
14	Closed fracture of vault of skull without mention of intracranial injury, with no loss of consciousness (disorder)
14	Closed multiple fractures of thoracic spine (disorder)
14	Conduction disorders (& [heart block]) (disorder)

14	Crush injury of head and neck (disorder)
14	Deep full thickness burn of a single finger with loss of body part (disorder)
14	Deep full thickness burn of a single finger without loss of body part (disorder)
14	Deep third degree burn of face, head AND/OR neck with loss of body part (disorder)
14	Diffuse large B-cell lymphoma (nodal/systemic with skin involvement) (disorder)
14	Fracture of base of skull (disorder)
14	Fracture of base of skull NOS (disorder)
14	Fracture of lumbar spine - no cord lesion (disorder)
14	Fracture of lumbar spine (disorder)
14	Fracture of lumbar spine with cord lesion (disorder)
14	Fracture of thoracic spine (disorder)
14	Fracture of thoracic spine with cord lesion (disorder)
14	Fracture: [base of skull] or [fossa: [anterior] or [middle] or [posterior]] or [sinus: [ethmoid] or [frontal]] or [bone: [occiput] or [sphenoid] or [temporal] or [orbital roof]] (disorder)
14	Heart block (disorder)
14	Heart block NOS (disorder)
14	Hodgkin's disease, lymphocytic depletion of intra-abdominal lymph nodes (disorder)
14	Hodgkin's disease, lymphocytic depletion of intrathoracic lymph nodes (disorder)
14	Hodgkin's disease, lymphocytic depletion of lymph nodes of multiple sites (disorder)
14	Hodgkin's disease, lymphocytic-histiocytic predominance of intra-abdominal lymph nodes (disorder)
14	Hodgkin's disease, lymphocytic-histiocytic predominance of intrathoracic lymph nodes (disorder)
14	Hodgkin's disease, lymphocytic-histiocytic predominance of lymph nodes of multiple sites (disorder)
14	Hodgkin's disease, mixed cellularity of intra-abdominal lymph nodes (disorder)
14	Hodgkin's disease, mixed cellularity of intrathoracic lymph nodes (disorder)
14	Hodgkin's disease, mixed cellularity of lymph nodes of multiple sites (disorder)
14	Hodgkin's disease, nodular sclerosis of intra-abdominal lymph nodes (disorder)
14	Hodgkin's disease, nodular sclerosis of intrathoracic lymph nodes (disorder)
14	Hodgkin's disease, nodular sclerosis of lymph nodes of multiple sites (disorder)
14	Hodgkin's granuloma of intra-abdominal lymph nodes (disorder)
14	Hodgkin's granuloma of intrathoracic lymph nodes (disorder)
14	Hodgkin's granuloma of lymph nodes of multiple sites (disorder)
14	Hodgkin's paraganuloma of intra-abdominal lymph nodes (disorder)
14	Hodgkin's paraganuloma of intrathoracic lymph nodes (disorder)
14	Hodgkin's paraganuloma of lymph nodes of multiple sites (disorder)
14	Hodgkin's sarcoma of intra-abdominal lymph nodes (disorder)
14	Hodgkin's sarcoma of intrathoracic lymph nodes (disorder)
14	Hodgkin's sarcoma of lymph nodes of multiple sites (disorder)
14	Hormone refractory prostate cancer (disorder)
14	Injury of dorsal nerve roots (disorder)
14	Injury to blood vessels of head and neck (disorder)
14	Kaposi's sarcoma of lymph nodes (disorder)
14	Leukemic reticuloendotheliosis of intra-abdominal lymph nodes (disorder)
14	Leukemic reticuloendotheliosis of intrathoracic lymph nodes (disorder)
14	Leukemic reticuloendotheliosis of lymph nodes of multiple sites (disorder)
14	Loss of teeth due to extraction (disorder)
14	Loss of teeth due to local periodontal disease (disorder)
14	Lymphosarcoma of intra-abdominal lymph nodes (disorder)

14	Lymphosarcoma of intrathoracic lymph nodes (disorder)
14	Lymphosarcoma of lymph nodes of multiple sites (disorder)
14	Malignant melanoma of head and neck (disorder)
14	Malignant tumour of eye: [Ca eye] or [malignant melanoma] or [retinoblastoma] (disorder)
14	Malignant tumour of lower labial mucosa (disorder)
14	Malignant tumour of testis: [carcinoma] or [seminoma] or [teratoma] (disorder)
14	Malignant tumour of tonsillar pillar (disorder)
14	Malignant tumour of upper labial mucosa (disorder)
14	Metastasis to head and neck lymph node (disorder)
14	Metastasis to multiple lymph nodes (disorder)
14	Mycosis fungoides of intra-abdominal lymph nodes (disorder)
14	Mycosis fungoides of intrathoracic lymph nodes (disorder)
14	Mycosis fungoides of lymph nodes of multiple sites (disorder)
14	Nodular lymphoma of intra-abdominal lymph nodes (disorder)
14	Nodular lymphoma of intrathoracic lymph nodes (disorder)
14	Nodular lymphoma of lymph nodes of axilla and upper limb (disorder)
14	Nodular lymphoma of lymph nodes of multiple sites (disorder)
14	On examination - clinically significant macular edema of left eye (disorder)
14	On examination - lower leg bone abnormal (disorder)
14	Open fracture base of skull with intracranial injury (disorder)
14	Open fracture base of skull without mention of intracranial injury (disorder)
14	Open fracture of base of skull with cerebral laceration AND contusion (disorder)
14	Open fracture of base of skull with intracranial hemorrhage (disorder)
14	Open fracture of base of skull with intracranial injury (disorder)
14	Open fracture of base of skull with intracranial injury, unspecified state of consciousness (disorder)
14	Open fracture of base of skull with intracranial injury, with concussion, unspecified (disorder)
14	Open fracture of base of skull with subarachnoid, subdural AND/OR extradural hemorrhage (disorder)
14	Open fracture of base of skull without intracranial injury (disorder)
14	Open fracture of base of skull without mention of intracranial injury, unspecified state of consciousness (disorder)
14	Open fracture of base of skull without mention of intracranial injury, with concussion, unspecified (disorder)
14	Open fracture of skull NOS with intracranial injury, with 1-24 hours loss of consciousness (disorder)
14	Open fracture of skull NOS with intracranial injury, with less than 1 hour loss of consciousness (disorder)
14	Open fracture of skull NOS with intracranial injury, with loss of consciousness of unspecified duration (disorder)
14	Open fracture of skull NOS with intracranial injury, with more than 24 hours loss of consciousness and return to pre-existing conscious level (disorder)
14	Open fracture of skull NOS with intracranial injury, with more than 24 hours loss of consciousness without return to pre-existing conscious level (disorder)
14	Open fracture of skull NOS with intracranial injury, with no loss of consciousness (disorder)
14	Open fracture of skull NOS without mention of intracranial injury, with 1-24 hours loss of consciousness (disorder)
14	Open fracture of skull NOS without mention of intracranial injury, with less than 1 hour loss of consciousness (disorder)
14	Open fracture of skull NOS without mention of intracranial injury, with loss of consciousness of unspecified duration (disorder)
14	Open fracture of skull NOS without mention of intracranial injury, with more than 24 hours loss of consciousness and return to pre-existing conscious level (disorder)
14	Open fracture of skull NOS without mention of intracranial injury, with more than 24 hours loss of consciousness without return to pre-existing conscious level (disorder)
14	Open fracture of skull NOS without mention of intracranial injury, with no loss of consciousness (disorder)
14	Open fracture of vault of skull with intracranial injury, with 1-24 hours loss of consciousness (disorder)
14	Open fracture of vault of skull with intracranial injury, with less than 1 hour loss of consciousness (disorder)

14	Open fracture of vault of skull with intracranial injury, with loss of consciousness of unspecified duration (disorder)
14	Open fracture of vault of skull with intracranial injury, with more than 24 hours loss of consciousness and return to pre-existing conscious level (disorder)
14	Open fracture of vault of skull with intracranial injury, with more than 24 hours loss of consciousness without return to pre-existing conscious level (disorder)
14	Open fracture of vault of skull with intracranial injury, with no loss of consciousness (disorder)
14	Open fracture of vault of skull without mention of intracranial injury, with 1-24 hours loss of consciousness (disorder)
14	Open fracture of vault of skull without mention of intracranial injury, with less than 1 hour loss of consciousness (disorder)
14	Open fracture of vault of skull without mention of intracranial injury, with loss of consciousness of unspecified duration (disorder)
14	Open fracture of vault of skull without mention of intracranial injury, with more than 24 hours loss of consciousness and return to pre-existing level of consciousness (disorder)
14	Open fracture of vault of skull without mention of intracranial injury, with more than 24 hours loss of consciousness without return to pre-existing conscious level (disorder)
14	Open fracture of vault of skull without mention of intracranial injury, with no loss of consciousness (disorder)
14	Open multiple fracture of thoracic spine (disorder)
14	Pathological fracture due to metastatic bone disease (disorder)
14	Peripheral T-cell lymphoma (disorder)
14	Pulmonary transfusion reaction of sudden onset (disorder)
14	Renal artery stenosis of unknown cause (disorder)
14	Reticulosarcoma of intra-abdominal lymph nodes (disorder)
14	Reticulosarcoma of intrathoracic lymph nodes (disorder)
14	Reticulosarcoma of lymph nodes of axilla and upper limb (disorder)
14	Reticulosarcoma of lymph nodes of multiple sites (disorder)
14	SÃ©zary's disease of intra-abdominal lymph nodes (disorder)
14	SÃ©zary's disease of intrathoracic lymph nodes (disorder)
14	SÃ©zary's disease of lymph nodes of multiple sites (disorder)
14	Sequelae of burn and corrosion classifiable only according to extent of body surface involved (disorder)
14	Sequelae of superficial injury and open wound of neck and trunk (disorder)
14	Spinal cord compression (disorder)
14	Tuberculosis of peripheral lymph nodes (disorder)
14	Type II diabetes mellitus with arthropathy (disorder)
14	Type II diabetes mellitus with diabetic cataract (disorder)
14	Type II diabetes mellitus with mononeuropathy (disorder)
14	Type II diabetes mellitus with nephropathy (disorder)
14	Type II diabetes mellitus with neuropathic arthropathy (disorder)
14	Type II diabetes mellitus with peripheral angiopathy (disorder)
14	Type II diabetes mellitus with polyneuropathy (disorder)
14	Type II diabetes mellitus without complication (disorder)
14	Warthin tumour of parotid gland (disorder)
13	(Arthropathy NOS of the lower leg) or (knee arthritis NOS) (disorder)
13	(Backache unspecified) or (back pain unspecified & [acute]) (disorder)
13	(Closed fracture of tibia and fibula, unspecified part, NOS) or (lower leg fracture NOS) (disorder)
13	(Defect of skull ossification) or (other specified skull or face bone anomaly) or (cranial dysostosis NEC) (disorder)
13	(Intervertebral disc: [disorders] or [displacement] or [slipped]) or (acute back pain) (disorder)
13	(Swelling: [limb] or [calf]) or (swollen: [lower leg] or [legs]) (disorder)
13	[X]Contusion of other and unspecified parts of lower leg (disorder)
13	[X]Crushing injury of other and unspecified parts of lower leg (disorder)
13	[X]Fracture of skull and facial bones, part unspecified (disorder)

13	[X]Fractures of other parts of lower leg (disorder)
13	[X]Injuries to the knee and lower leg (disorder)
13	[X]Injury of other blood vessels at lower leg level (disorder)
13	[X]Injury of other muscle(s) and tendon(s) of posterior muscle group at lower leg level (disorder)
13	[X]Injury of other muscles and tendons at lower leg level (disorder)
13	[X]Injury of other nerves at lower leg level (disorder)
13	[X]Injury of unspecified nerve at lower leg level (disorder)
13	[X]Open wound of other parts of lower leg (disorder)
13	[X]Other specified congenital malformations of skull and face bones (disorder)
13	[X]Other specified disturbances of cerebral status of newborn (disorder)
13	[X]Other specified injuries of lower leg (disorder)
13	[X]Other superficial injuries of lower leg (disorder)
13	[X]Sequelae of other specified injuries of neck and trunk (disorder)
13	[X]Sequelae of unspecified injury of neck and trunk (disorder)
13	Abrasion of lower leg, infected (disorder)
13	Abrasion or friction burn of face, neck and scalp, infected, NOS (disorder)
13	Abrasion, lower leg (disorder)
13	Acute lymphangitis of lower leg (disorder)
13	Acute osteomyelitis of lower leg (disorder)
13	Acute osteomyelitis of the lower leg (disorder)
13	Adult onset Still's disease (disorder)
13	Allergic arthritis of the lower leg (disorder)
13	Allergic contact dermatitis of lower leg (disorder)
13	Angina at rest (disorder)
13	Animal bite of lower leg (disorder)
13	Ankylosis of joint of lower leg (disorder)
13	Arthralgia of the lower leg (& [knee]) (disorder)
13	Arthropathy associated with mycoses, of the lower leg (disorder)
13	Arthropathy associated with other bacterial disease, of the lower leg (disorder)
13	Arthropathy associated with other viral disease, of the lower leg (disorder)
13	Arthropathy in Behcet's syndrome of the lower leg (disorder)
13	Arthropathy NOS, of the lower leg (disorder)
13	Arthropathy of the lower leg associated with helminthiasis (disorder)
13	Articular cartilage disorder of lower leg (disorder)
13	Back pain (disorder)
13	Back pain complicating pregnancy (disorder)
13	Benign neoplasm of blood vessel of lower leg (disorder)
13	Benign neoplasm of connective and soft tissue of lower leg NEC (disorder)
13	Benign neoplasm of intrapelvic lymph nodes (disorder)
13	Benign neoplasm of lymph nodes of lower limb (disorder)
13	Benign neoplasm of lymph nodes of neck (disorder)
13	Benign neoplasm of skin of lower leg (disorder)
13	Benign neoplasm of skull and facial bones (disorder)
13	Benign neoplasm of skull or facial bone NOS (disorder)
13	Benign neoplasm of superficial inguinal lymph nodes (disorder)
13	Benign neoplasm of supraclavicular lymph nodes (disorder)

13	Bladder neck obstruction (disorder)
13	Blister of lower leg (& traumatic) (disorder)
13	Blister of lower leg, infected (& traumatic) (disorder)
13	Boil of lower leg (disorder)
13	Bone infection NOS, of the lower leg (disorder)
13	Brain injury with open intracranial wound AND brief loss of consciousness (less than one hour) (disorder)
13	Brain injury with open intracranial wound AND loss of consciousness (disorder)
13	Brain injury with open intracranial wound AND moderate loss of consciousness (1-24 hours) (disorder)
13	Brain injury with open intracranial wound AND no loss of consciousness (disorder)
13	Brain injury with open intracranial wound AND prolonged loss of consciousness (more than 24 hours) AND return to pre-existing conscious level (disorder)
13	Brain injury with open intracranial wound AND prolonged loss of consciousness (more than 24 hours) without return to pre-existing conscious level (disorder)
13	Brain injury without open intracranial wound AND with brief loss of consciousness (less than one hour) (disorder)
13	Brain injury without open intracranial wound AND with loss of consciousness (disorder)
13	Brain injury without open intracranial wound AND with moderate loss of consciousness (1-24 hours) (disorder)
13	Brain injury without open intracranial wound AND with no loss of consciousness (disorder)
13	Brain injury without open intracranial wound AND with prolonged loss of consciousness (more than 24 hours) with return to pre-existing conscious level (disorder)
13	Brain injury without open intracranial wound AND with prolonged loss of consciousness (more than 24 hours) without return to pre-existing conscious level (disorder)
13	Brain stem contusion with open intracranial wound AND brief loss of consciousness (less than one hour) (disorder)
13	Brain stem contusion with open intracranial wound AND loss of consciousness (disorder)
13	Brain stem contusion with open intracranial wound AND moderate loss of consciousness (1-24 hours) (disorder)
13	Brain stem contusion with open intracranial wound AND no loss of consciousness (disorder)
13	Brain stem contusion with open intracranial wound AND prolonged loss of consciousness (more than 24 hours) AND return to pre-existing conscious level (disorder)
13	Brain stem contusion with open intracranial wound AND prolonged loss of consciousness (more than 24 hours) without return to pre-existing conscious level (disorder)
13	Brain stem contusion without open intracranial wound AND with brief loss of consciousness (less than one hour) (disorder)
13	Brain stem contusion without open intracranial wound AND with loss of consciousness (disorder)
13	Brain stem contusion without open intracranial wound AND with moderate loss of consciousness (1-24 hours) (disorder)
13	Brain stem contusion without open intracranial wound AND with no loss of consciousness (disorder)
13	Brain stem contusion without open intracranial wound AND with prolonged loss of consciousness (more than 24 hours) AND return to pre-existing conscious level (disorder)
13	Brain stem contusion without open intracranial wound AND with prolonged loss of consciousness (more than 24 hours) without return to pre-existing conscious level (disorder)
13	Brain stem laceration with open intracranial wound AND brief loss of consciousness (less than one hour) (disorder)
13	Brain stem laceration with open intracranial wound AND loss of consciousness (disorder)
13	Brain stem laceration with open intracranial wound AND moderate loss of consciousness (1-24 hours) (disorder)
13	Brain stem laceration with open intracranial wound AND no loss of consciousness (disorder)
13	Brain stem laceration with open intracranial wound AND prolonged loss of consciousness (more than 24 hours) AND return to pre-existing conscious level (disorder)
13	Brain stem laceration with open intracranial wound AND prolonged loss of consciousness (more than 24 hours) without return to pre-existing conscious level (disorder)
13	Brain stem laceration without open intracranial wound AND with brief loss of consciousness (less than one hour) (disorder)
13	Brain stem laceration without open intracranial wound AND with loss of consciousness (disorder)
13	Brain stem laceration without open intracranial wound AND with moderate loss of consciousness (1-24 hours) (disorder)
13	Brain stem laceration without open intracranial wound AND with no loss of consciousness (disorder)
13	Brain stem laceration without open intracranial wound AND with prolonged loss of consciousness (more than 24 hours) AND return to pre-existing conscious level (disorder)
13	Brain stem laceration without open intracranial wound AND with prolonged loss of consciousness (more than 24 hours) without return to pre-existing conscious level (disorder)

13	Burkitt's lymphoma of intrapelvic lymph nodes (disorder)
13	Burkitt's lymphoma of lymph nodes of head, face and neck (disorder)
13	Burkitt's lymphoma of lymph nodes of inguinal region and lower limb (disorder)
13	Burkitt's tumor of intrapelvic lymph nodes (disorder)
13	Burkitt's tumor of lymph nodes of head, face AND/OR neck (disorder)
13	Burkitt's tumor of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Burn erythema of lower leg (disorder)
13	Burn of lower leg (disorder)
13	Carbuncle of lower leg (disorder)
13	Carcinoma in situ of skin of lower leg (disorder)
13	Cellulitis and abscess of lower leg (disorder)
13	Chondrocalcinosis due to dicalcium phosphate crystals, of the lower leg (disorder)
13	Chondrocalcinosis due to pyrophosphate crystals of: [lower leg] or [knee] (disorder)
13	Chondrocalcinosis due to pyrophosphate crystals, of the lower leg (disorder)
13	Chondrocalcinosis unspecified, of the lower leg (disorder)
13	Chronic osteomyelitis of lower leg (disorder)
13	Chronic osteomyelitis of the lower leg (disorder)
13	Climacteric arthritis of the lower leg (disorder)
13	Closed anterior dislocation of distal end of femur (disorder)
13	Closed anterior dislocation of proximal end of tibia (disorder)
13	Closed crush injury, lower leg (disorder)
13	Closed dislocation of thoracic and lumbar spine not otherwise specified (disorder)
13	Closed fracture of acromial end of clavicle (disorder)
13	Closed fracture of base of neck of femur (disorder)
13	Closed fracture of distal end of radius (disorder)
13	Closed fracture of distal end of ulna (disorder)
13	Closed fracture of lower end of femur (disorder)
13	Closed fracture of lower end of forearm (disorder)
13	Closed fracture of lower end of humerus (disorder)
13	Closed fracture of lower end of radius AND ulna (disorder)
13	Closed fracture of lumbar spine with spinal cord lesion (disorder)
13	Closed fracture of proximal end of radius (disorder)
13	Closed fracture of proximal end of ulna (disorder)
13	Closed fracture of sternal end of clavicle (disorder)
13	Closed fracture of thoracic spine with spinal cord injury (disorder)
13	Closed fracture of thoracic spine with spinal cord lesion (disorder)
13	Closed fracture of thoracic spine with spinal cord lesion NOS (disorder)
13	Closed fracture of upper end of fibula (disorder)
13	Closed fracture of upper end of forearm (disorder)
13	Closed fracture of upper end of humerus (disorder)
13	Closed fracture of upper end of radius AND ulna (disorder)
13	Closed fracture of upper end of tibia (disorder)
13	Closed lateral dislocation of proximal end of tibia (disorder)
13	Closed medial dislocation of proximal end of tibia (disorder)
13	Closed multiple fractures of lower end of humerus (disorder)
13	Closed multiple fractures of upper end of radius (disorder)

13	Closed multiple fractures of upper end of ulna (disorder)
13	Closed posterior dislocation of distal end of femur (disorder)
13	Closed posterior dislocation of proximal end of tibia (disorder)
13	Closed subluxation of thoracic and lumbar spine NOS (disorder)
13	Closed traumatic dislocation of distal end of fibula (disorder)
13	Closed traumatic dislocation of distal end of tibia (disorder)
13	Closed traumatic dislocation of distal end of ulna (disorder)
13	Closed traumatic dislocation of proximal end of metacarpal bone of wrist (disorder)
13	Compartment syndrome of lower leg (disorder)
13	Congenital abnormality of skull and face bones (disorder)
13	Congenital complete heart block (disorder)
13	Congenital optic disc anomaly (disorder)
13	Congenital undergrowth of distal part of limb (& lower limb) (disorder)
13	Congenital undergrowth of proximal part of limb (& lower limb) (disorder)
13	Contracture of joint of lower leg (disorder)
13	Contusion of knee AND/OR lower leg (disorder)
13	Contusion of lower leg (disorder)
13	Contusion, knee and lower leg (disorder)
13	Contusion, knee and lower leg NOS (disorder)
13	Critical lower limb ischemia (disorder)
13	Crush injury, knee and lower leg (disorder)
13	Crush injury, knee and lower leg NOS (disorder)
13	Crushing injury of knee AND/OR lower leg (disorder)
13	Crushing injury of lower leg (disorder)
13	Crushing injury of skull and intracranial contents (disorder)
13	Crystal arthropathy NOS, of the lower leg (disorder)
13	Crystal arthropathy of lower leg (disorder)
13	Cut of lower leg (disorder)
13	Deep full thickness burn of multiple parts of the face, head or neck - with loss of body part (disorder)
13	Deep full thickness burn of multiple sites of face, head or neck without loss of body part (disorder)
13	Deep full thickness burn of neck, with loss of body part (disorder)
13	Deep full thickness burn of the eye without loss of body part (disorder)
13	Deep full thickness burn of the face, head or neck - without loss of body part (disorder)
13	Deep full thickness burn of the face, head or neck, with loss of body part (disorder)
13	Deep full thickness burn of unspecified part of the face or head - without loss of body part (disorder)
13	Deep full thickness burn of unspecified part of the face or head, with loss of body part (disorder)
13	Deep full thickness burn, with loss of body part, of face, head or neck NOS (disorder)
13	Deep full thickness burn, without loss of body part, of face, head or neck NOS (disorder)
13	Deep partial thickness burn of lower leg (disorder)
13	Deep third degree burn of abdominal wall with loss of body part (disorder)
13	Deep third degree burn of forehead AND/OR cheek with loss of body part (disorder)
13	Deep third degree burn of lower leg (disorder)
13	Deep third degree burn of lower limb with loss of body part (disorder)
13	Deep third degree burn of multiple sites of lower limb with loss of body part (disorder)
13	Deep third degree burn of neck with loss of body part (disorder)
13	Deep third degree burn of toe with loss of body part (disorder)

13	Degenerative joint disease of lower leg (disorder)
13	Degloving injury of lower leg (disorder)
13	Diffuse high grade B-cell lymphoma (disorder)
13	Disorder characterized by back pain (disorder)
13	Disorder of lower leg (disorder)
13	Dog bite of lower leg (disorder)
13	Effusion of joint of lower leg (disorder)
13	Enthesopathy of lower leg and ankle region (disorder)
13	Fetal blood loss from cut end of co-twin's cord (disorder)
13	Fishing hook in lower leg (disorder)
13	Follicular center B-cell lymphoma (nodal/systemic with skin involvement) (disorder)
13	Foreign body of skin of lower leg (disorder)
13	Fracture of acromial end of clavicle (disorder)
13	Fracture of distal end of fibula (disorder)
13	Fracture of distal end of tibia (disorder)
13	Fracture of neck and trunk (disorder)
13	Fracture of neck and trunk NOS (disorder)
13	Fracture of neck AND/OR trunk (disorder)
13	Fracture of proximal end of femur (disorder)
13	Fracture of proximal end of tibia (disorder)
13	Fracture of proximal end of tibia and fibula (disorder)
13	Fracture of skull and facial bones (disorder)
13	Fracture of sternal end of clavicle (disorder)
13	Fracture of upper end of fibula (disorder)
13	Fracture of upper end of humerus (disorder)
13	Fracture of upper end of tibia (disorder)
13	Frostbite with tissue necrosis of knee and lower leg (disorder)
13	Functional disorder of bladder neck and sphincter mechanism (disorder)
13	Glass in lower leg (disorder)
13	Gouty arthritis of the lower leg (disorder)
13	Hemarthrosis of lower leg (disorder)
13	Hemarthrosis of the lower leg (& [knee]) (disorder)
13	Hemarthrosis of the lower leg (disorder)
13	Hernia of abdominal cavity (disorder)
13	High grade B-cell lymphoma (disorder)
13	High grade T-cell lymphoma (disorder)
13	Hodgkin's disease NOS of intrapelvic lymph nodes (disorder)
13	Hodgkin's disease NOS of lymph nodes of head, face and neck (disorder)
13	Hodgkin's disease NOS of lymph nodes of inguinal region and lower limb (disorder)
13	Hodgkin's disease of intrapelvic lymph nodes (disorder)
13	Hodgkin's disease of lymph nodes of head, face AND/OR neck (disorder)
13	Hodgkin's disease of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Hodgkin's disease, lymphocytic depletion of lymph nodes of head, face and neck (disorder)
13	Hodgkin's disease, lymphocytic depletion of lymph nodes of head, face AND/OR neck (disorder)
13	Hodgkin's disease, lymphocytic depletion of lymph nodes of inguinal region and lower limb (disorder)
13	Hodgkin's disease, lymphocytic depletion of lymph nodes of inguinal region AND/OR lower limb (disorder)

13	Hodgkin's disease, lymphocytic-histiocytic predominance of lymph nodes of head, face and neck (disorder)
13	Hodgkin's disease, lymphocytic-histiocytic predominance of lymph nodes of head, face AND/OR neck (disorder)
13	Hodgkin's disease, lymphocytic-histiocytic predominance of lymph nodes of inguinal region and lower limb (disorder)
13	Hodgkin's disease, lymphocytic-histiocytic predominance of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Hodgkin's disease, mixed cellularity of lymph nodes of head, face and neck (disorder)
13	Hodgkin's disease, mixed cellularity of lymph nodes of head, face AND/OR neck (disorder)
13	Hodgkin's disease, mixed cellularity of lymph nodes of inguinal region and lower limb (disorder)
13	Hodgkin's disease, mixed cellularity of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Hodgkin's disease, nodular sclerosis of lymph nodes of head, face and neck (disorder)
13	Hodgkin's disease, nodular sclerosis of lymph nodes of head, face AND/OR neck (disorder)
13	Hodgkin's disease, nodular sclerosis of lymph nodes of inguinal region and lower limb (disorder)
13	Hodgkin's disease, nodular sclerosis of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Hodgkin's granuloma of lymph nodes of head, face and neck (disorder)
13	Hodgkin's granuloma of lymph nodes of head, face AND/OR neck (disorder)
13	Hodgkin's granuloma of lymph nodes of inguinal region and lower limb (disorder)
13	Hodgkin's granuloma of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Hodgkin's paraganuloma of lymph nodes of head, face AND/OR neck (disorder)
13	Hodgkin's paraganuloma of lymph nodes of head, face, and neck (disorder)
13	Hodgkin's paraganuloma of lymph nodes of inguinal region and lower limb (disorder)
13	Hodgkin's paraganuloma of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Hodgkin's sarcoma of lymph nodes of head, face and neck (disorder)
13	Hodgkin's sarcoma of lymph nodes of head, face AND/OR neck (disorder)
13	Hodgkin's sarcoma of lymph nodes of inguinal region and lower limb (disorder)
13	Hodgkin's sarcoma of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Human bite of lower leg (disorder)
13	Hyperemesis gravidarum before end of 22 week gestation with carbohydrate depletion (disorder)
13	Hyperemesis gravidarum before end of 22 week gestation with dehydration (disorder)
13	Hyperemesis gravidarum before end of 22 week gestation with electrolyte imbalance (disorder)
13	Infective arthritis NOS, of the lower leg (disorder)
13	Injury of blood vessels at lower leg level (disorder)
13	Injury of greater saphenous vein at lower leg level (disorder)
13	Injury of lesser saphenous vein at lower leg level (disorder)
13	Injury of multiple blood vessels at lower leg level (disorder)
13	Injury of multiple muscles and tendons at lower leg level (disorder)
13	Injury of multiple nerves at lower leg level (disorder)
13	Injury of muscle and tendon at lower leg level (disorder)
13	Injury of muscle(s) and tendon(s) of anterior muscle group at lower leg level (disorder)
13	Injury of muscle(s) and tendon(s) of peroneal muscle group at lower leg level (disorder)
13	Injury of nerves at lower leg level (disorder)
13	Injury of superficial nerves of head AND/OR neck (disorder)
13	Insect bite, nonvenomous, lower leg (disorder)
13	Joint ankylosis of the lower leg (& [knee]) (disorder)
13	Joint ankylosis of the lower leg (disorder)
13	Joint contracture of the lower leg (& [knee]) contracture (disorder)
13	Joint contracture of the lower leg (disorder)
13	Joint disorder NOS, of the lower leg (disorder)

13	Joint effusion of the lower leg (& [knee]) (disorder)
13	Joint effusion of the lower leg (disorder)
13	Juvenile granulosa cell tumour (disorder)
13	Kashin-Beck disease of the lower leg (disorder)
13	Laceration of lower leg (disorder)
13	Late effect of fracture of skull AND/OR face bones (disorder)
13	Letterer-Siwe disease of intrapelvic lymph nodes (disorder)
13	Letterer-Siwe disease of lymph nodes of head, face and neck (disorder)
13	Letterer-Siwe disease of lymph nodes of head, face AND/OR neck (disorder)
13	Letterer-Siwe disease of lymph nodes of inguinal region and lower limb (disorder)
13	Letterer-Siwe disease of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Leukemic reticuloendotheliosis of lymph nodes of head, face and neck (disorder)
13	Leukemic reticuloendotheliosis of lymph nodes of head, face AND/OR neck (disorder)
13	Leukemic reticuloendotheliosis of lymph nodes of inguinal region and lower limb (disorder)
13	Leukemic reticuloendotheliosis of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Lipoma of lower leg (disorder)
13	Localized osteoarthritis, unspecified, of the lower leg (disorder)
13	Localized, primary osteoarthritis of the lower leg (disorder)
13	Localized, secondary osteoarthritis of the lower leg (disorder)
13	Low back pain (disorder)
13	Lower leg fracture NOS (disorder)
13	Lower limb ischemia (disorder)
13	Lower motor neurone lesion (disorder)
13	Lung disease with systemic lupus erythematosus (disorder)
13	Lymphoplasmacytic B-cell lymphoma, nodal/systemic with skin involvement (disorder)
13	Lymphosarcoma of lymph nodes of head, face and neck (disorder)
13	Lymphosarcoma of lymph nodes of head, face AND/OR neck (disorder)
13	Lymphosarcoma of lymph nodes of inguinal region and lower limb (disorder)
13	Lymphosarcoma of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Malignant histiocytosis of intrapelvic lymph nodes (disorder)
13	Malignant histiocytosis of lymph nodes of head, face and neck (disorder)
13	Malignant histiocytosis of lymph nodes of head, face AND/OR neck (disorder)
13	Malignant histiocytosis of lymph nodes of inguinal region and lower limb (disorder)
13	Malignant histiocytosis of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Malignant lymphoma NOS of intrapelvic lymph nodes (disorder)
13	Malignant lymphoma NOS of lymph nodes of head, face and neck (disorder)
13	Malignant lymphoma NOS of lymph nodes of inguinal region and lower limb (disorder)
13	Malignant lymphoma of intrapelvic lymph nodes (disorder)
13	Malignant lymphoma of lymph nodes of head, face AND/OR neck (disorder)
13	Malignant lymphoma of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Malignant mast cell tumor of intrapelvic lymph nodes (disorder)
13	Malignant mast cell tumor of lymph nodes of head, face AND/OR neck (disorder)
13	Malignant mast cell tumor of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Malignant melanoma of lower leg (disorder)
13	Malignant neoplasm of bones of skull and face (disorder)
13	Malignant neoplasm of bones of skull and face NOS (disorder)

13	Malignant neoplasm of connective and soft tissue of lower leg (disorder)
13	Malignant neoplasm of skin of lower leg (disorder)
13	Malignant tumour of endocrine gland (disorder)
13	Malignant tumour of optic nerve (disorder)
13	Malignant tumour of soft tissue of lower limb (disorder)
13	Mantle cell B-cell lymphoma (nodal/systemic with skin involvement) (disorder)
13	Massive multi tissue damage lower leg (disorder)
13	Mast cell malignancy of intrapelvic lymph nodes (disorder)
13	Mast cell malignancy of lymph nodes of head, face and neck (disorder)
13	Mast cell malignancy of lymph nodes of inguinal region and lower limb (disorder)
13	Metal foreign body in lower leg (disorder)
13	Mixed cell type lymphosarcoma of intrapelvic lymph nodes (disorder)
13	Mixed cell type lymphosarcoma of lymph nodes of head, face, and neck (disorder)
13	Mixed cell type lymphosarcoma of lymph nodes of inguinal region and lower limb (disorder)
13	Multiple closed fractures of lower end of femur (disorder)
13	Multiple closed fractures of skull and face with intracranial injury of other and unspecified nature (disorder)
13	Multiple closed fractures of skull and face with other and unspecified intracranial hemorrhage (disorder)
13	Multiple closed fractures of skull AND/OR face with cerebral laceration AND/OR contusion (disorder)
13	Multiple closed fractures of skull AND/OR face with subarachnoid, subdural AND/OR extradural hemorrhage (disorder)
13	Multiple closed fractures of skull AND/OR face without intracranial injury (disorder)
13	Multiple fractures of lower leg (disorder)
13	Multiple open fractures of lower end of femur (disorder)
13	Multiple open fractures of skull and face with intracranial injury of other and unspecified nature (disorder)
13	Multiple open fractures of skull and face with other and unspecified intracranial hemorrhage (disorder)
13	Multiple open fractures of skull AND/OR face with cerebral laceration AND/OR contusion (disorder)
13	Multiple open fractures of skull AND/OR face with subarachnoid, subdural AND/OR extradural hemorrhage (disorder)
13	Multiple open fractures of skull AND/OR face without intracranial injury (disorder)
13	Multiple open wounds of lower leg (disorder)
13	Multiple superficial injuries of lower leg (disorder)
13	Myalgia/myositis - lower leg (disorder)
13	Mycosis fungoides of lymph nodes of head, face AND/OR neck (disorder)
13	Mycosis fungoides of lymph nodes of inguinal region and lower limb (disorder)
13	Mycosis fungoides of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Mycosis fungoides of the lymph nodes of head, face and neck (disorder)
13	Nail wound of lower leg (disorder)
13	Needle stick injury of lower leg (disorder)
13	Neoplasm of blood vessel of lower leg (disorder)
13	Neoplasm of lower leg (disorder)
13	Neoplasm of superficial inguinal lymph nodes (disorder)
13	Neoplasm of supraclavicular lymph nodes (disorder)
13	Neoplasm of uncertain behavior of skin of lower leg (disorder)
13	New onset angina (disorder)
13	Nodular high grade B-cell lymphoma (disorder)
13	Nodular lymphoma of lymph nodes of head, face and neck (disorder)
13	Nodular lymphoma of lymph nodes of head, face, and neck (disorder)
13	Oligoarticular osteoarthritis, unspecified, of lower leg (disorder)

13	On examination - left eye background diabetic retinopathy (disorder)
13	On examination - left eye completely blind (disorder)
13	On examination - left eye diabetic maculopathy (disorder)
13	On examination - left eye preproliferative diabetic retinopathy (disorder)
13	On examination - left eye proliferative diabetic retinopathy (disorder)
13	On examination - Left eye rubeosis iridis (disorder)
13	On examination - left eye stable treated proliferative diabetic retinopathy (disorder)
13	Open anterior dislocation of distal end of femur (disorder)
13	Open anterior dislocation of proximal end of tibia (disorder)
13	Open crush injury, lower leg (disorder)
13	Open dislocation of distal end of fibula (disorder)
13	Open dislocation of distal end of tibia (disorder)
13	Open dislocation of distal end of ulna (disorder)
13	Open dislocation of proximal end of metacarpal bone of wrist (disorder)
13	Open fracture of acromial end of clavicle (disorder)
13	Open fracture of distal end of radius (disorder)
13	Open fracture of distal end of ulna (disorder)
13	Open fracture of femur: [proximal, basicervical] or [of base of neck] (disorder)
13	Open fracture of lower end of femur (disorder)
13	Open fracture of lower end of forearm (disorder)
13	Open fracture of lower end of humerus (disorder)
13	Open fracture of lower end of radius AND ulna (disorder)
13	Open fracture of lumbar spine with spinal cord lesion (disorder)
13	Open fracture of proximal end of radius (disorder)
13	Open fracture of proximal end of ulna (disorder)
13	Open fracture of sternal end of clavicle (disorder)
13	Open fracture of thoracic spine with spinal cord injury (disorder)
13	Open fracture of thoracic spine with spinal cord lesion (disorder)
13	Open fracture of thoracic spine with spinal cord lesion NOS (disorder)
13	Open fracture of upper end of fibula (disorder)
13	Open fracture of upper end of forearm (disorder)
13	Open fracture of upper end of humerus (disorder)
13	Open fracture of upper end of radius AND ulna (disorder)
13	Open fracture of upper end of tibia (disorder)
13	Open lateral dislocation of proximal end of tibia (disorder)
13	Open medial dislocation of proximal end of tibia (disorder)
13	Open multiple fractures of upper end of radius (disorder)
13	Open multiple fractures of upper end of ulna (disorder)
13	Open posterior dislocation of distal end of femur (disorder)
13	Open posterior dislocation of proximal end of tibia (disorder)
13	Open wound of lower leg (disorder)
13	Open wound of lower leg with complication (disorder)
13	Open wound of lower leg with tendon involvement (disorder)
13	Open wound of lower limb (disorder)
13	Osteitis of lower leg (disorder)
13	Osteoarthritis NOS, of the lower leg (disorder)

13	Osteoarthritis NOS: [lower leg] or [knee] (disorder)
13	Osteoarthritis: [localised of the lower leg, unspecified] or [patellofemoral] (disorder)
13	Osteomyelitis of lower leg (disorder)
13	Other anomaly of skull or face bone NOS (disorder)
13	Other congenital anomalies of lower leg (disorder)
13	Other crystal arthropathies of the lower leg (disorder)
13	Other infections involving bone, of the lower leg (disorder)
13	Other infectious and parasitic diseases with arthropathy of the lower leg (disorder)
13	Other joint derangement NEC, of the lower leg (disorder)
13	Other obstetric pelvic organ damage with postnatal problem (disorder)
13	Other specified arthropathy of the lower leg (disorder)
13	Other specified joint disorders of the lower leg (disorder)
13	Pain in thoracic spine (disorder)
13	Palindromic rheumatism of the lower leg (disorder)
13	Partial failure replanted lower leg (disorder)
13	Pathological dislocation of lower leg (disorder)
13	Pathological dislocation of the lower leg (& [knee]) (disorder)
13	Pathological dislocation of the lower leg (disorder)
13	Pellet wound of lower leg (disorder)
13	Periostitis of lower leg (disorder)
13	Periostitis without mention of osteomyelitis, of the lower leg (disorder)
13	Poliomyelitis osteopathy of the lower leg (disorder)
13	Post-dysenteric reactive arthropathy of the lower leg (disorder)
13	Posterior end of inferior turbinate hypertrophy (disorder)
13	Postoperative complete heart block (disorder)
13	Post-phlebitic dermatosis of lower leg (disorder)
13	Predislocation status of hip at birth (disorder)
13	Primary cutaneous diffuse large cell B-cell lymphoma (disorder)
13	Primary cutaneous diffuse large cell B-cell lymphoma of the leg (disorder)
13	Primary localized osteoarthritis of lower leg (disorder)
13	Primary malignant neoplasm of blood vessel of lower leg (disorder)
13	Psychogenic back pain (disorder)
13	Puncture wound of lower leg (disorder)
13	Pyogenic arthritis of lower leg (disorder)
13	Pyogenic arthritis of the lower leg (disorder)
13	Pyogenic arthritis of: [knee] or [lower leg] (disorder)
13	Recurrent dislocation of lower leg (disorder)
13	Recurrent joint dislocation, of the lower leg (& [Knee] or [patella]) (disorder)
13	Recurrent joint dislocation, of the lower leg (disorder)
13	Reticulosarcoma of lymph nodes of head, face and neck (disorder)
13	Reticulosarcoma of lymph nodes of head, face, and neck (disorder)
13	Rupture of eye with partial loss of intraocular tissue (disorder)
13	Rupture of tendon of lower leg and ankle (disorder)
13	SÃ©zary's disease of lymph nodes of head, face and neck (disorder)
13	SÃ©zary's disease of lymph nodes of head, face AND/OR neck (disorder)
13	SÃ©zary's disease of lymph nodes of inguinal region and lower limb (disorder)

13	SÃ©zary's disease of lymph nodes of inguinal region AND/OR lower limb (disorder)
13	Second degree burn of lower leg (disorder)
13	Secondary and unspecified malignant neoplasm of inguinal and lower limb lymph nodes (disorder)
13	Secondary and unspecified malignant neoplasm of inguinal and lower limb lymph nodes NOS (disorder)
13	Secondary and unspecified malignant neoplasm of intrapelvic lymph nodes (disorder)
13	Secondary and unspecified malignant neoplasm of intrapelvic lymph nodes NOS (disorder)
13	Secondary and unspecified malignant neoplasm of lymph nodes of head, face and neck (disorder)
13	Secondary and unspecified malignant neoplasm of lymph nodes of head, face and neck NOS (disorder)
13	Secondary and unspecified malignant neoplasm of superficial cervical lymph nodes (disorder)
13	Secondary and unspecified malignant neoplasm of superficial inguinal lymph nodes (disorder)
13	Secondary and unspecified malignant neoplasm of superficial mesenteric lymph nodes (disorder)
13	Secondary and unspecified malignant neoplasm of superficial parotid lymph nodes (disorder)
13	Secondary and unspecified malignant neoplasm of superficial tracheobronchial lymph nodes (disorder)
13	Secondary and unspecified malignant neoplasm of supratrochlear lymph nodes (disorder)
13	Secondary localized osteoarthritis of lower leg (disorder)
13	Secondary malignant neoplasm of blood vessel of lower leg (disorder)
13	Secondary malignant neoplasm of intrapelvic lymph nodes (disorder)
13	Secondary malignant neoplasm of lymph nodes of lower limb (disorder)
13	Secondary malignant neoplasm of lymph nodes of neck (disorder)
13	Secondary malignant neoplasm of superficial inguinal lymph nodes (disorder)
13	Secondary malignant neoplasm of supraclavicular lymph nodes (disorder)
13	Sequelae of injuries of neck and trunk (disorder)
13	Severe multi tissue damage lower leg (disorder)
13	Sexually acquired reactive arthropathy of the lower leg (disorder)
13	Soft tissue lesion of lower leg and ankle (disorder)
13	Spinal cord/nerve root injury (disorder)
13	Splinter of lower leg, without major open wound (disorder)
13	Splinter of lower leg, without major open wound, infected (disorder)
13	Stab wound of lower leg (disorder)
13	Superficial frostbite of knee and lower leg (disorder)
13	Superficial injury of lower leg (disorder)
13	Superficial injury of lower leg NOS (disorder)
13	Superficial injury of lower leg NOS, infected (disorder)
13	Superficial partial thickness burn of lower leg (disorder)
13	Systemic lupus erythematosus (disorder)
13	Third degree burn of lower leg (disorder)
13	Total failure replanted lower leg (disorder)
13	Transient arthropathy of the lower leg (disorder)
13	Traumatic arthropathy of the lower leg (disorder)
13	Traumatic blister of lower leg (disorder)
13	Traumatic blister of lower leg, infected (disorder)
13	Traumatic brain injury with brief loss of consciousness (disorder)
13	Traumatic brain injury with loss of consciousness (disorder)
13	Traumatic brain injury with moderate loss of consciousness (disorder)
13	Traumatic brain injury with no loss of consciousness (disorder)
13	Traumatic brain injury with prolonged loss of consciousness (disorder)

13	Tuberculosis of the lower leg bone (disorder)
13	Tumour of optic nerve (disorder)
13	Unspecified malignant neoplasm of lymphoid and histiocytic tissue of intrapelvic lymph nodes (disorder)
13	Unspecified malignant neoplasm of lymphoid and histiocytic tissue of lymph nodes of head, face and neck (disorder)
13	Unspecified malignant neoplasm of lymphoid and histiocytic tissue of lymph nodes of inguinal region and lower limb (disorder)
13	Unspecified monoarthritis of the lower leg (disorder)
13	Unspecified osteomyelitis of the lower leg (disorder)
13	Unspecified polyarthropathy of the lower leg (disorder)
13	Unspecified thickness burn of lower leg (disorder)
13	Villonodular synovitis of the lower leg (disorder)
13	Wood splinter in lower leg (disorder)