

**Evaluating the Coverage of Health Informatics Terms in the  
MeSH Thesaurus**

**Critical Analysis of a Newly Created Health Informatics Glossary**

**By**

**Li Dong**

**B00496164**

**lz484229@cs.dal.ca**

**Performed at**

**Medical Informatics**

**5849 University Avenue**

**Halifax, Nova Scotia B3H 4H7**

**In partial fulfillment of the requirements of the Master of Health Informatics Program,**

**Dalhousie University**

**Report of Internship for the period February 16 – May 15, 2009**

**Date Submitted: May 31, 2009**

**Supervisor: Dr. Grace Paterson**

**Signature:** \_\_\_\_\_

**Date: May 31, 2009**

### **Acknowledgement and Endorsement**

This report has been written by me and has not received any previous academic credit at this or any other institution.

I would like to thank Dr. Grace Paterson and Dr. Swarna Weerasinghe for their supervision in this internship and Ms. Deirdre Harvey for her support during this internship.

Last but not the least, I am indebted to my outstanding professors in the Master of Health Informatics program at Dalhousie University who have opened my eyes to this new area of study and taught me valuable research and analytical skills that have not only benefited me in writing up this report but I am certain will continue to benefit me in all my future academic and professional endeavors: Dr. Grace Paterson of *Health Information Flow and Standards*, Dr. David Zitner of *Health Information Flow and Use*, and Dr. Syed Sibte Raza Abidi of *Knowledge Management of Health Informatics*, among others.

Li Dong

\_\_\_\_\_

Signature

\_\_\_\_\_

Date

## Executive Summary

As an integral part of the Master in Health Informatics (MHI) program at Dalhousie University, internship provides a good opportunity for MHI students to apply academic health informatics knowledge in healthcare work to real world with the purpose of improving the students' critical thinking ability through hand-on experiences. The 13-week internship was performed at Medical Informatics at Dalhousie University, Halifax, Nova Scotia. It consists of two components: (i) Updating *Medicine Critical Appraisal for Postgraduate Medical Education*, an online Blackboard Learning System course in Dalhousie University, and adding new contents in the field of Health Informatics; (ii) Developing a new glossary specialized in Health Informatics and evaluating the coverage of this glossary in MeSH thesaurus. During the internship, the intern acquired comprehensive knowledge of critical appraisal in medical and health informatics and honed critical thinking skills in identifying problems and coming up with health informatics solutions. In developing and evaluating a new glossary in health informatics as part of the internship, the intern found that the new glossary, based on key words from ITCH 2009 Conference Proceedings, couldn't match MeSH thesaurus very well. Although this is consistent with the results of previous studies, the intern recommended that further steps be taken to analyze the reasons in more depth. Finally, the intern identified certain gaps between what is needed in the updated course and what is currently available and recommended further actions to bridge the gap.

## Table of Contents

<b>1</b>	<b>Introduction .....</b>	<b>4</b>
<b>1.1</b>	<b>Overview .....</b>	<b>4</b>
<b>1.2</b>	<b>Introduction of the Organization .....</b>	<b>4</b>
<b>1.3</b>	<b>Introduction of the Internship.....</b>	<b>5</b>
<b>1.3.1</b>	<b>Part I: Updating BLS Course Contents.....</b>	<b>5</b>
<b>1.3.1.1</b>	<b>Background .....</b>	<b>5</b>
<b>1.3.1.2</b>	<b>Description of Work and Lessons Learned .....</b>	<b>6</b>
<b>1.3.2</b>	<b>Part II: Developing Glossary in Health Informatics and Evaluating Coverage of the Glossary in MeSH Thesaurus.....</b>	<b>6</b>
<b>1.3.2.1</b>	<b>Background .....</b>	<b>6</b>
<b>1.3.2.2</b>	<b>Description of Work and Lessons Learned .....</b>	<b>7</b>
<b>2</b>	<b>Health Informatics Relevance .....</b>	<b>7</b>
<b>2.1</b>	<b>Part I: Updating BLS Course Contents.....</b>	<b>8</b>
<b>2.2</b>	<b>Part II: Developing Glossary in Health Informatics and Evaluating Coverage of the Glossary in MeSH Thesaurus.....</b>	<b>8</b>
<b>3</b>	<b>Evaluating MeSH Coverage of Health Informatics Terms in Glossary.....</b>	<b>9</b>
<b>3.1</b>	<b>Literature Review.....</b>	<b>9</b>
<b>3.2</b>	<b>Research Questions .....</b>	<b>11</b>
<b>3.3</b>	<b>Methods.....</b>	<b>11</b>
<b>3.3.1</b>	<b>Materials .....</b>	<b>11</b>
<b>3.3.2</b>	<b>Method for Glossary Development .....</b>	<b>12</b>
<b>3.3.3</b>	<b>Method for Evaluation of the Glossary Coverage in MeSH.....</b>	<b>13</b>
<b>3.3.3.1</b>	<b>Making the Corpus.....</b>	<b>13</b>
<b>3.3.3.2</b>	<b>Manual Term Selection .....</b>	<b>13</b>
<b>3.3.3.3</b>	<b>Mapping to Vocabulary Sources .....</b>	<b>13</b>
<b>3.3.3.4</b>	<b>Coverage Evaluation .....</b>	<b>14</b>
<b>3.3.4</b>	<b>Results .....</b>	<b>14</b>
<b>3.3.4.1</b>	<b>Corpus and Terminology Data .....</b>	<b>14</b>
<b>3.3.4.2</b>	<b>Global Coverage and MeSH Coverage.....</b>	<b>15</b>
<b>3.3.4.3</b>	<b>Category Coverage.....</b>	<b>17</b>
<b>3.3.5</b>	<b>Discussion.....</b>	<b>18</b>
<b>3.3.5.1</b>	<b>Global Coverage and MeSH Coverage.....</b>	<b>18</b>
<b>3.3.5.2</b>	<b>Category Coverage.....</b>	<b>19</b>
<b>4</b>	<b>Conclusion.....</b>	<b>21</b>
<b>5</b>	<b>Recommendation.....</b>	<b>22</b>
	<b>Reference.....</b>	<b>23</b>

# 1 Introduction

## 1.1 Overview

The 13-week internship was performed at Medical Informatics at Dalhousie University, Halifax, Nova Scotia. It consists of two components:

- I. Updating *Medicine Critical Appraisal for Postgraduate Medical Education*, an online Blackboard Learning System course in Dalhousie University, and adding new contents in the field of Health Informatics. This was undertaken during the first eight weeks of the internship.
- II. Developing a new glossary specialized in Health Informatics and evaluating the coverage of this glossary in MeSH thesaurus. This was undertaken during the remaining five weeks of the internship.

These two parts, successfully completed, would lay the groundwork for the publication of *Critical Appraisal Handbook for Medical Education, Health Informatics, and Pharmacy Students*, a project undertaken by Dr. Grace Paterson and Dr. Swarna Weerasinghe.

## 1.2 Introduction of the Organization

Medical Informatics at the Faculty of Medicine of Dalhousie University was established in 1996. Applying computing and information technology to medical research, teaching, patient care and hospital administration, Medical Informatics covers the study of biomedical-systems technology, health-information systems, medical-decision-support systems, image and signal analysis, statistical techniques and modeling, education and training, health-care management, and human-machine interface.

In 2002, in collaboration with the Faculty of Computer Science, Dalhousie University, Medical Informatics at the Faculty of Medicine established the first graduate program in Health

Informatics in Canada, which integrates technology into the research of health care and studies the use of health information in support of clinical care and health services administration.

## **1.3 Introduction of the Internship**

### **1.3.1 Part I: Updating BLS Course Contents**

#### **1.3.1.1 Background**

Critical appraisal, an essential part of evidence-based clinical practice, is “the process of systematically examining research evidence to assess its validity, results and relevance before using it to inform a decision.”[1] It is one of the key skills used by postgraduate students in medicine, health informatics and pharmacy to critically appraise the effectiveness of health literatures. For clinical practitioners and health informaticians, critical appraisal is essential in making effective and accurate patient-oriented decisions based on existing health evidence. With the aim of teaching medical students critical appraisal skills and improving these skills for clinical residents, health informaticians and pharmacists, Dr. Grace Paterson and Dr. Swarna Weerasinghe from Medical Education Division of Faculty of Medicine at Dalhousie University came up with the idea of publishing a book to be titled *Critical Appraisal Handbook for Medical Education, Health Informatics, and Pharmacy Students*.

The online course *Medicine Critical Appraisal for Postgraduate Medical Education* was developed by the Department of Community Health & Epidemiology and the Division of Postgraduate Medical Education at Dalhousie University in 1999 with the purpose of improving critical appraisal skills for residents who rotate through hospitals in the three Maritime Provinces of Canada, and was last offered in 2004. The coverage of health informatics in the original course is limited since it was originally designed for clinical residents. The rapid growth of health science and the growing application of information technology in healthcare in recent years make it necessary to update the online course to reflect the new developments in health informatics. Building upon the updated course contents, the Handbook would be more relevant and better meet the growing needs of a wider readership.

### **1.3.1.2 Description of Work and Lessons Learned**

Updating the original online BLS course *Medicine Critical Appraisal for Postgraduate Medical Education* is the first component of the internship.

Upon reviewing the original course contents, the intern identified outdated contents including old definitions, irrelevant statistical methodologies, and malfunctioning links. The intern further investigated latest tools of critical appraisal such as checklists, worksheets, CATmaker (an online tool for generating Critical Appraisal Topics for health research), and VisualRx (an online calculator to help calculate NNT). Finally, the intern investigated and gathered relevant knowledge and tools used in the critical appraisal of articles on Economic Evaluation and Quality of Care, an added component to the updated course contents.

Through these experiences, the intern acquired comprehensive knowledge of critical appraisal in medical and health informatics and honed critical thinking skills in identifying problems and coming up with health informatics solutions.

## **1.3.2 Part II: Developing Glossary in Health Informatics and Evaluating Coverage of the Glossary in MeSH Thesaurus**

### **1.3.2.1 Background**

A glossary is an alphabetical list of specialized terms in a particular domain of knowledge with the definitions for those terms [2]. It is usually listed at the end of a book or included in the manual of a software. The terms included in the glossary are either specialized in the domain or newly introduced and not common for the users of the book or software. Adding health informatics contents to the original online course requires the simultaneous development of a new glossary expanded to cover health informatics.

Controlled vocabulary is used in subject indexing schemes, subject headings, thesauri and taxonomies in order to facilitate organizing knowledge for subsequent information retrieval

[3].From the professional perspective, a controlled vocabulary is more effective and efficient in indexing information retrieval than a natural language vocabulary.

A glossary with the highest level of consistency with a controlled vocabulary is most useful in indexing relevant articles. MeSH is an important controlled vocabulary in health science and is universally regarded as a powerful indexing tool in retrieving articles in health science. As such, it is essential to maximize the coverage of a glossary by MeSH Thesaurus to improve health informatics information retrieval.

### **1.3.2.2 Description of Work and Lessons Learned**

Developing a new glossary in health informatics to form part of the expanded glossary for the updated online course is another component of the internship. Using the methodology for developing the Cochrane Glossary, the intern identified and extracted key words in Health Informatics from ITCH 2009 Conference Proceedings, a collection of the latest publications in health informatics, mapped and retrieved their definitions from different sources such as key textbooks and controlled vocabularies. The intern then evaluated the coverage of the newly created glossary by these sources. During this stage of the internship, the intern learned the latest evolvement of relevant concepts and definitions in health informatics and obtained a basic understanding of the representativeness of several chosen vocabulary sources in health informatics, especially the effectiveness of MeSH as an indexing vocabulary in health informatics articles retrieval. Finally, the intern analyzed the results and raised constructive suggestions on improving the glossary as the next step of work.

## **2 Health Informatics Relevance**

Both components of the internship are relevant to the study of health informatics, as discussed in more details below.

## **2.1 Part I: Updating BLS Course Contents**

Health Informatics is an interdisciplinary science which integrates information technology and communication skills into medical care, aiming to improve health outcomes and processes for health consumers [4]. The word “informatics” is not limited to computer involvement; rather, it broadly covers information, technology, processes, analytical tools and techniques, governance and the skills, all approaches that can be employed to improve healthcare [5].

Critical appraisal skills, widely used for understanding and assessing the methods and results of health articles, also play an important role in health informatics. By way of example, reminder system uses artificial intelligence to improve timely delivery of preventive and acute medical care to patients [6, 7]. Safety and effectiveness are two major considerations in evaluating the system. Past relevant evidence could provide comprehensive information in this regard. However, lacking solid critical appraisal skills in either medical evidence or information technology evidence, many IT professionals find it challenging to identify the best evidence from previous experiences and develop truly worthwhile AI systems in clinical practice. Educating health informaticians and equipping them with strong critical appraisal skills could improve their ability in evidence-based decision making and allow them to provide better health care.

## **2.2 Part II: Developing Glossary in Health Informatics and Evaluating Coverage of the Glossary in MeSH Thesaurus**

Health Informatics is a new, evolving discipline. Numerous health informatics articles are being published scattering at a wide range of literatures such as Medical Science and Information Technology and Engineering. Effective information retrieval is critical in research and further advancement of the discipline.

The wide utilization of Internet offers an easily accessible approach. However, effective retrieving of health information is still problematic due to the lack of one home discipline [8].

Controlled vocabulary provides an alternative approach. However, problems arise when there is mismatch between the key words used in the publications and the terms in the controlled vocabulary. Zheng et al [9] showed that inconsistency between the terminology of the health website and the vocabularies used in retrieving health information could decrease the effectiveness and efficiency of health information retrieval.

Medical Subject Heading (MeSH) is a comprehensive controlled vocabulary in life science and is used to index MEDLINE journal articles and book catalogs and to facilitate articles searching through Pubmed. Evaluating the coverage of health informatics glossaries in MeSH helps to identify if there is gap between the glossary and the controlled vocabulary and to guide further research, if warranted, to bridge the gap and increase the effectiveness and efficiency of health informatics information retrieval.

## **3 Evaluating MeSH Coverage of Health Informatics Terms in Glossary**

### **3.1 Literature Review**

As a constantly evolving, multidisciplinary field related to health science, Health Informatics draws terminologies from a variety of disciplines such as Medicine, Information Technology, and Computer Science, which makes the use of the terminologies more complicated and problem-prone.

The formal work in the terminology of medical informatics commenced at National Library of Medicine (NLM), where the first L Tree in MeSH vocabulary was developed in 1960 to cover medical informatics terminology [10]. Despite its regular updates, there still remains a wide area of informatics not represented in the MeSH vocabulary [11].

Being a controlled vocabulary in health science, MeSH is widely used by the most important catalogues in health [12-16], and is universally accepted as a powerful indexing tool in searching health articles in MEDLINE via Pubmed [17]. One oft-cited advantage of MeSH indexing is that it is not limited to terms that match exactly. For terms not found in MeSH, the vocabulary will automatically generate a list of MeSH preferred terms with same or close meanings.

As with any other vocabularies, the extensiveness of MeSH thesaurus is critical in effective information retrieval. One previous study contributed to the extensiveness of MeSH coverage in Health Informatics. N. Ogg and others examined the MeSH thesaurus and the terminologies in Medical Informatics literature and found out that MeSH did not provide a satisfactory coverage of Medical Informatics terminologies [18]. Based on these findings, they suggested the development of a specialized vocabulary, which could reflect the multi-disciplinary nature of Health Informatics and could rapidly change as the discipline evolved.

Even though the existing MeSH has proven to be limited in the coverage of Health Informatics terminologies, it is still the best controlled vocabulary in the discipline. V. Mary and others assessed the coverage of MeSH and other specialized terminologies, e.g. UMLS and some gene banks, in Molecular Biology terminologies, and proved that MeSH still had the best coverage in that field [19].

For scientists and professionals in specific domains, using standardized terminologies in controlled vocabulary such as MeSH thesaurus will no doubt facilitate the retrieval of relevant information. However, for people conducting non-professional searches, the use of standard terminology might get the opposite result. Raza Abidi and others studied if consumer health queries formulated with MeSH terminology could improve the results of search query. They drew the negative answer. As Abidi explained in the article, the lower query results could be attributed to the fact that the subjects selected in the study were lay people who lack knowledge of medical terminology [20].

## 3.2 Research Questions

As discussed before, the evolution of a discipline will always generate new terms, which ideally should be represented in the glossary. A comprehensive glossary reflecting the latest development of the discipline can serve as a good database for authors to choose their terms from. The consistency between the glossary and the controlled vocabulary will maximize the effectiveness and efficiency of information retrieval.

This study attempts to evaluate the coverage of the new glossary in the MeSH thesaurus and to take the first step in answering the following questions:

- Does the MeSH vocabulary represent health informatics terminology?
- Does the new glossary represent the evolution of health informatics terminology?

## 3.3 Methods

### 3.3.1 Materials

#### **ITCH2009 Conference Proceedings**

*Information Technology and Communications in Health* (ITCH) is an international conference which focuses on the discussion of the impact of information technology on community health. The latest ITCH conference was held in 2009 with the theme of “Revolutionizing Health Care with Informatics: From Research to Practice.” The ITCH 2009 conference aimed at encouraging experts to demonstrate and share their experiences and knowledge in health informatics. A total of 84 articles on health informatics evolution were included in ITCH 2009 Conference Proceedings, from which 288 key words were chosen by the authors.

#### **Guide to Health Informatics 2<sup>nd</sup> Edition by Enrico Coiera**

*The Guide to Health Informatics* has been written for healthcare professionals who wish to understand the principles and applications of information and communication systems in

healthcare [6]. Last revised in 2003, it serves as a key textbook in many health/Medical informatics programs.

**Biomedical Informatics *Computer Applications in Health Care and Biomedicine*, Third Edition by E.H. Shortliffe**

As another key textbook focusing on the areas of Biomedical Informatics, Health Informatics and Medical Computing, *Biomedical Informatics Computer Application in Health Care and Biomedicine* was written for medical informatics specialist, healthcare administrators, professors and students. Last revised in June 2006, the book studies the role of computers in the provision of medical services and provides both a conceptual framework and a practical approach for the implementation and management of IT used to improve the delivery of health care [21].

**Medical Subject Heading (MeSH)**

As a controlled vocabulary composed of hierarchical terms, MeSH is published and updated by NLM with the purpose of facilitating the searching of biomedical journals at different specificity in PubMed. With the ability to index articles in PubMed from 4600 biomedical journal, MeSH is seen as one of the most powerful searching tools in health science.

### **3.3.2 Method for Glossary Development**

Following the Cochrane Glossary's approach developed by Cochrane Collaboration's Patient-Reported Outcomes Methods Group, the intern undertook the following steps:

- 1) Extracting and selecting Health Informatics relevant key terms from ITCH 2009 Conference Proceedings;
- 2) Retrieving definitions from original articles in ITCH 2009 Conference Proceedings;

- 3) Mapping key terms to the glossaries of *Guide to Health Informatics 2<sup>nd</sup> Edition* and *Biomedical Informatics Computer Applications in Health Care and Biomedicine, Third Edition* and identifying their definitions;

In order to evaluate the coverage of the glossary terms in MeSH, the intern took several additional steps not covered in Cochrane glossary development strategy:

- 1) Mapping key terms to MeSH database via Pubmed and identifying relevant definitions;
- 2) Identifying MeSH preferred terms and corresponding definitions for terms not exactly mapped.

### **3.3.3 Method for Evaluation of the Glossary Coverage in MeSH**

#### **3.3.3.1 Making the Corpus**

As the first step, the corpus was formed, drawing terms from tertiary sources in health informatics. The key words from all the journals in ITCH2009 Conference Proceedings were chosen to form the database of original terms, which were assigned to 15 categories representing different sub-disciplines in health informatics.

#### **3.3.3.2 Manual Term Selection**

Among all the original key words in the corpus, those not relevant to Health Informatics such as a city name (e.g. Hong Kong), a clinical terminology (e.g. cardiovascular disease), and a non-Health Informatics term (e.g. Student Recruitment) were manually removed. Only key words in the field of health informatics remained.

#### **3.3.3.3 Mapping to Vocabulary Sources**

The corpus was mapped to textbook glossaries to identify the exact matches (lexemic match included). The corpus was further mapped to the MeSH database to identify the terms and

definitions through two approaches: exact match and synonym match. The latter matches different terms with same or closely related concepts, suggested by MeSH as Preferred Terms.

### **3.3.3.4 Coverage Evaluation**

The coverage analysis was performed using three approaches:

- 1) Individual analysis of each term;
- 2) Global analysis of each vocabulary source; and
- 3) Category analysis of each sub-discipline.

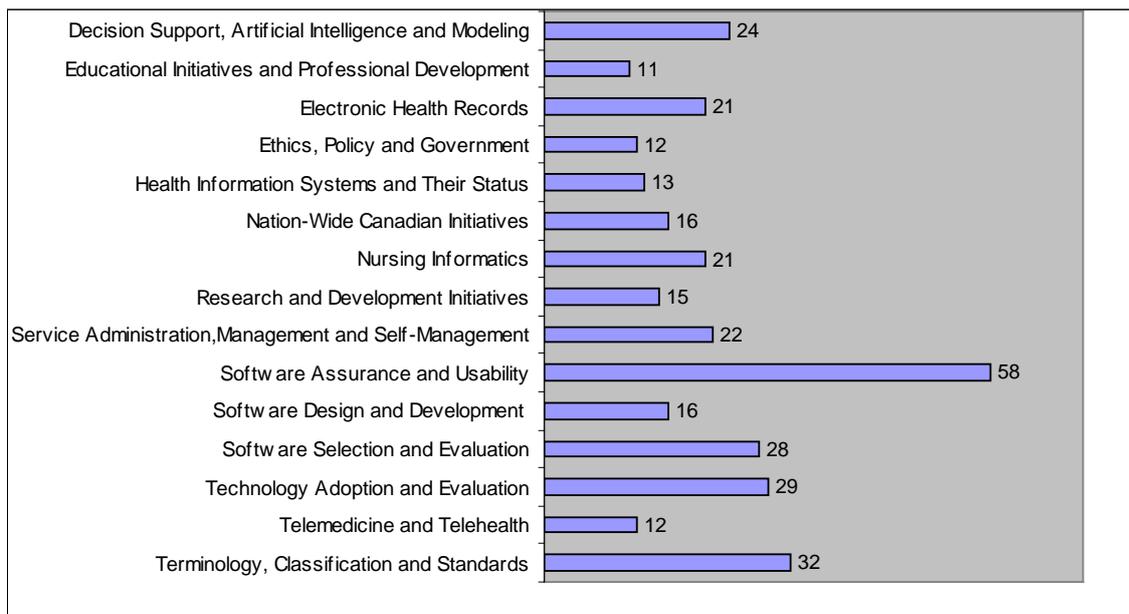
As discussed above, three key cover types were identified for individual term analysis:

- 1) Exact Match: The term in the glossary corresponds to an exactly matched single concept in the vocabulary (including lexemic Match);
- 2) Synonym Match: The term in the glossary corresponds to different terms in the vocabulary with same or close meaning (MeSH preferred terms); and
- 3) Not Mapped: Neither lexemic nor synonym terms were found.

## **3.3.4 Results**

### **3.3.4.1 Corpus and Terminology Data**

A total of 84 articles in ITCH 2009 Conference Proceedings were selected, with 288 key words (net of duplications) used by authors. After manual elimination of the terms not relevant to Health Informatics, 266 distinct terms remained in the corpus for further evaluation. They were assigned to 15 categories, each representing a sub-discipline of Health Informatics based on the classifications in ITCH2009 Conference Proceedings. Some terms could belong to more than one category. The original number of the terms by categories is in Table 1.



Table

1: Key Words by Categories

As can be seen in Table 1, among fifteen categories in Health Informatics area identified in ITCH2009 Conference Proceedings, the number of original terms range from 58 in “Software Assurance and Usability” to 11 in “Educational Initiatives and Professional Development”, out of a total of 266.

### 3.3.4.2 Global Coverage and MeSH Coverage

The global coverage of each vocabulary source is illustrated in Figure 1. General Match includes exact match or synonym match. If a term was mapped to more than one definition from multiple sources, it was only counted once. For the convenience of illustration, (C) and (S) are used to denote the two textbooks “*The Guide to Health Informatics*” and “*Biomedical Informatics*”, respectively.

Generally, 39% of the terms in corpus were matched to one of the three vocabulary sources. MeSH thesaurus covered 35% of terms by either exact match or synonym match. *The Guide to Health Informatics* has the lowest coverage of the terms in the new glossary (2%).

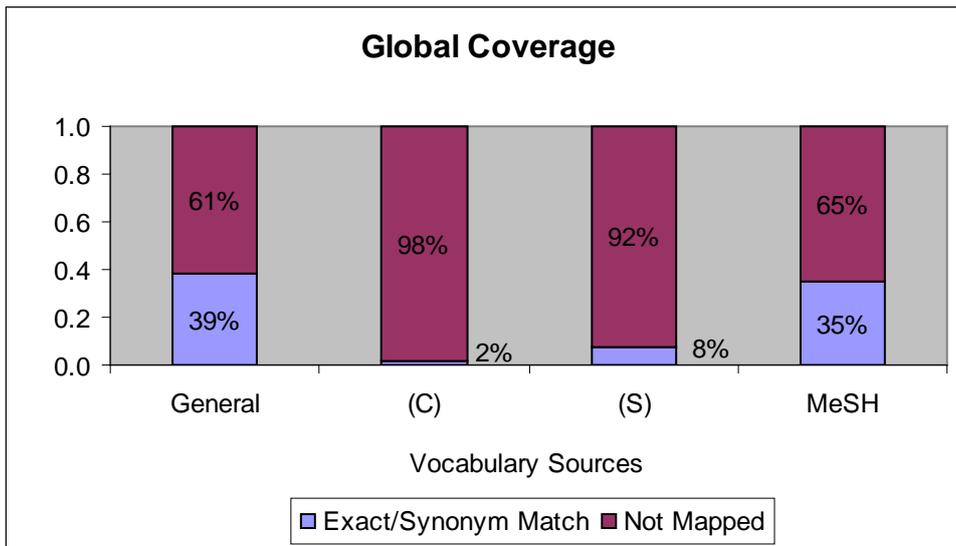


Figure 1: Global Coverage

Figure 2 shows the coverage of the glossary in the MeSH thesaurus. As it can be seen in the figure, only 11% of the terms could find exact matches in MeSH, another 25% could be mapped by synonym match as MeSH preferred terms.

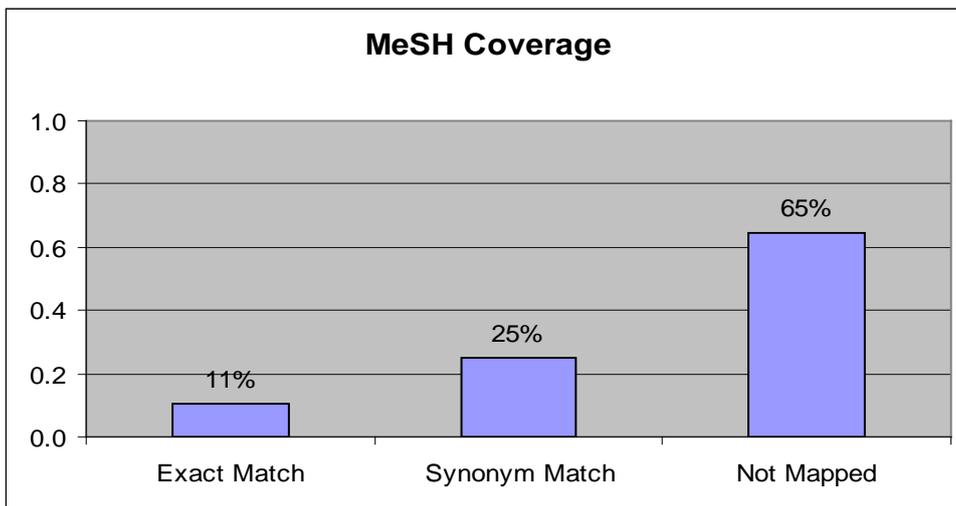


Figure 2: MeSH Coverage

Note: Percentage not adding to 100% due to rounding

### 3.3.4.3 Category Coverage

Among the 15 categories, *Educational Initiatives and Professional Development* has the highest coverage (55%). MeSH is the only vocabulary that could map the terms under this category. The terms in categories *Health Information Systems and Their Status*, *Software Design and Development*, *Technology Adoption and Evaluation* and *Telemedicine and Telehealth* couldn't map to either textbook. *Service Administration, Management and Self-Management* has the lowest MeSH coverage (14%). See figure 3.

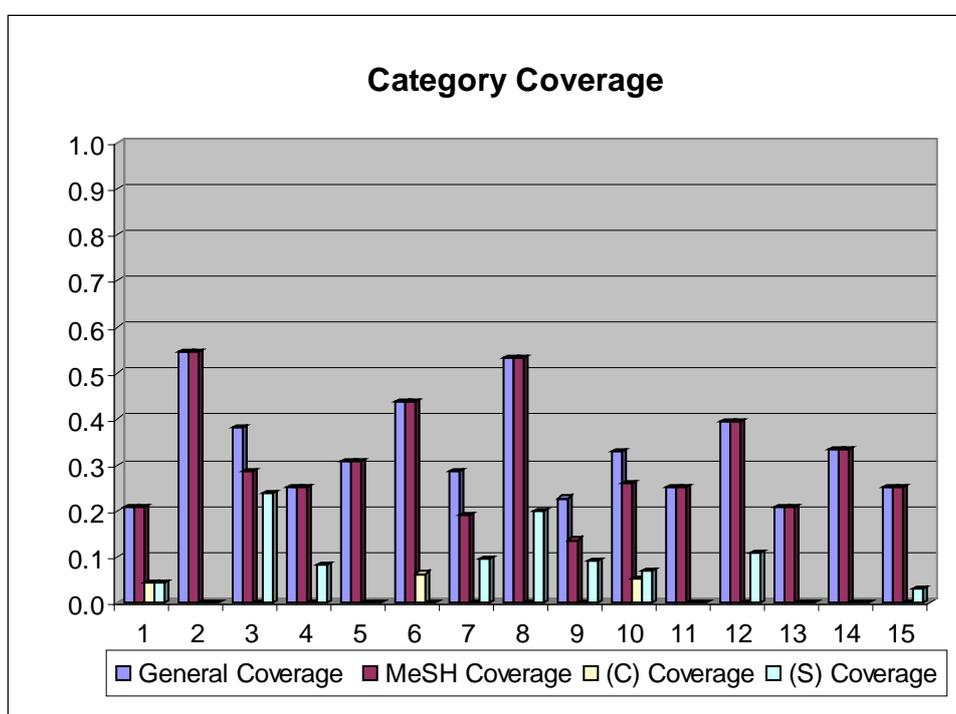


Figure 3: Category Coverage (by vocabulary sources)

Note: The numbers in X axis denote different categories listed below and the coverage includes both exact and synonym matches.

1. Decision Support, Artificial Intelligence and Modeling
2. Educational Initiatives and Professional Development
3. Electronic Health Records

4. Ethics, Policy and Government
5. Health Information Systems and Their Status
6. Nation-Wide Canadian Initiatives
7. Nursing Informatics
8. Research and Development Initiatives
9. Service Administration, Management and Self-Management
10. Software Assurance and Usability
11. Software Design and Development
12. Software Selection and Evaluation
13. Technology Adoption and Evaluation
14. Telemedicine and Telehealth
15. Terminology, Classification and Standards

### **3.3.5 Discussion**

#### **3.3.5.1 Global Coverage and MeSH Coverage**

MeSH has better results compared to the textbooks even though the coverage is still not very satisfactory (35%). Compared to MeSH, the two textbook glossaries have low coverage of the key words extracted from the latest articles in Health Informatics, especially *the Guide to Health Informatics*. The low coverage of textbook glossaries may be explained by two possible reasons: (i) The rapid evolution of Health Informatics renders the book contents not up to date as book revision is always time-consuming, especially *the Guide to Health Informatics*, which was last revised in 2003; (ii) As the key words in the ITCH 2009 Conference Proceedings were

assigned by the authors themselves, some might be less standardized, resulting in the low coverage.

MeSH has a better coverage of health informatics terms than the two textbooks. This may be attributed to MeSH's larger collection of concepts and the ability to generate MeSH preferred terms.

However, 65% non-coverage suggests that MeSH is not able to provide a perfect solution for mapping health informatics terms in the glossary. This finding is consistent with N. J. Ogg, M. Sievert's study. They found numerous major concepts not covered by current L Tree terms in MeSH after extracting and indexing medical informatics literature. Based on their study, they concluded that MeSH was not sufficient to cover the medical informatics field [18].

There are several possible explanations. First, this glossary was created based on the key terms from ITCH 2009 Conference Proceedings, which represents the latest development of Health Informatics and reflects the latest research interests in this field, yet MeSH was designed to index publications rather than to focus on research interests. Secondly, MeSH is a large vocabulary including more than 19,000 terms in health related field and is updated annually. At a certain point of time, it may not always reflect the most current research topics and techniques [17]. In addition, the revision may focus on creating a framework for standard development rather than on expanding coverage [11].

### **3.3.5.2 Category Coverage**

The result of differing coverage by category suggests that in some sub-disciplines of Health Informatics, the number of mapping concepts may be greater than in others. This can be explained by different stages of development for different sub-disciplines.

The study suggests that, among all the categories, *Electronic Health Records* is at a more mature stage than others since its coverage is consistent between MeSH and textbook vocabularies. On the other hand, *Educational Initiatives and Professional Development* is a

relatively new and evolving area since it has a high coverage in MeSH but no coverage in either of the textbooks.

As can be seen in figure 3, *Decision Support, Artificial Intelligence and Modeling; Nation-Wide Canadian Initiatives* and *Software Assurance and Usability* have the best coverage in textbook (S), which is explainable since the textbook by Shortliffe mainly focuses on computer application in health care service provision.

## 4 Conclusion

Critical appraisal is a necessary skill for students majoring in medicine, pharmacy and health informatics. It is essential for making best evidence-based decisions in providing health care service, from both clinical medicine and health informatics perspectives.

An extensive and standardized glossary is a key source for students to learn definitions and meanings of new terminologies. Consistency with the controlled vocabulary is critical in improving the effectiveness and efficiency of health informatics information retrieval.

The new glossary based on key words from ITCH 2009 Conference Proceedings cannot match the MeSH thesaurus well due to two possible reasons: (i) MeSH is not sufficient in representing the terminologies in Health Informatics; and (ii) The terms extracted from ITCH 2009 Conference Proceedings do not adequately represent Health Informatics terminology. Further research is needed to examine the reasons in more depth.

## 5 Recommendation

A good handbook should incorporate comprehensive and updated knowledge in the discipline. Updating the original course contents with the latest information is one important way to ensure that the Handbook is relevant and applicable. Some gaps are identified between the needed knowledge and current available information. CATmaker is one of them. Current CATmaker only covers medical researches in therapy, diagnosis, prognosis, and etiology. Articles in health informatics are not included. This is a limitation for health informatics students using this handbook. Further effort might focus on exploring approaches to incorporate health informatics studies into CATmaker.

In the newly developed glossary, a lot of terms in the glossary couldn't map to either current key textbooks or the MeSH vocabulary. Since the handbook is designed for students in specific disciplines to serve as an educational tool, not for lay persons, the glossary should be up to professional standards. Further testing the terms in the glossary, deleting lay terms, and using concepts preferred by controlled vocabularies will make the glossary more effective in information retrieval. Finally, mapping the glossary to other vocabulary sources such as UMLS and SNOMED and using concepts consistent with these controlled vocabularies will facilitate the application of this new glossary to Electronic Health/Medical Records when necessary.

## References

1. Alison Hill, Spittlehouse, *What is critical appraisal?* [www.evidence-based-medicine.co.uk](http://www.evidence-based-medicine.co.uk) , February 2001
2. <http://en.wikipedia.org/wiki/Glossary>
3. [http://en.wikipedia.org/wiki/Controlled\\_vocabulary](http://en.wikipedia.org/wiki/Controlled_vocabulary)
4. <http://www.healthinformatics.dal.ca/overview.php>
5. *Health Informatics Review Report*, Department of Health, 10 July 2008
6. Coiera EW (2003) *Guide to Health Informatics*, 2nd Ed, Arnold, London
7. Balas EA, Austin SM, Mitchell JA, Ewigman BG, Bopp KD, Brown GD. *The clinical value of computerized information services. A review of 98 randomized clinical trials.* Arch Fam Med 1996;5: 271-8
8. Gunther Eysenbach, Cameron Norman. *Introduction to CATCH-IT Reports: Critically Appraised Topics in Communication, Health Informatics, and Technology.* J Med Internet Res 2004; 6(4):e49
9. Zeng Q, Kogan S, Ash N, Greens R, Boxwala A. *Characteristics of consumer terminology for health information retrieval.* Methods of Information in Medicine 2002; 41(4): 289- 98
10. Li Z, Ogg N, Sievert M, and Mitchell J. *On the Growth and Trimming of the L Trees of MeSH.* Symposium on Computer Applications in Medical Care. 1993: 892
11. Rada R, Calhoun F, Mili E, Singer SJ, Blum B, and Orthnerr H. *A medical informatics thesaurus.* MEDINFO '86. Washington, DC. Amsterdam: North-Holland, 1986: 1164-1172
12. Nelson, S. J., Johnson, W. D. & Humphreys, B. L. *Relationships in medical subject headings.* In: Bean & Green (Eds). Relationships in the Organization of Knowledge. Dordrecht: Kluwer Academic Publishers, 2001, 171–84
13. Hersh, W. R., Brown, K. E., Donohoe, L. C., Campbell, E. M. & Horacek, A. E. *CliniWeb: managing clinical information on the world wide web.* Journal of the American Medical Informatics Association 1996, 3, 273–80
14. Norman, F. *Organising medical networks' information (OMNI).* Medical Informatics 1998, 23, 43–51

15. Boyer, C., Baujard, O., Baujard, V., Aurel, S., Selby, M. & Appel, R. D. *Health on the net automated database of health and medical information*. International Journal of Medical Informatics 1997, 47, 27–9
16. Deacon, P., Smith, J. B. & Tow, S. *Using metadata to create navigation paths in the HealthInsite Internet gateway*. Health Information and Libraries Journal 2001, 18, 20–9
17. Friedman PW, Winnick BL, Friedman CP, Mickelson PC, *Development of a MeSH-Based Index of Faculty Research Interests*, Proc AMIA Symp. 2000:265-9
18. N. J. Ogg, M. Sievert, Z. R. Li, and J. A. Mitchell, *Construction of a medical informatics thesaurus*. Proc Annu Symp Comput Appl Med Care. 1994; 900–904
19. Vincent Mary, Gwenaëlle Marquet, Pierre Le Beux, *MeSH and Specialized Terminologies : Coverage in the Field of Molecular Biology*, MEDINFO 2004
20. Abidi, S., Elvidge, K., Kharrazi, H., Shepherd, M., Watters, C., & Duffy, J. *Reformulation of consumer health queries with standard medical vocabulary*. 11th International Symposium on Health Information Management Research (2006)
21. Shortliffe, Edward H, *Computer Applications in Health Care and Biomedicine Series: Health Informatics*. 3rd ed., 2006, XXVI, 1037 p. 159. ISBN: 978-0-387-28986-1