

UMLS(Unified Medical Language System)의 증상용어와 국내의무기록에서 사용되는 증상용어와의 비교연구

A Comparative Study on the Sign or Symptom Concepts of the UMLS(Unified Medical Language System) and Clinical Terms in Korean Medical Records

Seung-Bin Han*, Jinwook Choi M.D., Ph.D.*, Jonghoon Chun, Ph.D†

* Department of Biomedical Engineering, College of Medicine, Seoul National University, Seoul 110-799, sbinhan@snu.ac.kr

† Division of Computer Science & Engineering, Myongji University, Seoul 120-728

Abstract: The UMLS is one of the most comprehensive medical terminology systems. We evaluated coverage of the UMLS as compared with Korean medical terms used at Seoul National University Hospital (SNUH). We tested for discordance of clinical terms between the chief complaint terms extracted from the SNUH discharge notes and those of the UMLS 'Sign or Symptom concepts. We found a degree of concordance between UMLS terms and the clinical terms used in Korean medical records. Approximately 35% of Korean clinical terms were matched to those of UMLS. The main reason for discordance was the different way in which clinical terms are described between the two vocabularies.

Keyword: UMLS, clinical terms, electronic medical terminology system, expressive pattern, concept

I. 배경

의료기관의 각종 교육, 연구, 임상자료의 데이터베이스로 인해 의료는 '정보 집중형 산업 분야'라 할 수 있다. 이는 의료의 질 향상을 위하여 진료정보를 검토하여 재활용할 필요가 높은 분야이며 의료혜택에 따른 결과의 관리를 위하여도 진료정보의 공유 및 재활용이 중요하다는 것을 의미한다. 또한 향후 전자의무기록시스템(Electronic Medical Records), 의사결정지원시스템(Clinical Decision Support System) [1], 음성인식시스템 등의 정보기술이 의료에 사용될 수 있기 위하여도 정보의 근간을 이루는 용어체계의 활용이 중요하다.

■ 의료환경에서 의학용어 사용실태

의무기록에 다양한 코드가 혼용되는 의료환경은 다음과 같은 문제점을 가지게 된다.

첫째, 같은 질병 또는 증상을 나타내는 코드인 경우에도 표현의 범위가 다를 수 있다.

둘째, 같은 질병에 대하여 각기 다른 코드로 표현하는 경우도 발생한다.

셋째, 표준 용어가 없기 때문에 의료기관 간 자료교환 및 정보공유에 걸림돌이 될 수 있다.

이러한 문제점은 의료연구 및 의료의 질 관리에 상당한 어려움을 초래할 수 있다. 향후 의료분야에 컴퓨터를 이용한 지능적인 자료처리를 위해서는 컴퓨터가 이해할 수 있는 형태의 의학용어모델을 활용하는 핵심기술에 대한 연구가 필요한데, 위와 같은 문제점들에 대한 해결방안을 제공하고 의료, 전체적으로는 생명과학분야의 전문 검색 시스템을 개발하고 지원하기 위한 연구들이 세계각국에서 진행되고 있

다. 그 중 미국국립의학도서관이 운영하는 통합용어모델(UMLS)가 가장 활발히 진행되고 있는 연구분야 중 하나이다.[2,3] 각종 시스템과 데이터베이스의 개발 및 제작 시 호환성을 가진 UMLS의 용어모델의 도입과 활용은 전자의무기록시스템의 활용이 보편화될 때 활용할 수 있는 표준화된 용어체계로서의 역할을 담당하여 정보 인프라로서 중요한 역할을 할 수 있다. 이는 각종 자료를 효율적으로 활용하기 위한 임상연구 및 지능적 의료정보검색 등에 활용될 수 있으며 이를 위하여 UMLS에는 프랑스어, 독일어, 핀란드어, 스페인어등과 같이 다국어어를 지원할 수 있도록 Word Index Table이 구성되어 있다. 향후 우리말로 기록된 의무기록 데이터를 정보검색이나 의사지원결정시스템등에서 활용하기 위해서는 한국어도 새로운 언어코드로 추가 되어야 할 필요가 있다.[4,5]

II. 방법

본 연구에서는 UMLS에 한글용어를 포함시키기 위한 기초작업으로 UMLS에 포함되어 있는 증상개념과 국내에서 주로 사용되는 증상용어의 개념일치성(concordance)을 조사·분석하였으며, 표현 차이의 주된 원인을 분석하여 국내에서 UMLS의 한글화 가능성을 제시하고 한글화 작업 시 고려해야 할 점들을 제시하는 것으로 일차연구를 진행하였다.

1. 의학용어 재료의 준비

본 연구에서는 국내에서 주로 사용하는 증상용어와 UMLS의 포괄성(coverage), UMLS의 한글화 작업 시 고려해야 할 두 용어체계 간 표현 방법의 차이를 알아보기 위한 실험으로 UMLS 2000 11th Metathesaurus와 의미망(semantic network)으로부터

134개의 semantic type을 분류하여 그 중 semantic type 'sign or symptom'에 해당되는 concepts를 사용하였다.

2. 국내 증상용어의 추출 및 정리

국내에서 주로 사용되는 증상용어는 서울대학병원의 임상 데이터웨어하우스 구축 시 환자 퇴원요약지로부터 환자주소(chief complaint)를 추출하여 정규화한 대표증상 5,073개를 활용하였으며, 한글 및 에러, 중복데이터를 제거한 3,186개를 실험을 위해 준비하였다.[6]

3. UMLS Metathesaurus로부터 증상 및 증후용어의 추출

용어시스템 서버를 2000년도 UMLS 자료를 이용하여 구축하였다. 용어시스템 서버로는 Pentium III 서버에 리눅스 운영체제에서 Mysql DBMS을 이용하여 서버기반의 용어시스템을 구축하였다. 'organism', 'body part, organ, or organ component', 'laboratory or test results', 'finding', 'medical device', 'hormone', 'lipid' 등 134개의 semantic type 중에서 'sign or sympto'에 해당하는 데이터를 MRSTY 테이블에서 찾고, 해당하는 데이터의 개념명을 MRCON에 기록되어 있는 내용 중 대표적인 용어(preferred term)를 사용하였다.

III. 결과

UMLS와 국내 증상용어의 mapping test를 수행한 결과, 서울대학병원의 의무기록 자료 중 증상용어를

3,186 개 중 'sign or symptom' 개념과 의미적으로 일치하는 데이터의 종류는 1,100개인 35%였으며, 나머지 58%에 해당하는 데이터는 증상보다는 검사, 수술, 진단명 등이었으므로 증상용어 매핑을 위한 UMLS 'sign or symptom' 개념과 일치되지 않았다. 그 외 나머지 7%는 국내 의무기록에서 사용하는 용어와 UMLS에서 포함하고 있는 용어의 개념이 달리 쓰이거나 UMLS에는 없는 용어를 사용하는 경우였다.

국내 증상용어와 UMLS 간의 주된 표현양상의 차이를 분석한 결과는 다음과 같다.

첫째, 국내 의무기록에서 사용하는 증상용어는 대부분 단순히 발병부위의 해부학적 위치에 의존적(usually dependant on the anatomical region)이며 '증상, 부위'에 근거한 표현 체계인 반면, UMLS에 포함되어 있는 용어는 포괄적이거나 단순한 용어의 범위 뿐 아니라 증상의 성상, 원인, 진행정도와 심한 정도(severity)까지 자세하고 다양한 표현을 포함한다.

둘째, 증상과 또는 증후군을 잘 확실히 분리된 개념으로 사용하지 않고, 질병명을 흔히 주소로 기록하며, 증상을 대신하는 용어로 사용하는 표현기준의 모호함을 발견할 수 있다.

우리가 흔히 증상 또는 증후라고 사용하는 의학용어의 개념들을 'sign or symptom' 인지 또는 질병으로서의 'disease or syndrome' 인지를 명확히 구분하지 않기 때문에, 국내 증상용어 데이터에서 발견할

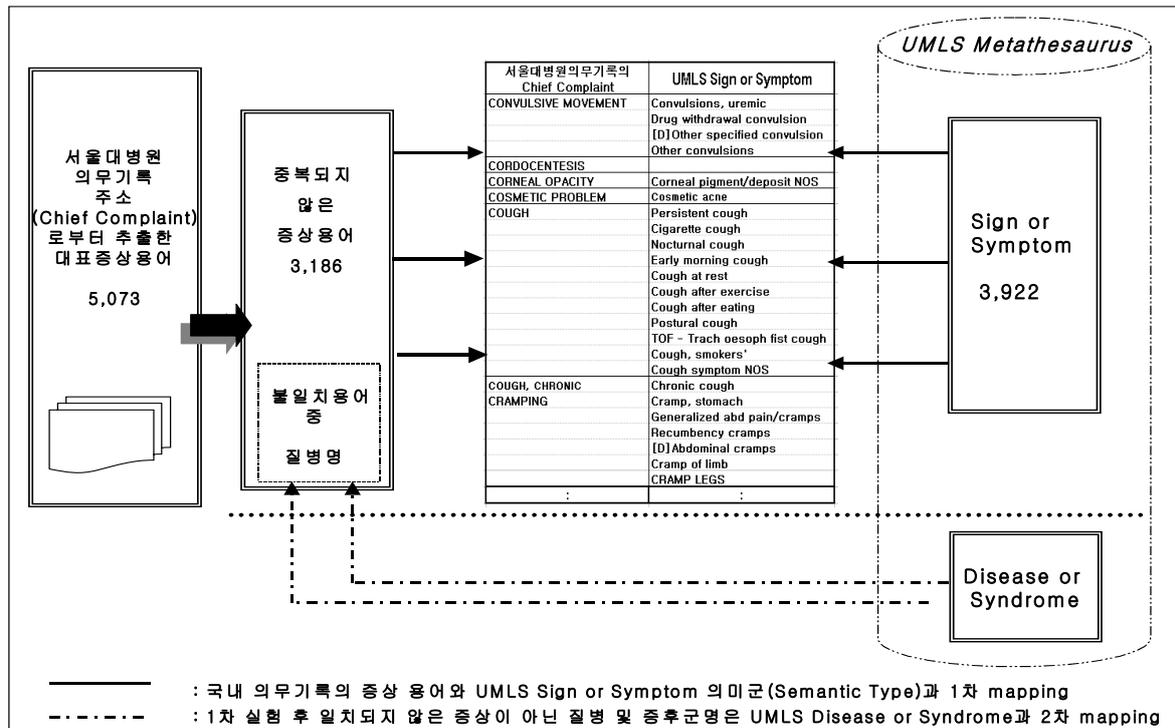


Figure 1. Methodology and Process of mapping test
 인기 위해 추출한 중복 없는 주소(chief complaint) 수 있는 'buccal abscess', 'liver abscess',

'paradental abscess', 'gallbladder abscess'와 같이 특정부위 또는 조직에 발생한 농양이라는 질병에 대하여 증상과 개념적으로 혼용하여 사용하고 있음을 발견할 수 있었다. 따라서 이와 같은 표현들은 UMLS 'sign or symptom' 개념용어와 거의 일치하지 않았다.

Table 1. Comparatives the expressive patterns of the clinical records in SNUH and UMLS sign or Symptom

서울대병원 의무기록의 Chief Complaint	UMLS Sign or Symptom
BACKACHE	Backache with radiation Low Back Pain, Mechanical Low Back Pain, Posterior Compartment Acute back pain with sciatica Back pain worse on sneezing Post-surgery back pain :
CHEST PAIN	[D]Retrosternal chest pain CHEST PAIN, DULL CHEST PAIN, CRUSHING Pleuropericardial chest pain Esophageal chest pain CHEST BURNING PAIN OF :
TREMOR	[D]Tremor NOS TREMOR, HAND Persistent Tremor TREMOR, HEAD Continuous Tremor TREMOR, TRUNCAL Intermittent Tremor Fine Tremor Postural tremor Rest tremor :
BLINDNESS	Blindness, Complete Blindness, Monocular Blindness, Cortical, Post-Ictal Blindness Blindness, Legal :

우리가 흔히 증상용어로 간주하여 사용하는 주된 용어 중 'deformity', 'skin defect', 'gallbladder mass', 'trauma', 'hernia', 'acromegaly', 'hyperglycemia', 'adenomiosis', 'myocardial infarction', 'amenorrhea', 'azotemia', 'blepharochalasia', 'cytopenia', 'endometriosis' 등 국내 의무기록에 흔히증상이라고 간주되어 사용되는 표현 들은 증상이 아닌 질병에 속하는 용어들이기 때문에 두 용어체계 간 mapping test에서 일치성을 많이 떨어지게 한 주된 원인이었다. 이러한 해부학적 위치를 포함한 용어들은 그 세부적 위치와 질병의 자세한 묘사와 함께 UMLS semantic type 'disease or syndrome' 의미군에 속하는 개념용어로 대부분 연결될 수 있었다. 수 있었다.

셋째, 용어를 사용하는 대상범위와 개념이 두 용어체계 간에 달리 적용되는 예를 발견할 수 있었다.

UMLS에서는 약해짐(weakness)라는 개념을 주로 근육에 사용하며 사지의 아래쪽, 위쪽의 부위뿐 아니라 그 증상의 성상, 원인을 함께 나타내 주는 것을 볼 수 있다. 또한 본 연구에서 국내증상용어와 mapping시키기 위해 대상으로 한 UMLS Semantic type 'sign or symptom' 뿐 아니라 semantic type

Table 2. Example the expressive patterns of the clinical records in SNUH and UMLS sign or Symptom - weakness

서울대병원 의무기록의 Chief Complaint	UMLS Sign or Symptom
ABDOMINAL WEAKNESS	Lower Extremity Weakness, Spastic
WEAKNESS, HAND	Muscle Pain or Weakness
WEAKNESS, HAND, FOOT	Muscle weakness: flexion and adduction of hip
WEAKNESS, LIMB	Muscle weakness: adduction of hip, extension of knee
WEAKNESS, LOWER EXTREMITY	WRIST MUSCLE WEAKNESS
WEAKNESS, UPPER EXTREMITY	EXTREMITY MUSCLE WEAKNESS, LOWER
GENERAL WEAKNESS	Diaphragmatic weakness General weakness/tiredness Paralysis/weakness WEAKNESS POSTURAL Upper facial weakness Weakness of limb :

'disease or syndrome'에서 'weakness' concept를 찾아 본 결과 국내 증상용어 중 'abdominal weakness'처럼 'abdominal'에는 'weakness'라는 개념을 적용시키지 않는 것 같다.

또한 국내 의무기록에서는 주로 '출혈, 부위', '수술 또는 분만 후 출혈'로 'bleeding'을 사용했으나, UMLS 'sign or symptom'에서는 'bleeding'의 개념을 주로 menstruation과 sexual intercourse에 사용하였으며, 우리가 주로 사용하는 'chief complaint'에서와 같은, 코피, 잇몸의 출혈, 분만 후 출혈 등은 UMLS 'disease or syndrome'의 category에서 찾을 수 있다. 그러므로 용어에 따라서 적용시키는 개념과 범위에 차이가 있는 것으로 생각된다.

넷째, 국내 의무기록에서 흔히 볼 수 있는 증상용어 표현 가운데 UMLS에 없는 개념들이 있다.

국내 증상용어 data set에서 'absence, anus', 'absence, breast'와 같이 의사가 환자 진료 시 특이적 신체 상태, 부위를 명시하는 표현을 볼 수 있다. 그러나 UMLS 'sign or symptom' concepts에는 위와 같은 '신체의 어떤 부분이 없다'는 "absence, **"의 개념이 없고 오히려 'absence of menstruation', 'absence of reflex', 'absence of tickle sensation', 'absence of pain', 'absence of pain sensation', 'absence of nausea and vomiting' 등과 같이 생리적인 신체현상 또는 병적인 상태가 없어졌음을 나타내는 개념으로 'absence' 라는 개념을 사용한다.

화상의 흉터를 명시하는데 있어서도 UMLS semantic type 'sign or symptom'에서는 국내 의무기록에서 자주 사용하는 'burn scar'는 찾을 수 없고 오직 'burning sensation' 즉 쓰라린 듯이 아픈 느낌을 표

현하기 위하여 'burn'을 사용하며 우리와 같은 화상에 대한 흉터의 의미를 찾기 위해서는 semantic type 'disease or syndrome'에서 각각 'burn', 'scar', 'cicatrix'의 분리된 용어로 찾아야 한다.

IV. 결론

본 연구에서는 국내에 UMLS의 적용 시 통합의학용어체계로서의 적절성과 활용가능성 및 향후 한글화 작업 시 고려해야 할 두 용어체계 사이의 차이점을 알아보기 위한 일차적인 연구로 국내에서 주로 사용되는 증상용어와 UMLS 'sign or symptom' 간 매핑을 통하여 두 용어체계 간 의미론적 상호호환성과 두 용어체계 간의 주된 표현방법의 차이를 살펴보았다.

첫째, 국내의무기록에서는 '증상', '증후', '증후군', '질병', '진단'에 대하여 명확한 개념을 정의하고 사용해야 할 필요가 있다.

국내의 의료진들이 환자 주소(chief complaint)를 기록할 때 환자가 말하는 것을 기록하기 보다는 진단명을 주된 증상으로 사용하는 문제가 있다. 실험을 위한 자료가 3차 의료기관인 서울대학교 병원의 데이터이기 때문에 환자는 병원에 내원했을 때는 이미 자신의 병명을 알고 있거나, 더 진보된 치료를 받기 위해서 내원한 것이므로 의료진은 환자의 주된 증상을 적기보다는 의심이 되거나 이미 진단된 질병명을 사용하는 경우가 많기 때문에 '증상 및 증후'란에 순수한 증상용어가 아닌 'known trophoblastic gestational tumor, for op spinal stenosis, evaluation old stroke, alleged polyp, cervix, DM management preop.' 등 진단명, 수술명 또는 처치명 등을 기록하기 때문에 순수한 증상용어와 혼용하여 사용하는 문제점을 발견할 수 있다.

둘째, 국내 의료분야에서 주로 사용하는 증상용어의 표현의 정확성이나 다양성이 많이 부족한 점을 지적할 수 있다.

실험결과 UMLS의 증상 및 징후 용어들과 일치되지 않았던 이유 중 대부분은 국내의무기록의 증상 표현이 대부분 질병이 발생하는 부위에 기초한 기록이었다. 그에 비해, UMLS는 하나의 개념 용어와 관련하여서도 환자의 상태와 증상 및 질병의 성상, 진행정도, 심리적인 상태, 병의 원인, 행정적인 문제에 관한 용어까지, 포괄적이면서도 다양하고 자세한 표현들을 살펴볼 수 있었다. 따라서 자세한 환자 상태의 기록을 위하여 임상적으로 사용하게 되는 증상용어의 표현을 다양하게 개발할 필요가 있으며, 향후 이와 같은 용어의 개발이 UMLS 또는 SNOMED[7]의 용어들을 참조하여 개발하는 것이 필요할 것으로 생각된다.

전자의무기록, 임상 의사결정 시스템 등을 의료환경에서 사용하기 위하여서는 먼저 의학정보 지식의 중추

가 되는 의학용어 서버의 구축이 진행되어야 한다. 이러한 용어체계 시스템을 기반으로 여러 가지 검사 정보, 진료데이터, 보조 시스템, 의사결정지원시스템 등의 병원정보 시스템이 통합될 수 있으며, 효율적인 전자의무기록 구축의 성공여부를 결정하는 주요한 결정인자가 된다고 말할 수 있다. 그러므로, 각종 시스템과 데이터베이스의 개발 및 제작 시 호환성을 제공하는 통합 의학용어체계인 UMLS를 이용하여, 국내 의료환경에서 사용할 수 있는 진단, 증상용어가 한글화된 용어시스템을 개발 하는 것이 많은 의료용어시스템을 개발하는데 선결되어야 하는 주요한 과제라 할 수 있다.

본 연구를 통하여 통합의학용어체계로서 UMLS가 한국의 용어모델을 담게 될 때 발생할 수 있는 용어 표현상의 차이점을 정리하였고, 현재 사용중인 용어 표현의 문제점들을 도출할 수 있었다. 또한 향후 의학분야의 정보인프라를 위하여 용어모델은 가장 중요한 위치에 있음을 다시 한 번 확인할 수 있는 계기가 되었고, 국내 의학용어체계개발을 위하여는 체계적인 접근과 보다 많은 연구자들이 관심을 기울여야만 국내 의학용어체계가 제 자리를 찾을 수 있음을 다시 한번 확인하였다.

참고문헌

1. Soumeya L, A UMLS-based Knowledge Acquisition Tool for Rule-based Clinical Decision Support System Development, JAMIA Vol 8, No 4, Jul/Aug 2001, 351-360
2. Bodenreider O., Evaluation of the UMLS as a medical Knowledge Source, JAMIA. 1998: 5:76-87
3. Betsy L. Humphreys, The Unifies Medical Language System, JAMIA. 1998;5:1-11
4. Unified Medical Language System, U.S. National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894, National Institutes of Health <http://www.nlm.nih.gov/research/UMLS/>
5. McCray AT, The UMLS knowledge source server: A versatile internet-based research tool, JAMIA 1996:164-168
6. Choi J. Analysis of Korean medical records using Datawarehouse. Journal of Korean Society of Medical informatics 2001; 7:1-11
7. The Systematized Nomenclature of Medicine(SNOMED), College of American Pathologists (CAP), <http://www.snomed.org/>