

한의학 용어의 UMLS 등재 - KIOM 용어정제연구 중 經穴名을 중심으로 - Inclusion of the Traditional Korean Medical Terms into the UMLS

김진현, 김상균, 장현철, 김민아, 오용택, 배순희, 김창석,
전병욱, 김재훈, 송미영
한국한의학연구원 정보연구센터

Jin-hyun Kim, Sang-kyun Kim, Hyunchul Jang, Minah
Kim, Oh Yong-taek, Sunhee Bae, Changseok Kim,
Byounguk Jeon, Jae-Hun Kim, Mi-young Song
Information Research Center, Korea Institute of
Oriental Medicine

요약

본 연구의 목적은 미국국립의학도서관(NLM)의 Unified Medical Language System(UMLS)에 한의학의 용어를 등재하는 것이다. 1차 등재 대상으로 한국한의학연구원(KIOM)에서 진행중인 한의학 용어 정제 연구 결과물 중 任督脈과 12正經에 속하는 360개의 經穴名의 Concept Unique Identify(CUI)를 선정하였다. UMLS Knowledge Source Server (UMLSKS)의 메타시소러스 검색을 통해 UMLS 내 經穴용어에 대한 terminology정보를 수집, 분석하였다. 이를 바탕으로 용어간 개념 비교를 통해 UMLS의 經穴용어와 KIOM의 經穴용어를 매핑하였다. 마지막으로 Rich Release Format(RRF)로 데이터를 저장하고, Unicode Transformation Format-8 (UTF-8)로 인코딩하여 NLM으로 송부하였다. NLM에서 용어 등재에 대한 적합성 여부를 판별한 후 2010AB버전에 "TKMT2010"이라는 소스명으로 한글로 된 한의학 용어가 최초로 UMLS에 등재되었다. 향후 UMLS와 같은 국제표준의학용어와 연계를 통해 한의학 용어가 다양한 의학용어체계와 상호호환성을 확보하고 표준화, 세계화할 수 있도록 지속적인 연구가 필요하다.

I. 서론

UMLS는 전세계 의료분야의 정보공유 및 교환을 위해 개발된 포괄적인 의학용어체계로서, 각각 상이한 목적을 위해 구축되었던 ICD-10, SNOMED, MeSH, RxNorm, NCBI 등 156개 medical vocabulary(2010AA 현재)를 통합하고 이들 간의 상호호환성과 운용성을 확보하기 위한 연구를 지속적으로 진행하고 있다[1].

서양의학 용어와 더불어 전통의학 용어 역시 의학용어의 한 영역으로서 UMLS내에 구축되어 있지만, 서양의학 용어 중심의 체계에서 동서양 의학의 개념적 차이에 의한 한계로 인해 그 수량이 상대적으로 적고 수록범위도 극히 제한적이다. 이러한 어려운 환경에도 불구하고 국내 한의계에서는 한의학 용어를 국제 의료용어체계와 연계하고자 하는 연구가 지속적으로 시도되고 있는데, 대표적인 연구로는 「脾胃論」에 기재된 용어들을 분류하고 UMLS의 Semantic Type과 비교분석한 연구, SNOMED를 사용하여 표준화된 한의학 임상용어를 제안하고 이를 입력할 수 있는 전자진료차트 모듈개발 연구, UMLS 기반의 한의학 참조용어를 구축한 보건의료정보 표준화 연구의 한방참조용어 개발과제 등이 있다.

이러한 노력과 더불어 KIOM의 정보연구센터에서는 2009년부터 5년간 [온톨로지 기반 한의학 지능형 정보체

계 연구를 진행중이며, 그 세부과제로 한의학 용어 정제 연구가 수행되고 있다. 2009년에서 2011년까지 3년에 걸쳐 총 30,000개의 대표 개념어를 수집, 정제하고 이들 용어를 효율적으로 통제, 활용할 수 있는 UMLS 기반 관리체계를 구축하는 것을 주된 연구내용을 한다. 더불어 구축된 용어의 일부를 2010년부터 연차별 단계적으로 UMLS에 등재하는 프로젝트를 추진하고 있으며, UMLS 내 전통의학 용어에 대한 선행 분석[1]을 바탕으로 2010년 1차로 등재된 經穴名의 등재 과정을 소개하고자 한다.

II. 본론

1. UMLS 내 經穴용어

UMLS 내 한의학 용어를 제안하기 전, 수집할 한의학 용어의 범위를 정하고 UMLS내 수록형태에 대한 조사 작업이 선행되었다. 우선 한의학의 가장 대표적인 분야인 經穴名을 1차로 선정하였고, 「WHO/WPRO 표준경혈 위치」[2]가 발간되어 명명과 위치의 표준이 정해진 任督脈과 12正經의 360개 經穴名을 대상으로 연구를 진행하였다.

UMLS내 經穴名의 수록형태는 모두 약호(abbreviation code)로 되어 있는데, 국내 한의계에서 약호보다 일반적

으로 통용되는 한자명, 한글명, 중국병음, 한국어 독음, 위치 등의 정보는 거의 수록되어 있지 않았다. 이들의 주된 vocabulary는 Read Codes와 SNOMED CT이며, semantic type은 모두 'Body Location or Region'로 분류되었다. CUI의 형태는 LU2(C0450466)와 같이 약호를 기본으로 하되, Lung acupuncture point LU1(C0450465)처럼 경락이 함께 기재되거나, ST14 (body structure) (C0450511)처럼 SNOMED CT의 분류명이 함께 기재되기도 하였다. 정經穴 중 GB11(頭竅陰), BL66(足通谷), BL67(至陰)의 經穴名은 UMLS내에 존재하지 않았으며, Liver acupuncture point LP3(C0450768)과 같이 오타가 CUI로 등록된 경우도 있었다 [1].

2. KIOM 용어 중 經穴名의 UMLS 등재

KIOM에서는 2009년부터 5년간 [온톨로지 기반 한의 학 지능형 정보체계 연구]를 진행중이며, 그 세부과제로 한의학 용어 정제연구를 수행하고 있다. 용어를 효율적으로 관리할 수 있는 용어관리시스템을 개발하여 2009년에서 2011년까지 3년에 걸쳐 총 30,000개의 대표 개념어를 수집, 정제하고 개념 중심의 UMLS 기반 관리체계를 구축하는 것을 주된 연구내용을 한다. 이와 더불어 한의학 용어 정제연구 결과물을 연차별 단계적으로 UMLS에 등재하는 프로젝트를 진행 중이다. 이 중 1차로 經穴名을 UMLS 2010AB release에 제안하였다. UMLS 내 누락된 GB11(頭竅陰), BL66(足通谷), BL67(至陰)과 오타로 수록된 Liver acupuncture point LP3(C0450768)를 제외한 14經穴 356개 經穴名의 한국어 번역본에 해당하는 KIOM 經穴名을 NLM에서 요구하는 UTF-8로 인코딩된 Rich Release Format(RRF)파일 형식에 맞게 구조화하였다. RRF는 NLM과 협력업체인 Apelon에서 UMLS Metathesaurus를 위해 개발한 고유한 파일 형식으로, 기존의 Original Release Format(ORF)에 비해 자료원의 투명성을 확보할 수 있다는 장점이 있어, 각각의 vocabulary 내에서 관계와 용어의 부가 정보들 사이에 대한 구분을 명확히 유지하고자 2004년 이후 RRF 타입으로 자료를 구조화하고 있다[3]. 더불어 Character-set을 UTF-8로 인코딩하는 것은 이 방식이 전세계 모든 문자를 동시에 표현가능하면서 상이한 언어 체계간 충돌을 방지하며, 다국어 입력에 수월하다는 장점이 있기 때문이다. 또한 용어와 함께 vocabulary의 소유권, 용어의 구조, 사용상 restriction의 수준, 업데이트 주기 등에 대한 제반 정보를 작성하여 제출하였으며, 용어 등재의 적합성 여부를 NLM에서 검토하였고, 2010AB release에 "TKMT2010"이라는 소스명으로 한글로 된 한의학 용어가 UMLS에 등재되었다.

The screenshot shows the UMLS Terminology Services interface. At the top, it says 'A service of the U.S. National Library of Medicine | National Institutes of Health'. Below that is the 'Unified Medical Language System' logo and 'UMLS Terminology Services'. The main content area is titled 'UMLS Metathesaurus License' and includes a scroll-down instruction: 'Scroll down and click "Accept and Continue" on each license page. Requesting a license and activating your account is a two-step process.' There are links for 'License Agreement for the Use of the UMLS Metathesaurus', 'Appendix 1', and 'Appendix 2'. The specific license is identified as 'TKMT2010 (new) - Traditional Korean Medical Terms, 2010; Korea; KOR'. Contact information for Jinhyun Kim is provided at the bottom.

▶▶ 그림 1. UMLS Terminology Service 내 TKMT2010의 수록정보

III. 결론

본 연구는 향후 연차별 단계적으로 진행될 한의학 용어 UMLS 등재의 1차 결과물로서, 經穴용어에 대한 UMLS 내 데이터를 조사하고 이를 KIOM 용어정제연구 결과물과 연계하여 한글로 된 한의학 용어를 등재했다는 점에서 큰 의의가 있다.

표준화된 한의학 용어를 통해 개념을 정의하고 이를 국제표준 의료용어 체계와 연계하는 연구는 학문의 발전과 한의학의 세계화에 초석이 되는 중요한 연구이다. 향후 한의학 분야의 효율적인 정보관리와 학문의 세계화를 위해 필수적인 한의학 표준용어모형을 정립하고 UMLS에 등재된 다양한 국제표준용어와 연계하여 상호호환성을 확보하기 위한 지속적이고 체계적인 연구가 필요할 것이다.

■ 참고 문헌 ■

- [1] 김진현, 김철, 장현철, 전병욱, 이상준, 김상균, 송미영, "UMLS내 전통의학 용어에 대한 고찰", 대한한의학원전학회지, 제23권, 제6호, pp15-26, 2010
- [2] World Health Organization 2008, WHO Standard Acupuncture Point Location in the Western Pacific Region, pp25-231, 2005.
- [3] 한승빈, 김성희, 최진욱, "UMLS Metathesaurus 2004의 새로운 파일구조 - Rich Release Format(RRF)의 소개, 2004 대한의료정보학회 추계학술대회 자료집, 2004.