

임상용어의 효율적 적용을 위한 참조용어 Subset 에디터의 개발

조훈[†], 김형희^{**}, 최병관^{***}, 최영연^{****}, 김화선^{*****}, 홍해숙^{*****}

요 약

하나의 의학용어체계가 모든 의학 개념을 표현할 수 없으므로 전자의무기록의 각 영역에 적합한 의학용어를 적용하고 용어들을 효과적으로 연계하는 것이 실제 임상에서 매우 유용한 방법이다. 표준용어의 편리하고 효율적인 활용을 위해서는 임상과나 임상과의 사용목적에 적합한 범주를 구분하고, 구분한 범주에 입각하여 사용가능성 높은 용어들을 추출하여 정리한 용어 Subset의 개발이 요구된다. 또한 Subset은 표준용어체계의 변경이나 수정, 사용자의 새로운 요구에 알맞도록 지속적으로 개발과 업그레이드가 가능해야 한다. 본 연구에서는 부산대학교병원 신경외과환자의 퇴원요약지를 대상으로 추출된 주호소, 증상, 진단 및 수술명, 과거 시술명의 추출된 자료를 추출하고 분석하였다. 그리고 subset 데이터베이스를 생성하고 추가되는 용어에 대해서는 임상의가 직접 참조용어와의 매핑을 통해 새로운 데이터베이스를 생성될 수 있도록 Subset 에디터를 개발하였다. 결론적으로 본 연구는 용어체계의 방대함으로 인해 생기는 문제점과 비효율성을 줄일 수 있는 효과적인 관리 방법이 될 것이다.

A Development of Reference Terminology Subset Editor for effective adaption of Clinical Vocabulary

Hune Cho[†], Hyung Hoi Kim^{**}, Byung Guan Choi^{***}, Young Yeon Choi^{****},
Hwa Sun Kim^{*****}, Hae Sook Hong^{*****}

ABSTRACT

It is highly useful in an actual clinical setting to apply appropriate medical terms to every area of electronic medical record (EMR) and link them effectively, as a single medical terminology system cannot cover all medical concepts. In order to use standardized terms conveniently and efficiently, it is required to categorize them depending on the purpose of individual departments or physicians and thereby develop organized subsets of extracted terms highly likely to be used. In addition, it is important to such a subset to make it possible to change or correct standardized terminology system and continue to develop and upgrade to meet renewed demands of users. In this paper, data including chief complaint, symptoms, diagnosis, operation, and history of previous treatments were collected from discharge summary of patients with Department of Neurosurgery at Busan National University Hospital for analysis. In addition, subset database was created, and for terms needed to be added, the physician directly performed mapping through connection with reference terminology server and developed subset editor for the purpose of creating new subset database. Therefore, it is expected that this can serve as a practical and effective management method to reduce problems and inefficiency caused by existing vast terminology system.

Key words: Electronic medical record(전자의무기록), Department of Neurosurgery(신경외과), Reference Terminology(참조용어), Mapping(매핑), Subset Editor(서브셋 에디터)

※ 교신저자(Corresponding Author) : 김화선, 주소 : 대구 중구 동인동 2-101(700-422), 전화 : 053)420-4896, FAX : 053)-423-1242, E-mail : pulala@paran.com

접수일 : 2007년 9월 4일, 완료일 : 2008년 1월 9일

[†] 경북대학교 의과대학 의료정보학과 교수
(E-mail : hunecho@knu.ac.kr)

^{**} 부산대학교 의과대학 진단검사의학과 교수
(E-mail : hhkim@pusan.ac.kr)

^{***} 부산대학교 의과대학 신경외과 교수

(E-mail : ns@doctor.com)

^{****} 경북대학교 의학전문대학원 의학과 3년
(E-mail : ychoi96@hanmail.net)

^{*****} 경북대학교 의과대학 의료정보학과 Post-Doc
^{*****} 경북대학교 간호대학 간호학과 교수

(E-mail : hshong@knu.ac.kr)

※ 본 논문은 2007년도 경북대학교 2007년 Post-Doc. 연수 지원에 의하여 연구되었음

1. 서 론

대규모 의료기관들의 최근 전자의무기록의 추세는 최첨단화된 의료지원시스템에 초점이 맞추어져 있다. 그 동안 종이기반의 의무기록관리시스템의 문제점은 첫째, 의무기록지 공유의 불가능, 둘째, 의무기록지의 분실이나 누락 및 보관 장소의 확보에 따른 관리상의 문제 셋째, 진료정보의 통합데이터베이스가 제대로 구축되지 못함에 따른 정보교환의 미비 및 의학 연구의 불편함을 그 주요한 문제점으로 내포하고 있다[1]. 전자의무기록은 이러한 문제점 해소를 통해 의료정보의 접근성을 향상시키고, 기록의 가속성과 완전성을 높이며, 정보의 통합성을 높여 사용자들이 쉽게 정보를 조회할 수 있게 한다. 또한 현장에서 오류를 감소시키고, 의사결정시스템을 효과적으로 적용할 수 있게 하며 가이드라인 혹은 주 임상경로를 통한 의료의 질 향상과 표준화를 이루게 한다.

전자의무기록은 의무기록을 단순한 기록으로 남겨두는 것이 아니라 재사용을 할 수 있는 정보로서 활용하는 측면에서 매우 중요한데, 이것은 컴퓨터에서의 인식문제와 더불어 전자의무기록 시스템에서 필수적인 조건이 된다. 다시 말하면 전자의무기록은 컴퓨터에서 인식 가능한 정보형태로 저장 관리되어야 하며 이것을 통해 기록의 재사용이 이루어져야 한다[2]. 하지만 전자의무기록의 효과는 단순히 시스템의 적용만으로는 이루어 질 수 없다. 잘 구성된 시스템 통합이 필요하며, 무엇보다도 구조화된 정보와 표준화된 의학용어가 적용된 전자의무기록을 필요로 한다. 현재 의학적 개념을 표현하기 위해 다양한 용어가 연구되고 있으나 모든 개념을 포괄할 수 있는 의학용어의 개발에는 많은 장애가 있다. 특히, 실제 진료환경에서 환자가 표현하는 용어를 정의된 표현 규칙으로 구조화할 수 있는 방법은 더욱 어려운 실정이다. 하나의 의학용어체계가 모든 의학 개념을 표현할 수 없으므로 전자의무기록의 모든 영역을 하나의 의학용어로서 표준화할 수는 없다. 임상에서 검색하여 사용하기에는 방대할 뿐만 아니라 표준용어에 대한 사용자들의 이해를 요구하므로 실질적으로 이용하는 데는 어려운 점이 있다[3].

대부분의 의학용어체계는 실제 임상에서 직접적으로 적용하는 것을 목적으로 구성된 것이 아니라, 참조용어체계(reference terminology system) 혹은

용어 간 개념 관계를 정의하여 용어 간의 통합을 목적으로 개발되어 있다. 그러므로 표준용어의 편리하고 효율적인 활용을 위해서는 임상과나 임상과의 사용목적에 적합한 범주를 구분하고 구분한 범주에 입각하여 사용가능성 높은 용어들을 추출하여 정리한 용어의 서브셋(Subsubset)의 개발이 요구된다. 또한 이 Subsubset은 표준용어체계의 변경이나 수정, 사용자의 새로운 요구에 알맞도록 지속적으로 개발과 업그레이드가 가능하여야만 한다[4].

전자의무기록을 위한 핵심기반기술연구로서 영국, 호주, 미국, 일본 등에서는 이러한 연구가 진행 중에 있으며(표 1) 2008년을 목표로 참조용어인 SNOMED CT(systematized nomenclature of human and veterinary medicine - clinical terms)를 중심으로 용어에 대한 패턴분석과 분석된 내용을 기반으로 자국의 실정에 맞는 용어모델을 개발 중에 있다. 네덜란드의 Open SDE(structured data entry)의 경우는 2005년 완성되어 2006년에 부분적으로 임상에서 적용하였고 그 결과들을 보고하고 있다[5-7].

미국의 경우 2004년 4월 27일 부시 대통령이 10년 이내에 상호운용적인 평생전자의무기록을 널리 적용할 것을 선포한 후 연방정부는 의료정보의 상호운용성(interoperability)을 이루기 위한 표준의 적용과 지원을 포함한 의료정보기술의 사용과 개발에 앞장서왔다. 그래서 전 국민의 평생전자의무기록을 위한 시스템 개발을 위해 국가차원에서 일괄적으로 SNOMED CT라이선스를 구매하여 점진적으로 임상에 적용 및 평가하고 있다.

2006년 미국의료정보협회(AMIA: American Medical Informatics Association)주관의 학회에서 미국방성 및 재향군인병원에서 사용 결과에 대한 발표에 따르면 SNOMECE CT의 민감도(sensitivity)는 63.8%, 최고 92.4%까지 각 임상 영역을 표현할 수 있다고 발표하였다[8]. 또한 국제적으로는 ISO TC 215에서 세계의 건강관리 산업에서 사용되는 전체적인 개념표현을 위한 모델화 작업을 통해 표준화하려는 노력을 하고 있다.

2. 연구 재료

표준용어의 사용은 병원정보시스템에 많은 이익을 제공하고 효율을 높여주지만, 다양한 임상분야에

표 1. 국외 참조용어에 대한 연구 사례

Project	Reference
Development of terminology set based on reference terminology	UK NHS SNOMED CT subset development project
Development of interface terminology	AU GP Vocabulary Project - Stage 2, Diabetic Terminology development
	Japan Medis-DC
Development of interface terminology and structured data entry	USA MEDCIN
	Netherlands OpenSDE
Definition of common constraints for specified clinical field	ISO TC 215 Conceptual framework for clinical findings
Reference Terminology &HL7 RIM	HL7 Terminfo

적용할 수 있도록 생성된 참조용어의 경우 그 범위와 코드의 양이 방대하여 직접적으로 의료현장에 적용 시키기는 어려우므로 임상과별로 Subset을 구성하여 사용하는 것이 바람직하다. 실제 참조용어를 병원에서 활용하기 위해서는 용어의 범주에 따라서 추출, 정리, 분석하는 작업을 위해 많은 시간과 인력을 필요로 할 뿐만 아니라, 추출된 용어들을 표준용어체계와 하나씩 매핑하는 작업 역시 고돌 전문화된 의료인력과 시간을 필요로 한다. 즉 많은 시간과 인력을 이용해 필요한 용어를 정리하지만 정리해 놓은 용어세트 역시 크기가 방대하여 실제 적용하기가 용이하지 않고 변화에 둔감한 시스템이 된다. 따라서 본 연구에서는 부산대학교병원 신경외과 환자의 1년 간의 퇴원요약지를 대상으로 추출된 주진단 및 수술명, 과거 시술명의 추출된 자료를 정리, 분석한 후 Subset을 개발하고 차후 추가되는 용어에 대해서는 임상의가 직접 Subset 에디터를 활용해서 참조용어와의 매핑을 통해 표준화된 용어체계를 생성할 수 있도록 하였다.

2.1 SNOMED CT (Systematized Nomenclature of Human and Veterinary Medicine - Clinical Terms)

본 연구에서는 참조용어로서 SNOMED CT를 선택하였다. 그 이유는 1963년부터 미국 병리학자들에 의해 체계화 된 용어체계인 SNOMED 온톨로지 개념을 도입하여 SNOMED CT로 발전하기에 이르렀다. SNOMED CT는 1999년 미국병리학회(college of American pathologists)와 영국 National Health

Service(NHS)가 SNOMED Reference Terminology (SNOMED RT)와 Clinical Terms Version 3 (CTV3)을 통합해서 만든 국제표준 의학용어체계로서, 의학 개념을 표현할 수 있도록 하는 정확하고 포괄적인 임상참조용어(clinical reference terminology)이다[9,10]. SNOMED는 질병의 여러 가지 특성을 코드화하도록 설계되었으며, 용어를 합쳐서 좀 더 복잡한 개체나 개념을 만들 수 있는 특성을 가지고 있다. 이것은 하나의 개념에 대하여 여러 가지의 조합방식들이 존재할 수 있다는 것을 의미하며, 개발자와 사용자에게 확장성을 제공한다는 측면에서 장점으로 부각시킬 수 있다. SNOMED CT는 개념(concepts), 유사어(descriptions), 관계(relationships)를 포함하고 있는 핵심 테이블(core tables)과 ICD-10 (international classification of diseases of tenth revision)과 같은 분류체계와 연계 테이블(cross-mapping Mechanism), 특정 영역 혹은 기관에 적합하게 용어의 개념을 구분 지을 수 있도록 하는 하위 세트 도구(subset mechanism), 각 개념 및 테이블 구조의 버전을 관리하는 과거력 도구(history mechanism), 각 기관에 적합하게 SNOMED CT를 적용할 수 있도록 하는 개발 도구(developer toolkit)들로 구성되어 있다. SNOMED CT의 핵심 테이블은 개념(concepts) 테이블, 유사어(descriptions) 테이블, 관계(relationships) 테이블로 이루어져 있다. 개념, 유사어, 이들과의 관계는 SNOMED CT를 이루는 핵심 구조로서 각각은 유일한 식별자인 개념코드(conceptID), 유사어코드(descriptionID), 관계코드(relationshipID)를 가지며, 개념코드를 중심으로 연계되어 있다.

2.2 자료의 준비

2005년4월부터 2006년 3월까지 부산대학교병원 신경외과 퇴원요약지를 대상으로 하였다. 입력 내용은 퇴원요약지에 있는 현병력(present illness), 과거력(brief history, past history), 진단명(diagnosis), 평가(assessment), 수술명(operation name)으로 각각의 용어를 입력 받았다.

2.3 용어의 추출 및 정리

입력 받은 용어에 대해 일정한 기준에 의거해서 정리 작업을 진행하였다. 우선 각 항목별로 입력 받은 용어의 중복을 제거하고, SNOMED CT와 매핑하기 위해서 개념 단위로 분리하는 작업을 하였다. DVT와 같은 약어는 deep vein thrombosis로, lung ca와 같은 축약형은 lung cancer와 같은 원형으로 풀어주었다. 또한 pituitary macroadenoma with hemorrhage와 같은 복합개념은 pituitary macroadenoma + hemorrhage로 분리하였다. 진단에 사용된 용어를 최소 개념 단위로 만들기 위해서 용어의 개념을 시간적 혹은 관리적인 면에서 변환시키는 개념 변화자(qualifier)나 주 개념을 임상적으로 더 자세히 설명하는 개념 구분자(modifier)를 주 개념으로부터 분리해 주었다. 또한 일반적으로 축약해서 쓰는 용어는 풀어서 표현했다(예, tibia fx → tibia fracture, T-O → temporal lobe-occipital lobe). 이러한 작업을 마치고 중복을 없앤 1,266개의 용어를 데이터로 사용하였다.

3. 참조용어와의 매핑 결과

3.1 용어의 매핑

의학용어체계에 대한 용어 매핑에 관한 국내 연구로는 SNOMED CT의 퇴원요약지 임상용어 표현정도[11]의 방법에서 보고된 바 있다. 우리는 앞서 언급한 기준으로 정리한 용어에 대해 SNOMED CT의 개념과 매핑하였다. SNOMED CT의 개념과 자동적으로 매핑(pre-coordinated mapping)이 되지 않는 용어들은 CLUE browser를 사용하여 4가지의 규칙을 정하여 매핑 작업을 진행하였다. 일정한 규칙이 없는 조합이 항상 의미를 가진 조합을 만드는 것이 아니므로 자유로운 조합방식으로 생성된 코드에 대

한 의미가 명확하지 못할 수 있기 때문이다. 첫 번째 방법은 SNOMED CT에 하나의 개념으로 표현되어 있지 않은 경우, 두 개 이상의 개념을 후 조합(post-coordinated mapping)하여 원래 용어와 매핑을 시도하였다. 예를 들어, SNOMED CT에는 brain stem glioma라는 개념이 없기 때문에 brain stem(CID=15926001)과 glioma(CID=115240006)을 조합하여 하나의 개념으로 만든 후 매핑하였다. 또한 brain tumor Rt F도 없는 개념이기 때문에, brain tumor(CID=254935002), Rt F(right frontal lobe, CID=314141005)를 조합하여 매핑하였다. 두 번째 방법은 동일한 개념임에도 진단에 사용한 용어와 SNOMED CT 용어의 형태가 다른 경우, 해당 개념을 찾아서 (conceptual mapping) 연결시켜 주었다. 예를 들어, 진단에 사용한 용어는 liver abscess로 되어 있는데, SNOMED CT에는 abscess of liver(CID=27916005)인 경우, abscess of liver로 해당 용어를 매핑하였다. 세 번째 방법은 SNOMED CT의 개념과 용어가 후 조합도 안되고, 동일한 개념도 없으며 일부만 매핑이 되는 경우로 부분적으로 같은 개념만 연결하였다 (partial mapping). 예를 들어, degenerative cervical spondylosis는 후 조합도 안되고, 동일한 개념도 없기 때문에 cervical spondylosis(CID=387800004)로만 연결시켜 주고, degenerative는 SNOMED CT에 없는 개념으로 제외시켰다. 마지막으로 SNOMED CT에 개념이 없는 경우로 분류하였다.

3.2 부분적으로 매핑 된 용어의 분류

SNOMED CT에서 세 번째 방법으로 분류되었던 (partial mapping) 용어들은 다시 세 가지의 형태로 분류되었다. 첫 번째는 개념 구분자의 성격을 갖는 용어들이 매핑되지 않는 경우이다. 예를 들면, pyogenic spondylitis L3/4의 경우, pyogenic은 'pyogenic abscess(CID=402125000)', 'pyogenic pericarditis (CID=15016005)'만 있고, pyogenic 또는 pyogenic spondylitis는 찾을 수 없어 spondylitis(CID=84172003)와 L3(L3 vertebra, CID=181843001), L4 (L4 vertebra, CID=181844007)만 매핑하고 pyogenic은 매핑할 수 없는 개념으로 처리하였다. 두 번째는 진단명은 조합된 개념인데 SNOMED CT에는 입력된 용어와 같은 선 조합된 개념은 없고, 후 조합을 위한 독립된 용어가 부분적으로 SNOMED CT와 일

표 2. 신경외과 추출 용어와 참조용어 간 매핑 결과

	Pre-coordinated mapping terms	Post-coordinated mapping terms	Conceptual mapping terms	Partial mapping terms	Non-mapping terms	Total
Num	192	660	18	379	17	1266
%	15.17	52.13	1.42	29.94	1.34	100

치하지 않는 경우이다. 예를 들어, acute SDH falx의 경우에 선 조합된 개념은 없었고 acute(CID=53737009), SDH(CID=95453001)는 매핑이 되었다. 하지만 falx는 SNOMED CT에서 찾을 수 없었다. 세 번째는 조합된 개념의 진단명에 대한 선 조합된 개념이 없고, 부분적으로 매핑되지 않는 부분이 개념 변화자와 독립된 용어, 두 가지 모두를 포함하는 경우이다. 예를 들어, spinal epidural tumor(metastasis)에서 tumor(CID=108369006)는 매핑이 되었다. 그러나 spinal epidural의 경우 spinal epidural space(CID=54092001), spinal epidural abscess(CID=63627007)으로 찾을 수 있었지만 spinal epidural만으로는 찾을 수 없었다. 또한 metastasis도 metastasis to brain(CID=94225005), metastasis stage(CID=277208005)으로 찾을 수 있지만, metastasis만으로는 찾을 수 없어 세 번째 분류에 포함시켰다.

3.3 매핑의 결과

진단 용어와 SNOMED CT의 매핑을 위해 용어 정리 작업을 거쳐 중복된 단어와 철자 오류가 있는 단어를 정리하여 1,266개의 용어가 추출되었다. 이 용어 중 870개(68.72%)를 SNOMED CT의 용어로 표현할 수 있었고, 나머지 396개(31.28%)는 해당 용어를 부분적으로 찾을 수 있거나 전혀 찾을 수 없는 경우였다. 각 항목별로는 post-coordinated mapping term이 52.13%로 가장 많은 비중을 차지했으며, conceptual mapping이 1.42%로 가장 적은 비중을 차지했다(표 2).

4. Subset 에디터의 구현

4.1 재료의 선택

매핑 결과에서 완전하게 용어로 표현할 수 있는 870개(68.72%)를 Subset 에디터의 데이터로 선택하

였다. 나머지 396개(31.28%)는 해당 용어를 부분적으로 찾을 수 있거나 전혀 찾을 수 없는 경우는 차후 개발된 에디터를 통해서 임상이가 직접 환자의 데이터를 매핑 할 수 있도록 설계하였다. 이는 개발자의 임의가 아닌 실제 데이터를 사용하는 임상이의 의견이 적극적으로 반영되어야 하기 때문이다.

수집된 신경외과 Subset 데이터는 그림 1과 같이 Ns_subset 테이블로 구성하였으며, SNOMED CT는 독립된 용어 서버(vocabulary server)를 구축하고, 추가로 변환(conversion) 테이블을 구성하여, Subset 데이터의 용어가 진단과 관련되었을 경우는 ICD-10, 시술 및 수술과 관련되었을 경우는 ICD-9-CM(ICD, 9th revision, clinical modification) 코드와 매핑 될 수 있도록 하였다. 그 이유는 ICD 계열의 코드는 병원수가체계와 직접적인 연관이 있을 뿐 아니라 사망률 및 이환률의 연간 통계를 위해 보고해야 할 의무가 있기 때문이다[12].

Conversion 테이블은 2005년 3월 소화기계 질환의 참조용어에 대한 한글번역을 수행하면서 SNOME CT의 Map Targets 테이블과 ICD 계열 코드와 변환시키는 Conversion 테이블 간의 매핑 정확도(100%)는 웹 인터페이스 프로토타입을 개발하면서 확인한 바 있다(그림 2). ICD 계열과의 매핑은 수가와 밀접한 연관이 있는 코드이므로 병원 경영에 있어서는 매우 필수적인 용어체계이다.

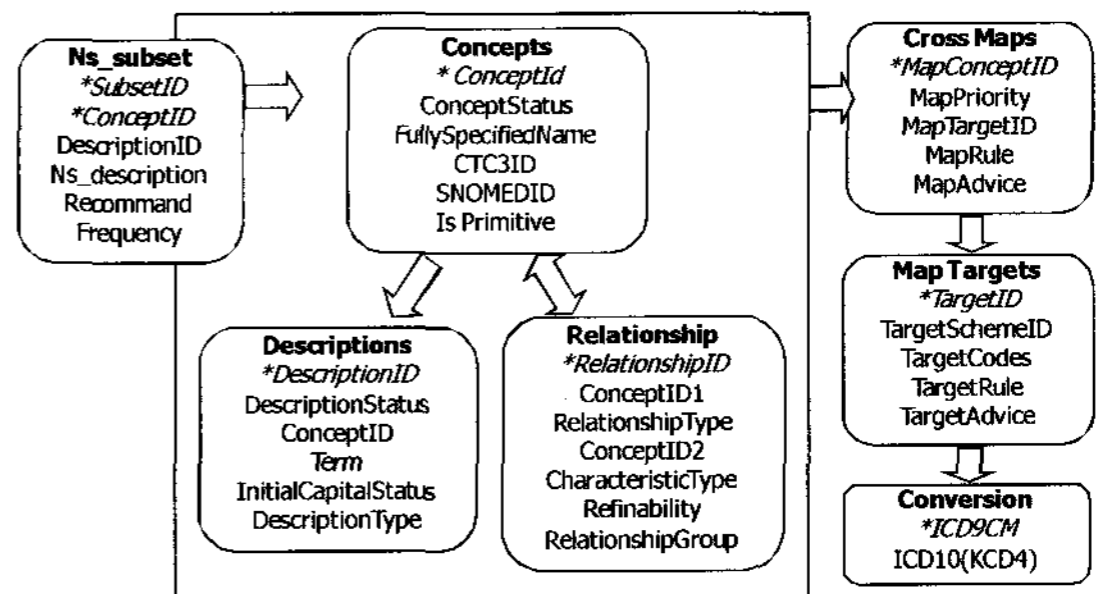


그림 1. SNOMED CT와 매핑 데이터 구조

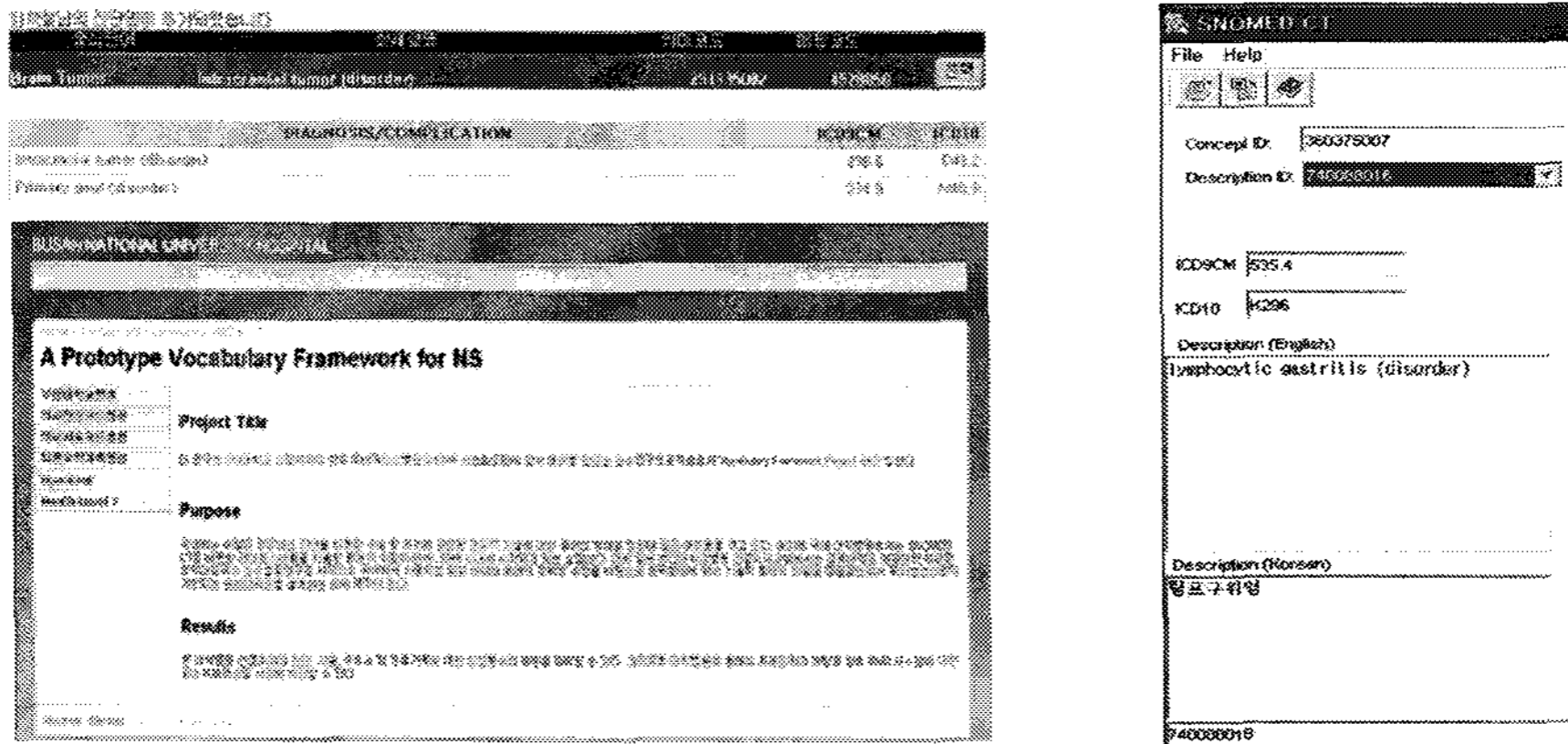


그림 2. 변환 테이블에 의한 ICD와 SNOMED CT 간의 코드 간 변환

4.2 시스템 구현

개발 도구로는 Microsoft에서 제공하는 개발 프로그램인 Microsoft Visual Studio .NET 2003의 C#.NET을 사용하였다. C#.NET은 개발자가 응용 프로그램을 작성할 수 있도록 하는 간단하고 표준적인 모델을 제공하기 때문이다. 또한 프로그램상의 데이터를 저장하고 관리하기 위하여 Microsoft의 SQL 2005를 참조용어의 서버로 사용하였다.

작업을 통하여 추출된 용어는 그림 3의 시스템 구조와 같이 Subset 데이터베이스라고 명명되며 관리된다. 보통의 경우 Subset 데이터베이스가 하나의 각 임상과별 Subset이 되어 실제 의료 환경에서 사용되며 사용의 효율성을 높이고 관리하기 편리하다. 그 이유는 병원정보시스템의 용어서버를 먼저 검색할 경우 용어집단의 크기가 매우 크기 때문에, 검색의 효율성이 떨어지게 된다. Subset 데이터베이스에 있는 용어집단의 크기는 사용자가 결정할 수 있으므로

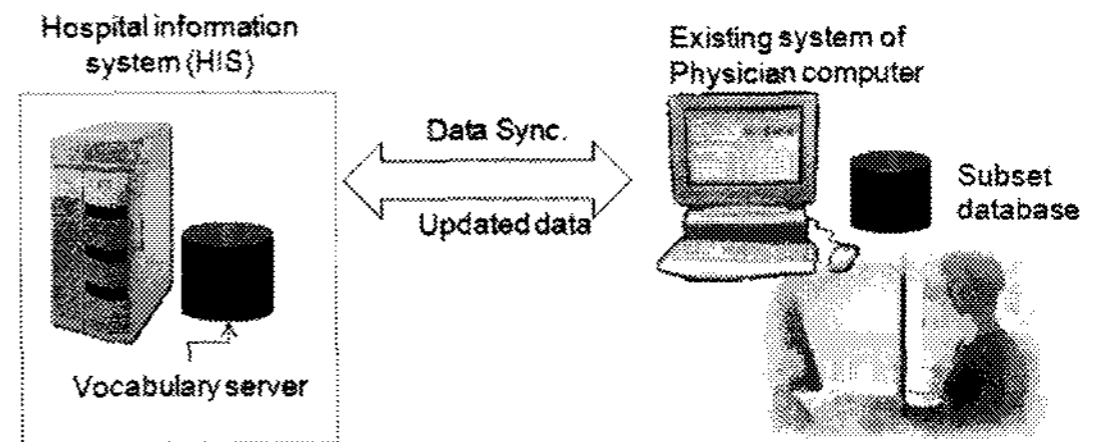
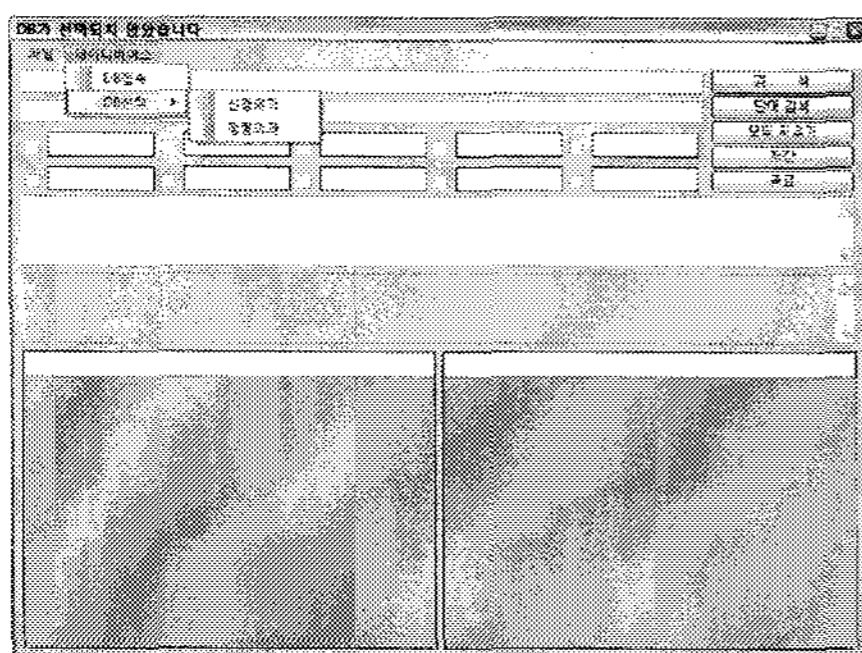
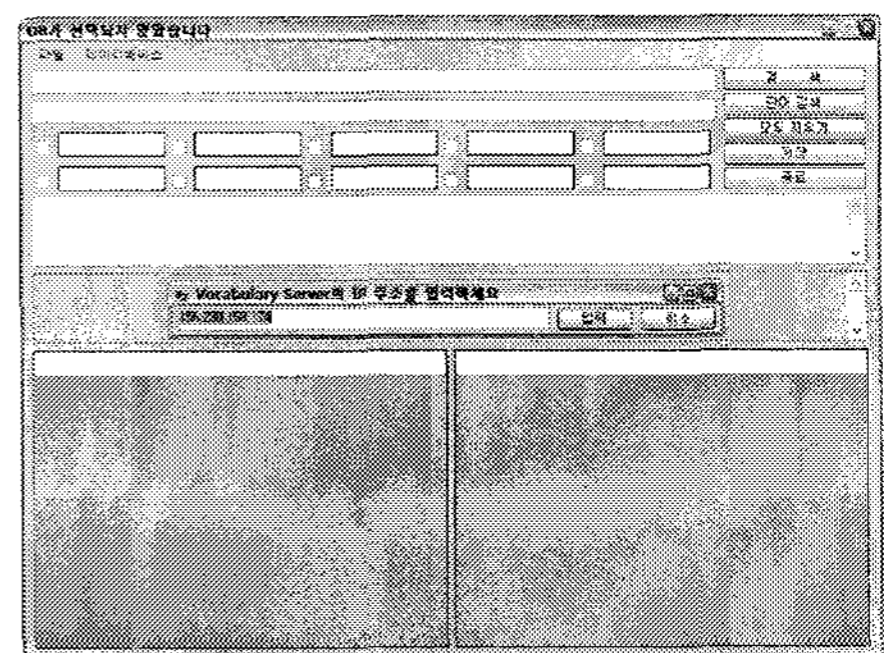


그림 3. 시스템 구조

시스템의 능력에 따라 최대한의 효율성을 가져다 줄 수 있다. 퇴원요약을 위한 의무기록을 작성하기 위해 신경외과 임상외는 개인 컴퓨터에서 환자 정보를 다운로드 하는 것으로 업무를 시작하게 된다. 만약, 최초 검색 시 Subset 데이터베이스에서 검색이 되면 사용자는 아주 빨리 용어를 사용할 수 있게 되며 해당 용어의 평가치를 갱신하여 앞으로도 계속해서 쓰일 수 있음에 대한 가능성을 유지하게 된다. 그림 4의 (a)는 사전에 수집된 용어의 분류로 생성된 subset



(a)



(b)

그림 4. Subset 데이터베이스(a)와 용어서버로의 접속(b)

데이터베이스인 신경외과로 접속하는 화면이다. 검색할 용어가 없을 경우 그림 4의 (b)와 같이 용어서버로 접속을 해서 새로운 용어를 검색하고 업데이트 할 수 있는 기능을 가지고 있다.

용어서버로의 접속과정을 통해 사용자는 다양한 개념과 일치하는 매핑 과정을 거쳐 새로운 개념의 조합을 생성함으로써 기존의 Subset 데이터베이스를 업데이트하거나 수정할 수 있다. 이러한 일련의 과정은 그림 5로 간략히 도식적으로 나타낼 수 있다.

새로운 개념의 조합을 생성하여 저장할 경우에도 그림 6과 같이 Subset 데이터가 임의로 갱신되지 못하도록 하였다.

5. 결론

본 연구는 의학용어체계를 효율적으로 직접 임상에서 사용하기 위해 부산대학교병원의 신경외과환

자의 퇴원요약지에서 사용하고 있는 용어와 SNOMED CT의 개념을 매핑을 통해 비교, 분석 한 결과를 기준으로 참조용어기반의 Subset 에디터를 개발하였다. 참조용어 기반으로 용어를 매핑하는 과정에서 어떤 식으로 개념을 추출해야 할 것인지 개념을 구분하는 기준 설정에 있어 어려움이 있었다. 이 경우에는 각 항목별로 SNOMED CT와 비교하기 쉬운 형태인 개념 단위로 입력하였다. 국내 의무기록지에서 사용되는 용어들을 참조용어들과 매핑하는 연구는 서울대학교병원을 중심으로 SNOMED CT의 퇴원요약지 임상용어 표현정도 연구에서 최근 2년간 수행해 온 연구이다[11].

본 연구는 추출된 자료를 정리, 분석한 후 Subset을 개발하고 차후 추가되는 용어에 대해서는 임상의가 직접 Subset에서 참조용어와의 매핑을 통해 용어체계를 생성될 수 있도록 하였다. 참조용어인 SNOMED CT는 부산대학교병원에서 사용하고 있는 퇴원요약지 용어 중 부분적으로 매핑된 것을 포함하여 98.72%의 개념들이 부산대학교병원에서 사용하고 있는 신경외과 퇴원요약지 용어를 비교적 잘 표현하고 있음을 확인 할 수 있었다. 하지만 시스템에 포함된 것은 부분 매핑 된 것을 제외한 사전 작업을 통하여 추출되고 정확히 매핑 된 68.72%의 데이터를 사용함으로써 실제 의료 환경에서 명확히 사용될 수 있는 신뢰도가 있는 데이터가 되도록 하였다.

국내에서는 전문용어체계의 중요성을 인식하고, 전문용어언어공학연구센터(Korean Terminology Research Center for Language and Knowledge Engineering, KORTERM)가 1998년에 설립되어 의학분야의 전문용어 기초자료 데이터베이스를 구축하고자 하였으나 의료분야의 방대함과 극도의 전문성 때문에 연구 계획 및 성과가 크게 나타나지 않았다. 연구가 좋은 결실을 가져오지 못하는 큰 이유는 의학에서 사용하는 언어의 자체 개발이 단독연구가 아닌 의학적인 지식, 의료전반에 대한 폭넓은 이해, 언어 및 공학적 지식, 정보학적인 지식 등 다학제간의 전문적인 이해를 하는 전문가의 부족과 협조체제가 이루어지기 힘들었기 때문일 것이다.

실제 포괄적인 의료환경에서 사용되고 있는 의무기록과 자연어는 일반 환경의 자연어와는 달리 매우 전문적 용어로 묘사되고 있어 이를 포괄하는 용어에 대해서는 지나치게 간과해 왔던 점이 국내에

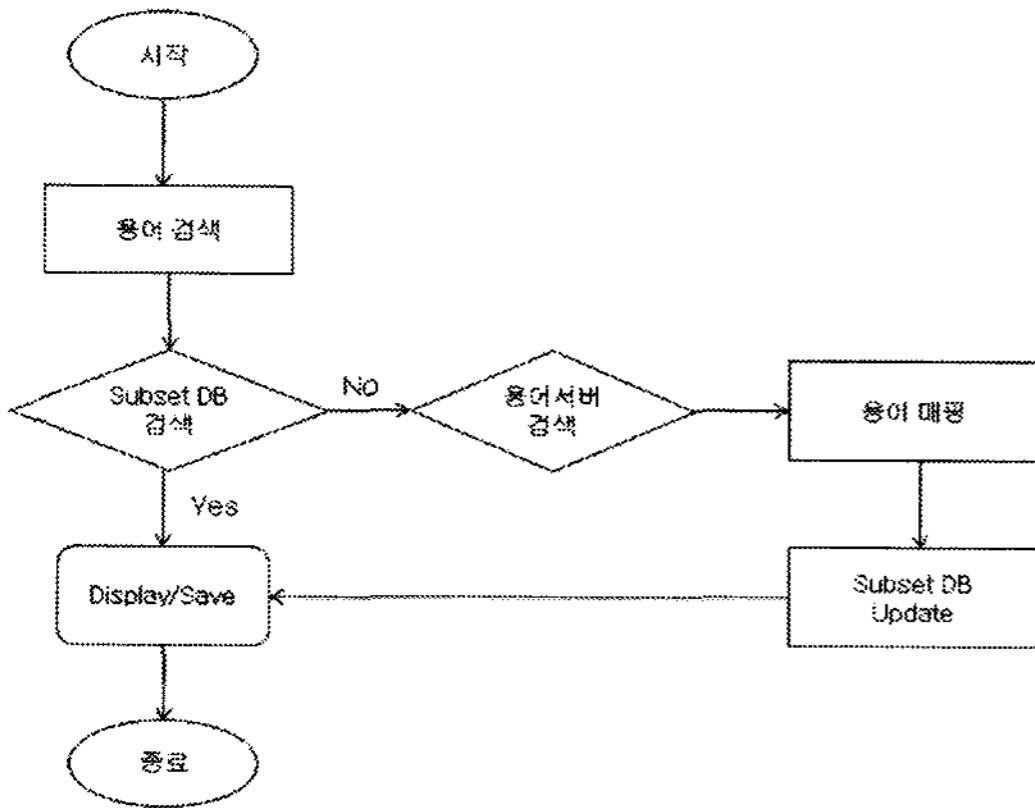


그림 5. 용어 검색 및 생성의 흐름

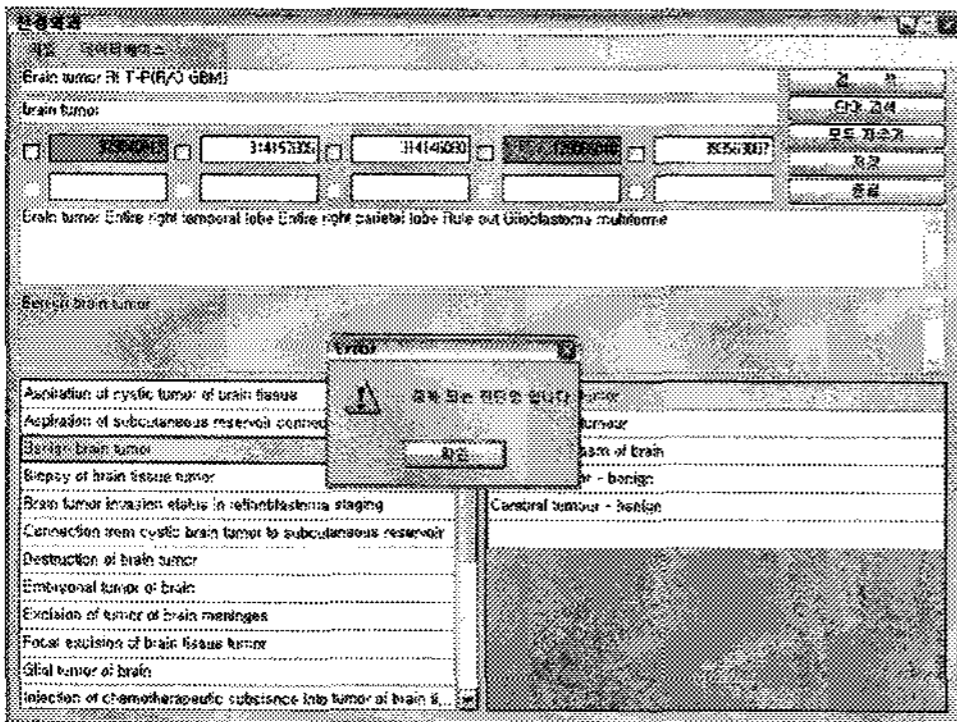


그림 6. 중복코드 생성 방지를 위한 예외 처리

전자의무기록의 확산을 방해하는 최대의 걸림돌이 되고 있다.

SNOMED CT는 1963년부터 국가에서 관심을 갖고 국민의 건강과 관련된 코드체계 연구를 시작한 이후 40년 동안 진행시켜온 연구의 결실로서 정보화 기술과 함께 참조용어로서 전 세계에서 신뢰하고 있는 표준용어체계로 인식되었다.

국내에서는 보건복지부산하 EHR(electronic health record)핵심공통기술연구개발사업단 중심으로 2007년 11월부터 ‘참조용어기반 임상영역 용어세트 개발 및 EMR 적용을 위한 구조화된 데이터 입력 도구 개발’의 연구를 공공의료정보화 사업기관인 국민건강보험공단 일산병원을 대상으로 시작하였다.

이러한 점에서 볼 때 본 연구는 국내에서 첫걸음을 시작한 참조용어기반의 용어세트 개발을 위해 대학병원의 의료환경에 적용하였을 뿐 아니라 실제 임상영역의 신경외과 전문의에 의해 현재 사용되고 있다는 점에 큰 의의가 있다.

차후 연구과제로는 초기 단계인 참조용어를 기반으로 실제 유용한 데이터를 병원데이터베이스에 수집하여 의사결정지원시스템을 개발하는 것이다. 기존에는 수가를 위한 ICD 계열 코드에 의존하여 환자가 가지고 있던 다양한 정보들이 누락된 채 저장되는 문제로 의료인이 기술하는 대부분의 중요한 정보가 시스템에서 저장되지 못했다.

물론, SNOMED CT가 의료환경에서 사용되는 용어를 100% 완벽히 포괄해주지는 않지만 이에 대한 연구를 기반으로 향후 국내 의료환경에 맞는 용어체계를 구축하는 연구의 밑거름이 될 것이다.

참 고 문 헌

- [1] Health Level Seven, "HL7 Version 3 Development Framework (HDF)," Health Level Seven Inc, 2005.
- [2] K. Donnelly, "SNOMED-CT: The advanced terminology and coding system for eHealth," *Stud Health Technol Inform*, Vol.121, pp. 279-290, 2006.
- [3] K. Giannangelo and L. Berkowitz, "SNOMED CT helps drive EHR success," *J Ahima*, Vol. 76, pp. 66-67, 2005.
- [4] 이종혁, 임상용어코드체계의 효율적 활용을 위한 Subset 관리방법 연구, 경북대학교 석사학위 논문, 2005.
- [5] Z. Min, D. Baofen, M. Weeber, and A. M. van Ginneken, "Mapping OpenSDE domain models to SNOMED CT. Applied to the domain of cardiovascular disease," *Methods Inf Med*, Vol.45, pp. 4-9, 2006.
- [6] R. K. Los, A. M. van Ginneken, and J. van der Lei, "Extracting data recorded with OpenSDE: possibilities and limitations," *Int J Med Inform*, Vol.74, pp. 473-80, 2005.
- [7] A. C. Venema, A. M. van Ginneken, M. de Wilde, and A. J. Bogers, "Is OpenSDE an alternative for dedicated medical research databases? An example in coronary surgery," *BMC Med Inform Decis Mak*, Vol.7, p. 31, 2007.
- [8] S. H. Brown, P. L. Elkin, B. A. Bauer, D. Wahner-Roedler, C. S. Husser, Z. Temesgen, S. P. Hardenbrook, E. M. Fielstein, and S. T. Rosenbloom, "SNOMED CT: utility for a general medical evaluation template," *AMIA Annu Symp Proc*, pp. 101-105, 2006.
- [9] The Systemized Nomenclature of Medicine (SNOMED) and College of American Pathologists (CAP), Available from: <http://www.snomed.org>. 2007.
- [10] H. J. Chin and S.H. Kim, "Standardization of Main Concept in Chief Complaint based on SNOMED CT for Utilization in Electronic Medical Record," *J Kor Soc Med Informatics*, Vol.9, No.3, pp. 235-247, 2003.
- [11] S.H. Kim, S. B. Han, J.W. Choi, "A comparative study on concepts of the SNOMED CT and clinical terms in discharge summary of SNUH," *J Kor Soc Med Informatics*, Vol. 10. pp. 53-56, 2006.
- [12] S. Bowman, "Coordinating SNOMED-CT and ICD-10," *J Ahima*, Vol.76, pp. 60-61, 2005.



조 훈

1980년 서울대학교 수학과 학사
 1986년 남캐롤라이나주립대학
 전산학 석사
 1992년 유타주립대학 의료정보
 학 박사
 1999년~현재 경북대학교 의료정
 보학과 교수

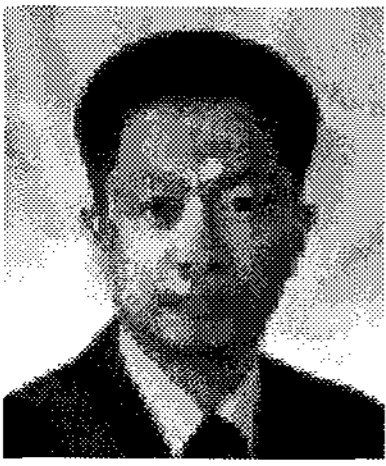
2002년~현재 과학진흥재단 의료정보위원, 산업자원부
 기술표준원 정보위원
 2003년~현재 대한의료정보학회 부회장 및 부편집위원
 장
 관심분야 : 알고리즘, 의료영상시스템, MFER, 유비쿼터
 스 헬스케어 로봇시스템



김 형 회

1991년 부산대학교 의학과 학사
 1994년 부산대학교 의과대학 임
 상병리학 석사
 2002년 부산대학교 의과대학 임
 상병리학 박사
 2004년~현재 경북대학교 의과대
 학 의료정보학 박사과정

2000년~현재 부산대학교 의학전문대학원 진단검사의
 학과 부교수
 2007년~현재 부산 U-city 포럼 U-Health 분과위원장
 관심분야 : 표준용어체계, 유비쿼터스 병원정보
 시스템



최 병 관

1993년 부산대학교 의학과 학사
 2001년 부산대학교 의과대학 신
 경외과학 석사
 2005년 부산대학교 의과대학 신
 경외과학 박사
 현재 부산대학교 의학전문대학원
 신경외과학과 조교수

2007년~현재 부산대학교병원 의료정보과 TFT
 관심분야 : 임상용어(clinical Terminology)



최 영 연

2000년 서울시립대학교 생명과
 학과 학사
 2002년 서울시립대학교 생명과
 학과 석사
 2005년~현재 경북대학교 의학전
 문대학원 의학과 3년 재
 학 중

관심분야 : 생명정보 및 임상정보



김 화 선

1991년 마산대학 간호학부 학사
 2003년 인제대학교 컴퓨터공학
 과 석사
 2007년 경북대학교 의료정보학
 과 박사
 2007년~현재 경북대학교 의과대
 학 의료정보학과 Post-
 Doc

관심분야 : 병원정보시스템, 온톨로지, 참조정보모델, 임
 상문서구조



홍 해 숙

1977년 경북대학교 간호학과 학
 사
 1979년 경북대학교 간호학과 석
 사
 1990년 대구대학교 생물학 박사
 1979년~현재 경북대학교 간호대
 학 교수

관심분야 : 방문간호, 유전정보, 간호용어체계, 기초간호
 과학