MeSH를 확장한

한국보건의료정보분야 주제어 연구

- 대한의료정보학회지를 중심으로

연세대학교 보건대학원

보건정보관리학과

권 애 경

MeSH를 확장한
한국보건의료정보분야 주제어 연구

- 대한의료정보학회지를 중심으로

지도 채 영 문 교수

이 논문을 보건학 석사학위 논문으로 제출함.

2001년 12월 일

연세대학교 보건대학원

보건정보관리학과

권 애 경

권애경의 보건학 석사학위논문을 인준함

심사위원 _____ 인

심사위원 _____ 인

심사위원 _____ 인

연세대학교 보건대학원

2001년 12월 일

감사의 글

항상 부족하게 느껴졌던 배움에 대한 아쉬움 때문에 자신을 학문과 지식으로 채워보고자 시작했던 대학원 생활을 이제 마감하게 되었습니다. 어느덧 2년 반의 세울을 넘기게 되면서, 지난 학업의 성과를 정리하는 의미에서 부족하나마 이 글을 내놓게 되었습니다. 지난 대학원 기간동안 괄목상대하기보단 일신우일신하며 생활하기를 바랬던 저에게 부족함과 아쉬움이 남습니다. 하지만 이 논문을 쓰면서 배우고 생각하고 느꼈던 많은 것들이 앞으로의 일들에 큰 보탬이 될 것이라 느낍니다.

먼저 지금까지의 결실이 완성될 수 있드록 따뜻한 가르침으로 조언해주신 모든 분들게 진심으로 감사의 말씀을 올립니다. 깊은 관심과 배려로 이 글이 완성되기까지 지도해주신 채영문 교수님, 바쁘신 와중에도 귀한 시간을 내어 조언을 아까지 않으셨던 공재근 단장님, 홍준현 과장님께도 깊은 감사를 드립니다.

오늘이 있기까지 늘 격려를 아끼지 않으신 허선 교수님, 김수녕 교수님께도 감사를 드리며, 바쁜 업무중에도 결실을 맺을 수 있도록 배려해주신 보건의료기술연구기획평가단 가족들께도 감사를 드립니다. 그리고 많은 도움을 주신 호승희 선생님을 비롯한 대학원 후배님께 진심어린 감사의 마음을 전합니다.

마지막으로 한겨울 늦은 귀가길을 걱정하시던 아버님, 어머님과 사랑하는 언니, 동생, 가족들에게 감사한 마음을 전하고 싶습니다. 그리고 언제나 사랑으로 격려해준 사랑하는 이에게 고마운 마음을 전합니다.

2001 년 12 월
권 애 경 올림

목 차

표 목 차

그 림 목 차

국 문 요 약

첨단화된 연구의 폭발적인 양적 증가와 새롭고 급변하는 보건의료정보는 보건의료과학과 정보기술 및 컴퓨터공학 분야의 관련 테크놀로지와 이론들을 결합한 학제간 협력분야임에도 불과하고 체계적으로 운영 관리하는데 그동안 우리나라 보건의료인들은 상대적으로 타 분야 전문가에 비해 무관심해 왔다. 그러나 중장기적으로는 새로운 연구지원환경 구축 및 한국형 국립 보건의료과학기술도서관 설립·운영의 추진의 필요성이 꾸준히 제기되어 온 것 또한 사실이다. 보건의료과학기술도서관 설립을 위해서는 지속적이고 체계적인 준비와 관련 전문인력양성, 통합 및 분산 연구 정보시스템 기반기술 연구 등 정보인프라 구축이 요구된다.

보건의료과학기술도서관 설립의 바탕이 될 국내 보건의료분야 문헌의 데이터베이스 구축과 검색을 위해서는 통제된 용어의 사용이 절대적으로 필요하다. 즉, 용어의 통일성과 일관성을 유지하기 위해서는 논문을 발행할 때 키워드 선정시부터 통제된 용어의 사용이 이루어져야 한다. 이러한 인식 하에 보건의료분야 학술지, 특히 의학학술지의 대부분이 저자가 자신의 논문의 키워드를 선정할 때 MeSH 용어를 사용하도록 투고규정에 권장하고 있지만 아직까지 보건의료분야 200여개의 학회중 단지 23개 학회만 용어집 혹은 색인집을 발표하고 있는 실정이다.

따라서 본 연구에서는 대한의료정보학회지를 중심으로 논문의 저자들이 부여한 저자용어와 동일한 논문에 대하여 MEDLINE 레코드에 부여된 MeSH 용어를 비교하여, 저자용어와 MeSH 용어의 일치 정도, 저자용어와

MeSH 용어의 차이점 등을 분석한 후 이를 바탕으로 보건의료정보분야 주제어집을 제시하였다. 1995년부터 200년까지 대한의료정보학회지에 발표된 172편의 논문을 대상으로 분석한 결과, 부분적으로 일치한다고 볼수 있는 경우는 25.56%였으며, 완전히 일치하는 용어는 단지 11.72%에 불과하였으며, 이를 기반으로 제시되어진 보건의료정보분야 주제어의 수는 12,431개 였다.

본 논문에서는 국내 보건의료정보분야 정보공유기반을 확립하기 위한 작은 초석인 대한의료정보학회지를 기초로 한 보건의료정보분야 주제어집(용어집)을 제시하였다. 물론 이렇게 구축된 주제어집은 미국 국립의학도서관에서 개발된 MeSH와 국내에서 새로 추가되는 용어를 보완하는 등의 계속적인 수정·보완이 이루어져야 할 것이다. 이러한 다각적인 노력을 통하여서만 수준 있는 MeSH 용어 부여작업이 가능하고, 국제적으로 유통되는 의학정보의 생산이 비로소 이루어지게 될 것이다.

Ⅰ. 서 론

1. 연구의 배경 및 의의

의학은 날로 발전해가고 있고 기존의 잡지에서 나오는 논문과 용어의 수도 날로 증가해 가고 있다. 또한 널리 발표되거나 읽혀지지 않는 논문이 실린 잡지까지 포함한다면 훨씬더 많은 논문이 될 것이다. 이런 방대한 양의 논문을 우리가 쉽게 찾아보고 참고하여 올바른 방향으로 논문을 이해하고 쓰려면 색인단어가 제대로 정해져야 한다. 자신의 논문을 대표하고 검색에 유용하게 쓰이는 주제어의 중요성은 근거중심의학의 발전과 함께 증가하고 있으며 영문 주제어의 중요성 또한 증가하고 있다(Lowe 등, 1994).

각 논문의 요약에 실리는 주제어는 대체로 두 가지 목적을 가진다. 그 하나는 자신의 논문을 대표할 만한 단어를 제시함으로써 논문의 주의미가 무엇인지 보여주는 것이고 다른 하나는 논문검색에 유용하게 사용하고자 하는 것이다. 근자에 와서 이러한 목적 중에서 후자의 비중이 더 커지고 있는 것은 근거중심의학(evidence based medicine)의 발전에 힘입은 바가 크다(김수영, 1998). 근거중심의학은 문제를 구조화하고 문제에 대한 문헌검색을 하고 각 문헌을 비판적으로 평가하는 것을 핵심내용으로 하는 의학의 새로운 조류이다. 그런데 이러한 근거중심의학의 발전은 논문검색의 중요성이 한층 커지는 결과를 가져 왔으며 따라서 검색에 중요한 역할을 담당하는 주제어의 중요성도 더욱 커지게 되었다.

일반적으로 학회지의 논문초록 양식은 연구배경, 방법, 결과, 고찰의 순으로 기술하고 중심단어를 삽입한 것이다. 국내에서 논문에 국문초록대신 영문초록을 삽입한 이유는 우리나라의 연구성과를 외국에 알리고 학회지가 인덱스 메디쿠스 같은 국제색인 잡지에 실릴 때를 대비한 것이다(대한의료정보학회, 1999). 그런데

논문의 얼굴로 취급되어 중요성을 인정받고 있는 국문초록과는 달리 영문초록의 경우 논문의 구색을 맞추는 정도로 취급되어 주제어를 부여하는 것에 대한 체계가 정해져 있지 않아 주제어를 통한 문헌조사는 거의 불가능한 실정이다. 반면에 기존의 국제적인 의학잡지들은 거의 대부분이 영문주제어를 정할 때 MeSH 용어로 하고 있다(서울대학교 의학도서관, 1999).

대부분의 학회지의 투고규정에는 영문주제어는 가능하면 MeSH 용어를 사용하도록 권고하고 있다(International Committee of Medical Journal Editors, 1997). 그처럼 권고한 이유는 국제적 조류 때문이기도 하지만 논문검색의 두 가지 방법인 원문검색과 MeSH 검색 중에서 MeSH가 논문을 정확하고 빠짐없이 찾아주기 때문에 앞으로의 검색에 대비하기 위해서이다(대한의학학술지 편집인협의회). 미국국립의학도서관(National Library of Medicine, NLM)의 의학용어 주제어집(Medical Subject Headings, MeSH)은 MEDLINE(http:/ / www.ncbi.nlm.nih.gov/ PubMed/ , PubMed)과 같은 컴퓨터기반 생의학 데이터베이스의 중요한 인덱스로 사용되어져 왔다. MEDLINE 데이터베이스는 NLM에 의해 만들어지고 유지되어 오고있으며, 1966년이래 7백만 이상의 생의학관련서적을 담고 있다. 대다수의 학회지 투고규정에는 주제어를 부여할 때 MeSH 주제어를 사용하도록 하고 있으나 이러한 원칙이 잘 지켜지지 않고 있어 검색에 유용하게 쓰이지 않고 있다(김병국 등, 1998).

MEDLINE은 3,500개 이상의 저널을 포함하고 있으며, 이중 약 3분의 2는 영어로 쓰여진 저자초록을 포함하고 있다. 이것을 우리가 적절히 사용한다면, MeSH는 의학논문을 찾고 검증하는데 효과적인 도구로 사용될 수 있다(Fremer 등, 995). 하지만 이전에 선행되어야 할 가장 중요한 과제는 아마도 영문초록의 주제어를 찾기 쉽고 접근하기 쉬운 MeSH 용어로의 통일화 과정이다. 학회지에

실린 논문의 영문주제어가 과연 얼마나 정확한 MeSH 용어를 사용하고 있는지를 검토해 보는 것이 꼭 필요할 것이며 각 주제어가 실제 논문 내용을 얼마나 잘 나타내고 있는가를 검증한 후 결과를 바탕으로 MeSH 용어집에는 들어있지 않으나 학회지에서 그 특성상 주제어로 색인되는 것이 다른 연구자가 찾아보기 쉬울 뿐 아니라 그 논문을 대변하는 것으로 인정받을 만한 것들을 선별하는 작업이 필요하다. 이를 위하여 국내의 논문들을 바탕으로 주제어 선정에 관한 정확성을 평가하는 작업으로는 1992년부터 1997년까지 가정의학회지에 발표된 논문 중 무작위 60편을 추출하여 저자가 논문의 주제어를 적절하게 사용하였는지 분석하였다. 2명의 가정의학전문의가 합의하여 분석대상논문의 중심단어를 선정하고 저자용어를 판단한 결과 MeSH용어와 완전히 일치하는 경우가 17.9%라고 보고하였다(김병성 등, 1999). 또 다른 연구로는 문헌정보학전공자가 대한기생충학회지가 MEDLINE등재 학회지인 점을 활용하여 MEDLINE의 MeSH 전문가가 부여한 MeSH 용어를 논문 단위로 대조하여 일치여부 및 차이점을 초래하는 여러 요인을 비교하고 분석한 것이 있다(이춘실 등, 2000). 그러나 이들 연구는 선택된 단어들의 의학전문분야의 특수성과 전문성, 주제어 선정의 타당성을 고려하지 못하고 MeSH 용어와의 일치에 국한시킨 단점을 가지고 있다.

본 연구는 그들의 연구와 그 규모와 분석방법론에서 큰 차이가 있다. 대한의료정보학회지 1권 1호(1995년)에서 6권 4호(2000년)까지 전체의 논문 172편을 대상으로 하여 무작위추출시의 오류를 줄였으며, 각각의 중심단어가 MeSH 용어와 갖는 차이점을 저자가 사용한 용어를 기준으로 MeSH browser를 이용하여 변환 및 분석하였다. 분석결과를 기초로 2002년 MeSH 중 보건의료정보분야 용어를 구분하고, 대한의료정보학회지에서 다빈도로 사용되어진 용어들 중 MeSH에서 반영하지 못한 용어를 추가하여 "한국보건의료정보분야 주제어집"을 제시하였다.

2. 연구의 목적

본 연구는 대한의료정보학회지 1권 1호(1995년)에서 6권 4호(2000년)까지의 논문 172편을 대상으로, 각각의 중심단어가 MeSH 용어와 갖는 차이점을 분석하고, 이 결과를 기초로 MeSH 중 보건의료정보분야 용어를 구분하여 "한국보건의료정보분야 주제어집"을 제시한다. 따라서 본 연구의 구체적인 목적은 다음과 같다.

첫째, 1995년 1권 1호부터 2000년 제 6권 4호까지의 대한의료정보학회지 논문의 영문 주제어의 MeSH 용어 사용실태를 조사한다.

둘째, 저자가 설정한 영문주제어를 기준으로 MeSH browser를 통하여 MeSH 용어로 변환한 후 기존 영문주제어와의 정확도를 평가한다.

셋째, MeSH의 15개 범주 중 보건의료정보 관련분야를 설정한 후, MeSH를 바탕으로 한 한국보건의료정보분야 주제어집을 제시한다.

Ⅱ. 이론적 배경

1. 의학주제어 개요

의학주제어(Medical Subject Headings, MeSH)는 미국 국립의학도서관(National Library of Medicine, NLM)에서 책, 논문 등을 분류하기 위해 사용하고 있는 계층화된 통제 의학용어 분류집(hierarchically controlled)이다. MeSH는 현재 세계 수많은 기관에서 의료 정보와 논문, 책, 자료 색인을 위해 사용하고 있다. 현재 37,862 개의 용어와 100,132개의 additional heading(chemical)을 포함하고 있으며 해마다 수천 개의 새로운 단어가 추가되고 있다. NLM의 인

덱스는 다수의 MeSH vocabulary의 MeSH 용어로 이루어져 있으며, 대부분의 MEDLINE 검색시스템은 검색시에 MeSH를 이용하여 할수 있도록 되어 있다. 결국 MeSH는 의학문헌 검색의 중요한 관문이며 그 중요성은 MEDLINE 데이터베이스가 의학전문가들에게 정보의 근원으로써 결정적인 역할을 할수록 더욱 증가될 것이다(Lowe 등 1994). 표 1의 리스트는 MeSH 용어를 사용하는 데이터베이스의 예이다. MeSH에서 하나의 개념은 하나의 용어로 대표되어지며, 중요한 개념이 새로이 생성되거나 중대한 변경이 발생하게 되면, NLM은 새로운 용어를 추가한다. 새로운 논문에 MeSH를 이용한 인덱싱은 NLM의 indexer들이 논문전체중 대표적인 단어(대부분 10-12)를 선택, 첨가하여 이루어진다(National Library of Medicine, 1998).

검색자들은 그 MeSH 용어를 이용하여 다른 논문에서 같은 용어로 부여된 것들을 빠르게 검색할수 있다. MEDLINE 검색시, '중요 개념(Major Concept)'을 이용한 검색으로 얻을수 있는 이득은 한정과 지정을 통하여 관계없는 것들의 검색을 줄여주는데 있다. 이런 전략의 효율성은 대략적으로 개념위주의 검색을 하지 않을때에 비해 약 41%의 불필요한 자료를 걸러준다 (Mitchell 등, 1992).

표 1. Major NLM Databases Indexed with MeSH

Database	Contents
AIDSLINE	Citations to the AIDS literature
AIDSTRIAL	Active and closed clinical AIDS trial
BIOETHICS	Citations to biomedical ethics literature
CANCERLIT	Citations to the cancer literature
HEALTH	Health care administration and plan
MEDLINE	Citations to the biomedical literature
TOXLINE	Citations to the toxicology literature

가. MeSH 색인의 필요성

MeSH는 "통제된(controlled)" 언어이고 "계층화(hierarchically)"된 언어이다. MeSH가 통제되어 있다는 의미는 하나의 개념에 하나의 주제어를 부여한다는 의미이다. 예를 들어 B lymphocytes와 B cell은 같은 단어이지만 B lymphocytes로 논문이 색인되어 있지 않으면 B lymphocytes로 검색하면 이 논문을 찾을 수 없게 된다. 비슷한 예로 다음과 같은 제목을 가진 논문을 찾는 것이 필요하다고 할 경우 이탤릭 처리된 각 단어는 모두 같거나 비슷한 의미를 가지고 있지만 실제 논문에서는 다른 단어를 사용하고 있다.

- ***Childhood victimization*** and the development of personality disorders.
- ***Childhood maltreatment*** increases risk for personality disorders during early adulthood
- The medical evaluation of the sexually ***abused child***: lessons from a decade of research
- Estimation of stress in ***child neglect*** from thymic involution.

이 경우 이러한 의미에 대해 *child abuse*라는 용어로 색인하기로 약속이 되어 있으면 검색하는 사람은 하나의 검색어로 필요한 논문을 모두 찾을 수 있게 된다. 이 경우 *child abuse*라는 단어는 MeSH에 해당한다(Darmoni SJ, 1995). MeSH로 색인하고 MeSH로 검색을 해야 하는 가장 큰 이유는 검색의 민감도(sensitivity)와 특이도(specificity)를 높이기 위함이다. 검색의 민감도란 필요한 논문을 찾을 수 있는 정도를 말하고 특이도란 적절하지 않은 논문을 배제할 수 있는 정도를 말한다. MeSH로 검색하고 색인 하게 되면 검색의 민감도와 특이도는 모두 올라가게 되므로 검색의 효율을 높여주는 효과가 있다(O'Rourke 등, 1999).

MeSH의 또 다른 특성은 Tree structure로 되어 있다는 것이다. MeSH 용어들은 비슷한 단어끼리 하나로 묶여 있는데 이를 tree라고 한다. MeSH tree의 가장 위에는 15개 범주(표 2)가 있고 그 아래에는 각각의 의미에 가지처럼 MeSH 용어가 배열되어 있다. 하나의 MeSH 용어는 두 개 이상의 가지에 속해 있을 수 있다. 이런 tree 구조(그림 1)는 검색에 있어서 두 가지 이점을 제공한다. 하나는 MeSH로 검색을 할 때 broader(more general) 혹은 narrower(more specific)로 검색을 할 수 있다는 것이고 다른 하나는 확장(explosion)이라 불리는 검색이 가능하다는 점이다(그림 2). 확장이란 더 specific한 검색을 하기 위해 현재 계층 아래에 있는 MeSH 용어까지 같이 검색을 한다는 의미이다(Anita 등, 2001).

표 2. MeSH의 15개 범주

번호		설 명
1	A.	Anatomy
2	B.	Organisms
3	C.	Diseases
4	D.	Chemicals and Drugs
5	E.	Analytical, Diagnostic and Therapeutic Techniques and Equipment
6	F.	Psychiatry and Psychology
7	G.	Biological Sciences
8	H.	Physical Sciences
9	I.	Anthropology, Education, Sociology and Social Phenomena
10	J.	Technology and Food and Beverages
11	K.	Humanities
12	***L.***	***Information Science***
13	M.	Persons
14	N.	Health Care
15	Z.	Geographical Locations

Information Science [L01]
Medical Informatics [L01.700]
Medical Informatics Computing [L01.700.568]
Computer Literacy [L01.700.568.070]
Computer Systems [L01.700.568.080] +
Computing Methodologies [L01.700.568.110]
Algorithms [L01.700.568.110.050]
Artificial Intelligence [L01.700.568.110.065] +
Automatic Data Processing [L01.700.568.110.085] +
Computer Graphics [L01.700.568.110.108] +
Computer Simulation [L01.700.568.110.160]
Image Processing, Computer-Assisted [L01.700.568.110.308] +
Mathematical Computing [L01.700.568.110.680] +
Signal Processing, Computer-Assisted [L01.700.568.110.800]
Software [L01.700.568.810] +

그림 1. MeSH Tree 구조의 예

Information Science [L01]
Medical Informatics [L01.700]
Medical Informatics Computing [L01.700.568]
Computer Literacy [L01.700.568.070]
Computer Systems [L01.700.568.080] +
Computing Methodologies [L01.700.568.110]
Algorithms [L01.700.568.110.050]
Artificial Intelligence [L01.700.568.110.065] +
Automatic Data Processing [L01.700.568.110.085] +
Computer Graphics [L01.700.568.110.108] +
Computer Simulation [L01.700.568.110.160]
Image Processing, Computer-Assisted [L01.700.568.110.308] +
Mathematical Computing [L01.700.568.110.680] +
Signal Processing, Computer-Assisted [L01.700.568.110.800]
Software [L01.700.568.810] +

그림 2. MeSH Explosion의 예

나. MeSH의 특징과 형태

1) MeSH 용어의 형태적 특성

MeSH 용어에는 몇 가지 특징이 있다. 우선 기관에 대해서는 Anglo-saxon으로 표현하고 부가적인 형태로 쓰일 때는 Latin이나 Greek으로 표현하다. 그래서 renal disease 혹은 nephrologic disease 보다는 kidney disease로 표현한다. 또한 알파벳 순서로 나열되어 있어서 전체를 볼 때는 알파벳순서로 찾아보기를 한다. 여러 단어로 되어 있는 경우 비슷한 개념을 같이 묶기 위해서 순서를 바꾸는 경우가 흔하다(예: "Management Decision Support Systems", "*Decision Support Systems, Management*"). 하지만 바꾸지 않는 경우도 있으므로 반드시 MeSH browser를 이용한 검색 및 확인이 필요하다. MeSH 용어는 용어에 따라 단수 혹은 복수 형태를 취하고 각각은 나름대로의 이유가 있다(예: Information Science, Medical Informatics). 어떤 경우는 명사형을 취하고 어떤 경우는 형용사형을 취한다(예: Viral proteins, Virus cultivation).

2) Cross-reference

Cross-reference 란 검색 중에 고려해야 하는 관계용어이다. 여기에는 몇 가지 종류가 있다. 첫째, Entry term은 MeSH 용어는 아니지만 해당 MeSH의 동의어 혹은 비슷한 용어이다. Cancer나 Tumor 등은 MeSH 용어는 아니지만 이 단어로 검색해도 Neoplasm으로 검색한 것과 비슷한 결과를 얻도록 연결되어 있다. 둘째로 see related(also)는 현재 MeSH 용어와 개념적으로 연관되어 있는 다른 용어로 이런 MeSH 용어로도 검색을 해 보라는 의미이다. 이 경우 Antibody, Neoplasm은 see also에 해당한다. 셋째로 consider also가 있다. 이것은 언어학적 줄기가 같

은 것으로 이 경우에는 Cancer, Carcino, Onco 등이 여기에 해당한다(National Library of Medicine, 1997).

3) Annotations

MeSH에는 여러 가지 주석이 있다. Annotation은 indexer, catalogers, searchers를 위한 기술적인 것이고 History Notes는 해당 MeSH 용어가 과거 어떠한 이름을 가졌고 어떻게 변화했는지에 대한 자료를 제시하는 것이다. 또한 Online Note는 검색하는 사람을 위해서 필요한 사항을 기록하고 있다. Scope Note는 indexer가 색인작업을 할 때 참고하거나 검색자가 그 뜻을 알 수 있도록 해당 용어에 대한 개략적인 설명을 한 것이다. 표 3은 Annotated MeSH로 Management Information Systems에 대해 MeSH browser로 검색한 결과이다.

표 3. Annotated MeSH의 예

MeSH Heading	Management Information Systems
Tree Number	L01.700.508.300.680
Tree Number	N04.452.515
Annotation	**for automated systems do not coord with AUTOMATIC DATA PROCESSING**
Scope Note	Systems designed to provide information primarily concerned with the administrative functions associated with the provision and utilization of services; also includes program planning, etc.
Allowable Qualifiers	N04.452.515
Previous Indexing	Information Systems (1966-1982)
Previous Indexing	Organization and Administration (1968-1982)
History Note	87(83); was see under INFORMATION SYSTEMS 1983-86
Unique ID	D008331

4) Check tags

Check tags(그림 3)는 특정 관심집단에게 유용한 정보를 제공하기 위해 부여한다. 여기서 특정집단이란 임상가, 실험연구자, 약물문헌 사용자 등을 말한다. check tag는 모두 MeSH 용어이며 indexers가 논문의 특성에 따라서 부여한다. check tag에는 다음과 같은 것이 있으며 검색 시에는 limit option으로 많이 쓰인다.

Check Tags의 예

· Human	· Animal
· Female, Male	· Age of Human
· Pregnancy	· Special Animal
· In VITRO	· CASE Report
· Comparative Study	· History 등

Check Tags Cording Sheet

A ☐ PREGN	J ☐ CATS	V ☐ HUMAN	f ☐ 15th CENT
B ☐ INF NEW(to 1mo)	K ☐ CATTLE	W ☐ MALE	g ☐ 16th CENT
C ☐ INF(1-21mo)	L ☐ CHICK EMBRYO	X ☐ FEMALE	h ☐ 17th CENT
D ☐ CHILD PRE(2-5)	M ☐ DOG	Y ☐ IN VITRO	i ☐ 18th CENT
E ☐ CHILD(6-12)	O ☐ GUINEA PIG	Z ☐ CASE REPT	j ☐ 19th CENT
F ☐ ADOLESC(13-18)	P ☐ HAMSTERS	b ☐ COMP STUDY	k ☐ 20th CENT
G ☐ ADULT(19-44)	Q ☐ MICE	c ☐ ACIENT	l ☐ NIH/PHS SUP
H ☐ MID AGE(45-64)	S ☐ RABBITS	d ☐ MEDIEVAL	m ☐OTHER US GOVT SUP
I ☐ AGED(65+)	T ☐ RATS	e ☐ MODERN	n ☐ NON-US GOVT SUP
	U ☐ ANIMAL		

그림 3. Check Tags의 예 및 Cording Sheet

5) Subheadings

MeSH는 80여개의 Subheadings(부주제어)라고 불리는 용어를 포함한다. Subheadings는 각 MeSH 용어의 여러 측면 즉 치료, 수술, 역학, 원인 등에 대하여 좁혀서 기술하는 용어이다. MeSH Subheadings의 정확한 사용은 검색자들이 검색능력을 향상시킬수 있도록 도와준다. Subheadings를 이용해서 색인 혹은 검색하게 되면 해당 주제를 더욱 더 좁혀서 나타내게 할 수 있다. 미국의학도서관의 공식 명칭은 'Qualifiers'이며, 약어로도 많이 쓰인다. Subheadings는 일반적으로 MeSH 용어와 같이 사용되어, 해당 주제를 더욱 더 좁혀서 검색하게할 수 있다. 대부분의 MEDLINE 검색시스템은 subheadings를 통한 검색을 제공하여, 좀 더 구체적이며 한정적인 검색을 도와준다. 표 4는 subheadings의 전체 목록이다. Subheadings도 비슷한 것끼리 같이 집단을 이루어 트리구조를 이루고 있으며 이러한 이유로 확장(explosion)이 가능하다(표 5). Subheadings는 주로 main heading/ *subheading의 형식으로 표현되며(예: Internet/ *standards), 해당 MeSH 용어에 대해 허용되는 subheadings 표 5처럼 제한되어 있다(National Library of Medicine, 1997).

표 4. MeSH Subheadings

Abnormalities AB	Manpower MA
Administration and Dosage AD	Metabolism ME
Adverse Effects AE	Methods MT
Agonists AG	Microbiology MI
Analogs and Derivatives AA	Mortality MO
Analysis AN	Nursing NU
Anatomy and Histology AH	Organization and Administration OG
Antagonists and Inhibitors AI	Parasitology PS
Biosynthesis BI	Pathogenicity PY
Blood Supply BS	Pathology PA
Blood BL	Pharmacokinetics PK
Cerebrospinal Fluid CF	Pharmacology PD
Chemical Synthesis CS	Physiology PH
Chemically Induced CI	Physiopathology PP
Chemistry CH	Poisoning PO
Classification CL	Prevention and Control PC
Complications CO	Psychology PX
Congenital CN	Radiation Effects RE
Contraindications CT	Radiography RA
Cytology CY	Radionuclide Imaging RI
Deficiency DF	Radiotherapy RT
Diagnosis DI	Rehabilitation RH
Diagnostic Use DU	Secondary SC
Diet Therapy DH	Secretion SE
Drug Effects DE	Standards ST
Drug Therapy DT	Statistics and Numerical Data SN
Economics EC	Supply and Distribution SD
Education ED	Surgery SU
Embryology EM	Therapeutic Use TU
Enzymology EN	Therapy TH
Epidemiology EP	Toxicity TO
Ethnology EH	Transmission TM
Etiology ET	Transplantation TR
Genetics GE	Trends TD
Growth and Development GD	Ultrasonography US
History HI	Ultrastructure UL
Immunology IM	Urine UR
Injuries IN	Utilization UT
Innervation IR	Veterinary VE
Instrumentation IS	Virology VI
Isolation and Purification IP	
Legislation and Jurisprudence LJ	

표 5. Families of Subheading Explosions

adverse effects
- poisoning
- toxicity

analysis
- blood
- cerebrospinal fluid
- isolation and purification
- urine

anatomy and histology
- blood supply
- cytology
  - pathology
  - ultrastructure
- embryology
  - abnormalities
- innervation

chemistry
- agonists
- analogs and derivatives
- antagonists and inhibitors
- chemical synthesis

complications
- secondary

cytology
- pathology
- ultrastructure

diagnosis
- pathology
- radiography
- radionuclide imaging
- ultrasonography

embryology
- abnormalities

epidemiology
- ethnology
- mortality

etiology
- chemically induced
- complications
- secondary
- congenital
- embryology
- genetics
- immunology
- microbiology
- virology
- parasitology
- transmission

metabolism
- biosynthesis blood
- cerebrospinal fluid
- deficiency
- enzymology
- pharmacokinetics urine

microbiology
- virology

organization and administration
- economics
- legislation and jurisprudence
- manpower
- standards
- supply and distribution
- trends
- utilization

pharmacology
- administration and dosage
  - adverse effects
  - poisoning
  - toxicity
- agonists
- antagonists and inhibitors
- contraindications
- diagnostic use
- pharmacokinetics physiology
- genetics
- growth and development
- immunology
  - metabolism
  - biosynthesis
  - blood
  - cerebrospinal fluid
  - deficiency
  - enzymology
  - pharmacokinetics urine

physiopathology
- secretion

statistics and numerical data
- epidemiology
  - ethnology
  - mortality
- supply and distribution
- utilization

surgery
- transplantation

therapeutic use
- administration and dosage
- adverse effects
- contraindications
- poisoning

therapy
- diet therapy
- drug therapy
- nursing
- prevention and control
- radiotherapy
- rehabilitation
- surgery transplantation

6) Publication Types Journal Article

Publication Type(PT)으로 색인을 해 두면 나중에 좁혀서 검색할 때 많은 도움이 된다. publication type는 자체가 MeSH 용어로 되고 limit option에 이용된다. 공식적인 PT에는 그림 7과 같은 것들이 있다.

표 6. Publication Types

· Addresses	· Bibliography
· Biography	· Classical Article
· Clinical Conference	· Clinical Trial
· Clinical Trial, Phase I	· Clinical Trial, Phase II
· Clinical Trial, Phase III	· Clinical Trial, Phase IV
· Congresses	· Controlled Clinical Trial
· Randomized Controlled Trial	· Comment
· Consensus Development Conference	· Consensus Development Conference, NIH
· Corrected and Republished Article	· Dictionary
· Directory	· Duplicate Publication
· Editorial	· Festschrift
· Guideline	· Historical Article
· Interview	· Journal Article
· Lectures	· Legal Cases
· Letter	· Meeting Abstract
· Meta-Analysis	· Multicenter Study
· News	· Newsletter Article
· Overall	· Periodical Index
· Practice Guideline	· Published Erratum
· Retracted Publication	· Retraction of Publication
· Review	· Review, Academic
· Review, Multicase	· Review of Reported Cases
· Review Literature	· Review, Tutorial
· Scientific Integrity Review	· Technical Report
· Twin Study	

2. TEXT-Based 검색과 MeSH-Based 검색의 비교

Free-text 검색은 제목이나 초록에서 일부분의 단어를 이용한 비 통제된 용어이다. 이것은 어떤 표준이나 협정없이 저자가 지정하는 것을 의미하므로 별도의 노력이 필요치 않다. 따라서 검색자들은 저자와 검색자가 생각한 중요단어의 개념이 달라서 부적절한 검색을 하더라도, MeSH 용어를 이용한 검색보다 free-text 검색을 사용한다. 그러나 free-text 검색을 사용하면 검색의 민감도가 떨어지게 된다. 예를 들어 검색자가 Hyperlipedemia를 가지고 검색을 했을 때, 저자가 좀더 좁은 개념인 Hypercholesterolemia를 사용한다면 단지 Hyperlipedemia라는 단어를 사용하지 않았다는 이유로 많은 관련된 자료를 놓치게 된다(Lowe 등, 1994). MeSH 용어를 이용하여 Hyperlipedemia에 대해 검색을 하면 MeSH 구조가 "explosion"을 지원하므로 Hyperlipedemia, Hypercholesterolemia, Hyperlipoprpte-inemia, Hypertriglyceridemia로 저자가 사용한 모든것을 검색한다.

가. 정확도(Precision)와 회귀도(Recall)

- The Sensitivity and Specificity of Searching

MeSH등의 인덱싱용어를 이용한 검색에 비하여 free-text검색은 recall rate가 낮다(Funk ME 등, 1983). 회귀도의 정의는 아래와 같다.

$$\text{Recall} = \frac{\text{Number of relevant citations retrieved}}{\text{Number of relevant citations in database}}$$

용어론(terminology)적 관점에서 보면 recall은 검색의 민감도 및 정확도와 관련이 있다. 용어론(terminology)적 관점에서 정확도(precision)는 검색의 특이도(specificity)를 의미하며 검색한것들이 얼마나 관계성이 있는지를 재어보는 척도가

된다.

$$\text{Precision} = \frac{\text{Number of relevant citations retrieved}}{\text{Total Number of citations retrieved}}$$

MeSH가 Explosion을 제공함에 따라, 정확도는 free-text검색이 일반적으로 좀 더 높다. 즉, Recall을 최대로 높이면 부정확한 자료의 검색가능성도 높아진다. 따라서 MEDLINE 검색자들은 정확도(precision)와 민감도(recall)의 적절한 균형을 이룬 효율적인 배분을 해야한다. MeSH를 이용하여 정확도를 높이는 방법은 subheading등의 limit option등을 사용하는 것이다.

나. MeSH Tree Structures

MeSH는 37,862개의 용어로 이루어져 있으며, "MeSH tree structure"라는 복잡한 계층구조를 가지고 있다. 이 계층구조는 가지(branch)의 집합으로 연결되어 있으며, 이런 트리구조는 특이도(specificity)를 증대시킨다. 이런 트리구조는 MEDLINE을 검색할 때 대단히 유용한 전략으로 활용되어진다. 대부분의 검색자가 지정한 개별적인 용어를 MEDLINE이 지정한 정확한 용어로 발전시키도록 하며, 이런것들을 이용하여 부정확한 자료의 검색을 줄일수 있다. 또한 MeSH tree는 recall을 향상시켜 검색의 범위를 확장할수도 있다. 결국 이런 트리구조는 검색에 있어서 두 가지 이점을 제공한다. 하나는 MeSH로 검색을 할 때 broader(more general) 혹은 narrower(more specific)로 검색을 할 수 있다는 것이며 다른 하나는 확장(explosion)이라 불리는 검색이 가능하다는 점이다(Fremer, 1995).

다. 통제어로서의 언어학적인 문제

의학지식의 양과 복잡성이 증대되는 것처럼 언어의 지식에 대한 표현도 복잡해지고 많아진다. 각각의 전문분야는 그들만의 독특한 부주제어를 가지게 되고 그것들은 비전문가들과의 장벽이 될 수 있다. 분자생물학, 생명공학, 의료정보학 등의 기술이 발달함에 따라 각 분야에서 전문가와 비전문가사이의 언어의 벽은 점점 더 깊어질 것이다. 대규모의 생의학 데이터베이스에서 정보의 검색을 위해서는 검색자들이 대상영역에 적당한 언어의 표현을 통한 질의가 반드시 필요하다. 즉 검색자 자신만의 용어를 적당한 영역의 용어로 변환시키는 것이 정보검색의 기본적인 문제이다. MeSH는 MEDLINE에서 규정하고 있는 언어이다. 이것을 전부 습득하기는 매우어렵다. 인쇄된 참조페이지들의 양이 2,300페이지 이상이며 무게는 약 5.5kg에 이른다. 결국 MeSH가 제공하는 entry term이 있다해도 경험이 없고, 자주 MEDLINE을 이용하지않는 검색자들에게는 적당한 MeSH용어를 검색하는것은 여전히 어려울 것이다. MeSH와 같은 통제어를 사용시에 기본적인 어려움은 개념을 표현하는 정확한 용어를 찾아내는 것이다. 많은 관련된 개념은 하나의 표준 MeSH 용어로 연결되어진다(Mitchell 등, 1992).

MEDLINE 검색시 정확한 표준 MeSH 용어를 찾아내지 못하는 문제를 해결하는 방법으로 MeSH는 거대한 규모의 entry term의 집합을 가지고 있다. 이러한 entry term은 비록 MeSH 용어가 아니더라도 검색자가 일반적으로 생의학에서 사용되는 용어로 MEDLINE의 검색이 가능토록 해준다. 그러나 entry term도 MeSH 용어 단어자체의 오류를 보정해주거나, 단·복수 혼용등의 방법적 미숙등을 보충해주는 역할을 중점적으로 수행하고 있으므로 실제로 의미론적인 문제를 해결하는 방안은 아니다.

3. 의학주제어 색인 방법

가. Coordination 개요

Coordination이란 두 개 이상의 색인어를 적절히 조합하여 쓰는 것을 말하며, 표 7과 같은 형태가 있다.

표 7. Coordination의 형태 및 예

형 태	예
1. Main Heading + Main heading	* Liver * Pneumonia
2. Main Heading + Check tag	* hepatitis child
3. Main Heading + Subheadings	Hepatitis / * Prev
4. Pre-Coordinated main heading a. two main headings originally b. Main heading + Check tag c. Main heading + Subheadings	 *Liver glycogen = *liver + *glycogen *Schizophrenia, child = * Schizophrenia child *communicable disease control = communicable disease /* prev

앞의 예에서 보는 바와 같이 Pre-coordinated heading은 원래는 두 개 이상의 주제어이었지만 문헌에서 동반되는 경우가 많아서 하나로 결합되어 있는 것을 말하며 표 8과 같은 형태가 흔하다.

표 8. Precoordination의 형태 및 예

형 태	예
an organ + disease	Stomach diseases
an organ + neoplasm	Stomach neoplasms
an organism + infection	Staphylococcal infections
an animal + disease	Dog diseases
a disease + a site	Hypertension, Portal

나. 일반적인 원칙

MeSH 색인의 기본적인 원칙은 가능하면 specific하게 한다는 것이다. 예를 들어 spectinomycin에 대한 논문은 spectinomycin으로 색인 해야 나중에 spectinomycin으로 찾을 수 있다. 만일 검색하는 사람이 모든 항생제에 대해 찾는 것을 원하면 antibiotics로 색인하지 않았더라도 antibiotics로 검색하여 explosion 시키면 이 논문도 같이 검색이 된다. 이런 원칙으로 다음과 같이 색인 한다.

1) 부주제어(subheadings)로는 색인하지 않는다 (예: glucose-metabolism에 대한 논문은 Metabolism으로 색인하지 않는다).
2) 가능한 가장 특이한 주제어를 선택한다(예: neomycin에 대한 논문은 antibiotics로 색인하지 않는다).
3) 나이에 따른 질환, 생리현상에 대한 논문은 나이로 색인하지 않는다(예 : infant gout에 대한 논문은 infant로 색인하지 않고 check tag를 부여한다).
4) 실험동물을 사용한 실험에 관한 논문을 동물로 색인하지 않고 check tag를 부여한다(예 : rat에서 유발한 arthritis에 관한 논문).
5) 실험 방법에 관한 논문의 경우 질환이나 특이물질 등의 주제어를 찾고 실험방법으로는 색인하지 않는다(gama globulin의 electrophoresis에 대한 논문은 electrophoresis로 색인하지 않는다).
6) 기관별 질환에 대한 논문은 기관-질환의 분류로 색인 한다(cecum의 질환은 cecum으로 색인하지 않고 cecal disease로 색인 한다).
7) 미생물에 의한 감염질환은 organism-infection으로 색인 한다(대장균 감염은 escherichial coli infection으로 escherichial coli로 색인하지 않는다).
8) 일반적인 병리현상 등은 해당 기관명-pathology로 색인한다(췌장의 necrosis에 관한 논문은 pancreas-pathology로 색인하고 necrosis로 색인하지 않는다).

다. 다빈도 Coordinations

표 9의 쌍들은 대부분 다음과 같이 동시에 색인되는 경우가 많다.

표 9. 다빈도 Coordination의 예

유 형	예	유 형	예
(Disease A)/	etiology	(Disease A)/	complications
(Disease B)/	complications	(Disease B)/	complications
(Disease)/	drug therapy	(Disease)/	chemically induced
(Drug)/	therapeutic use	(Drug)/	adverse effect
(Disease)/	pathology	(Disease)/	etiology
(Organ)/	pathology	(Technique)/	adverse effect
(Disease)/	microbiology	(Enzyme)/	metabolism
(Organ)/	microbiology	(Organ)/	enzymology
(Organism)/	isolation	(Disease)/	enzymology
(Organ)/	drug effects	(Organism)/	drug effects
(Drug)/	pharmacology	(Drug)/	pharmacology
(Organ)/	metabolism	(Organism)/	metabolism
(Drug)/	metabolism	(Drug)/	metabolism
(Organ)/	chemistry	(Disease)/	diagnosis
(Drug)/	analysis	(Drug)/	diagnostic use
(Organ)/	metabolism	(Organ)/	radiation effects
(Drug)/	pharmacokinetics	(Drug)/	specific radiation
(Disease)/	metabolism	(Disease)/	metabolism
(Organ)/	metabolism	(Organ)/	chemistry
(Drug)/	metabolism	(Drug)/	analysis

라. 질환약제의 MeSH indexing

질환은 다음과 같은 형태로 색인 한다.

1) Precoordinated organ-disease term으로 한다(예: Brain diseases, skin diseases).

2) 감염의 경우 organism-diseases term으로 한다(예: salmonella infections) 대부분 organ-disease과 동반되어 색인 한다(예: salmonella infection;

liver diseases).

3) precoordinated organ-organism-disease term으로 한다(예: Tuberculosis, Renal).

4) specific organ + precoordinated more general organ-disease term으로 한다(예: ciliary body disease에 대해서는 ciliary body, Uveal diseases).

5) 특정 질병 이름을 그대로 쓴다(예: Kidney failure, Acute; aganulocytosis)

6) Temporal lobe disease처럼 organ-disease가 MeSH에 없는 경우에는 precoordinated term이 있을 때까지 tree를 거슬러 올라간다(예: drug therapy of temporal lobe의 경우 temporal lobe; brain diseases/ drug ther로 한다).

7) Temporal lobe Neoplasm처럼 organ-neoplasm이 MeSH에 없는 경우에는 Temporal lobe ; Brain Neoplasm으로 색인 한다.

8) 만일 논문이 Drug therapy of Citrobacter Cellulitis인 경우 Citrobacter, Enterobacteriaceae infections/ drug ther, Cellulitis/ drug ther 로 색인한다.

9) Neoplasm의 경우 조직형과 위치를 색인 한다. 예를 들어 Basal cell carcinoma of the skin의 경우 carcinoma, basal cell; skin neoplasm으로 색인한다.

10) 약제의 경우 해당질병, pharmacologic activity를 같이 색인 한다(예: Lithium carbonate in the treatment of manic disorders의 경우 lithium carbonate / ther use ; Manic disorder / drug ther ; antimaniac agents / ther use로 색인 한다).

Ⅲ. 연구 방법

1. 연구의 틀

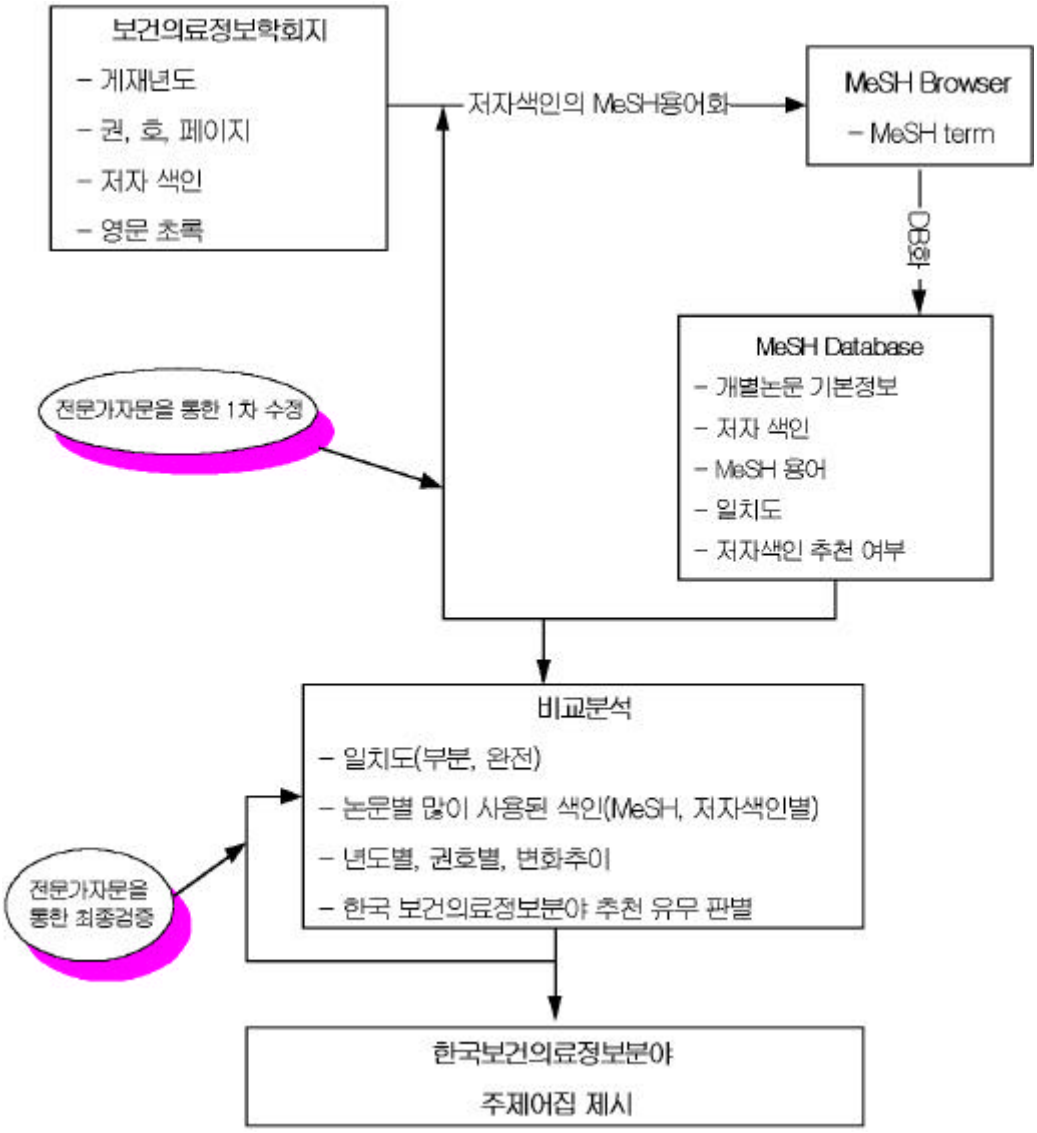

그림 4. 연구의 틀

2. 연구 대상

1995년 1권 1호부터 2000년 제 6권 4호 까지 6년동안 대한의료정보학회지에 게재된 논문의 영문주제어를 연구대상으로 영문주제어의 개수와 다사용 주제어를 조사하고 이러한 주제어가 얼마나 MeSH와 일치하는지를 MeSH browser를 이용하여 분석한다.

3. 분석 방법

가. 분석의 틀

1) 기술적 분석

대한의료정보학회지에 게재된 논문을 대상으로 저자가 사용한 영문주제어를 수집·분석하여 몇 개의 용어가 사용되는지, 저자가 가장 많이 사용한 용어는 무엇인지 분석한다.

2) MeSH 사용여부

저자가 사용한 개별 영문주제어를 NLM에서 제공하는 MeSH browser를 이용하여 분석한다. 각 주제어를 MeSH 용어, MeSH 용어와 사소한 차이가 있는 것(단수와 복수 혼동, 어순이 틀린 것), MeSH 용어로 볼수 없는 것의 세 가지로 구분하여 각각의 분율을 구한다.

3) 용어선택의 정확도

1995년 1권 1호부터 2000년 제 6권 4호까지의 원저는 172편이다. 원저의 영문초록에 실린 중심단어를 참조하여 영문초록을 읽고 가장 적절하다고 생각되는 중심단어를 MeSH 용어를 기준으로 하여 기존 중심단어의 정확도를 평가한다. 중심단어는 가능한 3개이상을 선정하도록 하며 만일 적절한 중심단어로 추천이

되었으나 MeSH browser에 나와 있지 않은 용어는 따로 구분하여 보건의료정보 분야에서 필요한 용어로 추천하고자 별도의 표시를 한다.

나. 수집 및 분석을 위한 데이터베이스 구축

분석을 위한 기초자료수집을 위한 데이터베이스를 구축한다. 데이터베이스는 논문의 기본정보 및 영문주제어를 변환된 MeSH 용어정보등을 포함하며 데이터베이스의 개요는 다음과 같다.

1) **데이터베이스 테이블 관계도**(Table Relationships)

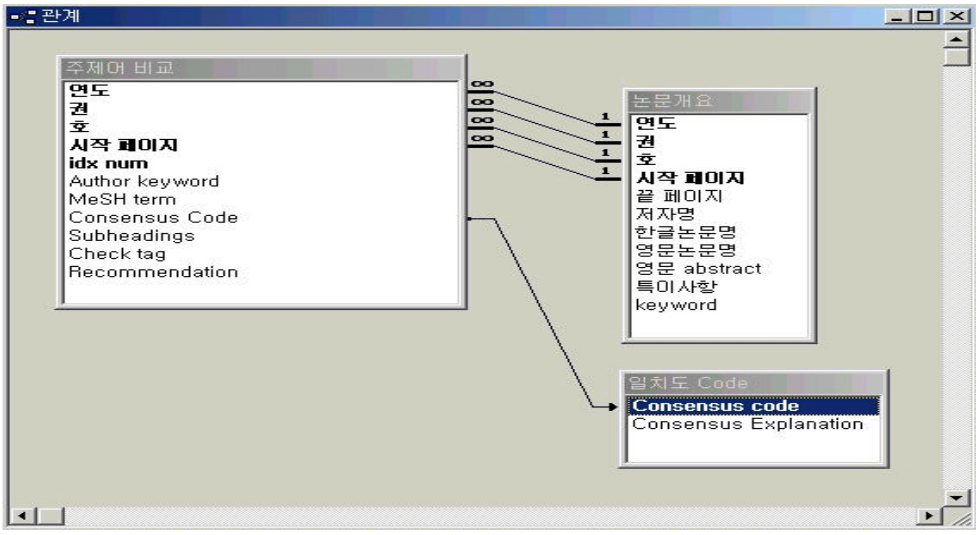

그림 5. 데이터베이스 테이블 관계도

2) **테이블 구성**

수집 및 분석을 위한 테이블은 크게 두 부분, 일곱 개의 테이블로 구성된다. 첫번째 구성부분은 대한의료정보학회지 관련정보를 수집한 데이블로 개별논문의 발행년도, 권, 호, 논문명, 초록등을 포함한 논문개요 테이블과 저자용어와

MeSH 용어의 변환과정, 상관관계 및 일치정도를 포함하는 주제어 비교 테이블과 주제어비교 테이블에서 숫자화하여 표현된 일치정도에 대한 자세한 설명을 담고 있는 테이블인 일치도 코드 테이블로 구성된다. 두번째 구성부분은 MeSH를 이용한 대한의료정보학회지와 MeSH 용어의 상관성 및 주제어 제시를 위해 반드시 필요한 MeSH에 대한 정보를 담고 있는 테이블다. 이때 필요한 테이블은 크게 네가지로 첫째 MeSH 용어의 정보 용어의 Tree number를 가지고 있는 주 테이블인 Descriptor 테이블, 부정확한 MeSH용어 사용의 보정을 위한 Entry term 테이블, MeSH용어중 특별히 분류된 흔히 Subheadings로 불리나 NLM공식 명칭은 Qualifier인 Qualifier 테이블, 용어의 기술적인 부분을 포함하고 있는 Check Tag 테이블로 구성되어 있으며 각 테이블의 자세한 구성은 아래와 같다.

가) 논문개요 테이블

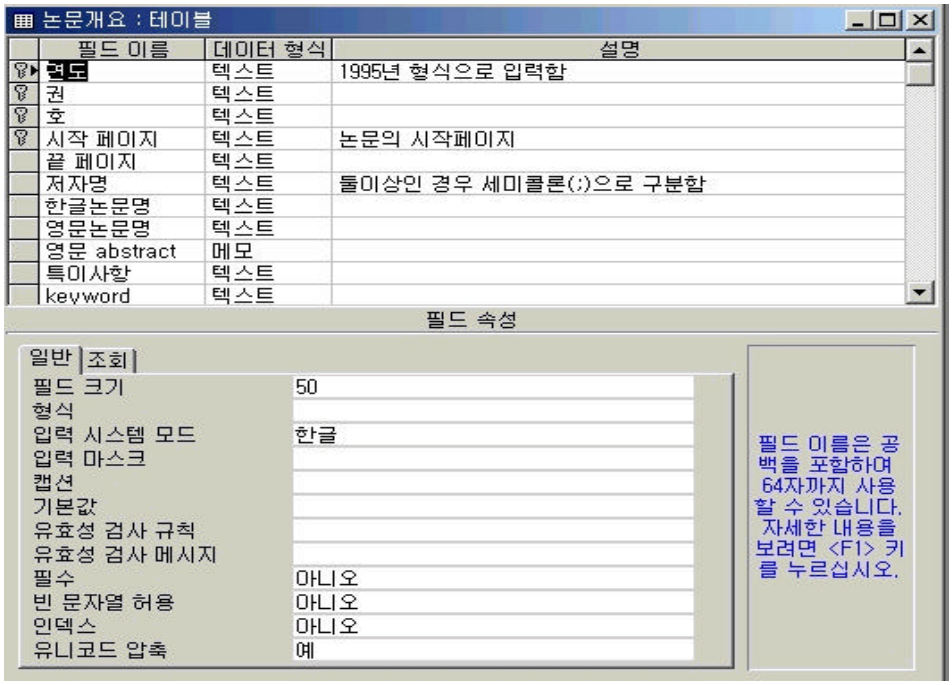

그림 6. 논문개요 테이블 구조 및 필드 설명

나) 주제어 비교 테이블

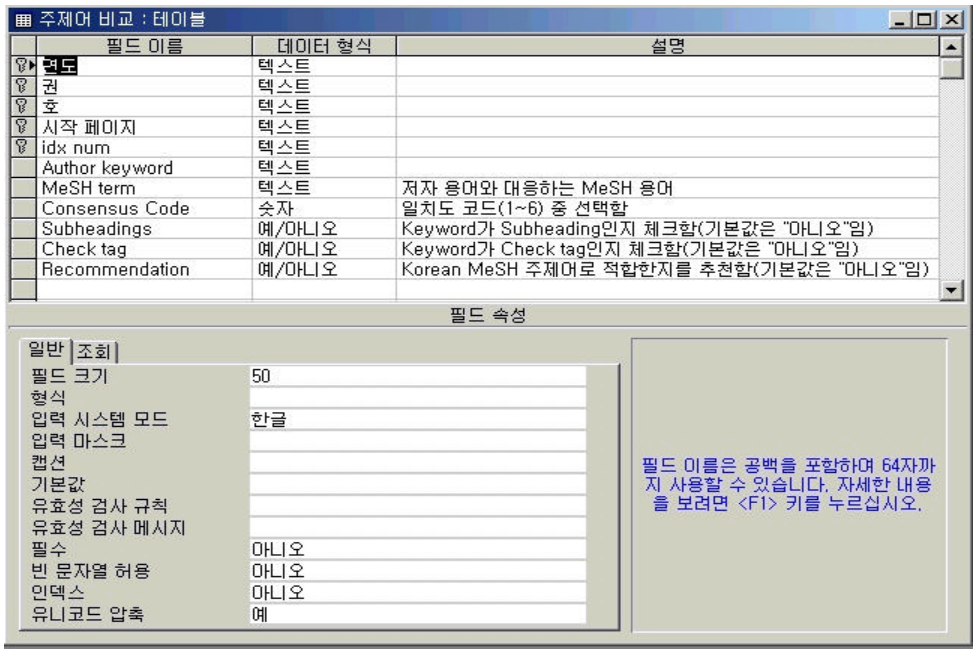

그림 7. 주제어 비교 테이블 구조 및 필드 설명

다) 일치도 코드 테이블

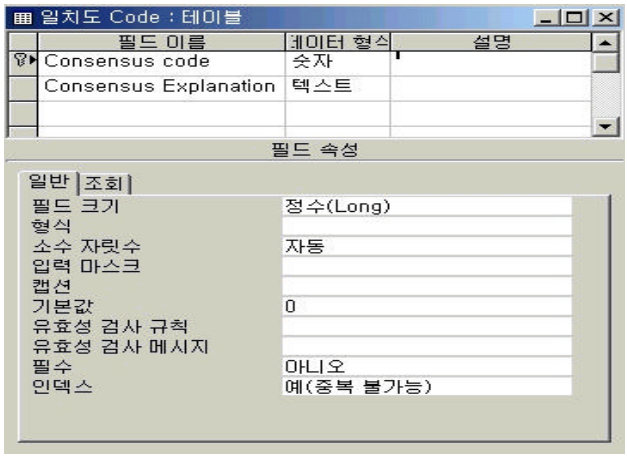

그림 8. 일치도 코드 테이블 구조 및 필드

라) Descriptor **테이블**

Descriptor : 테이블

필드 이름	데이터 형식	설명
MeSH Heading	텍스트	
Tree Number	텍스트	

필드 속성

일반 | 조회

필드 크기	255
형식	
입력 시스템 모드	영문
입력 마스크	
캡션	
기본값	
유효성 검사 규칙	
유효성 검사 메시지	
필수	아니오
빈 문자열 허용	아니오
인덱스	아니오
유니코드 압축	예

그림 9. Descriptor **테이블 구조 및 필드 설명**

마) Entry term **테이블**

Entry Term : 테이블

필드 이름	데이터 형식	설명
MeSH Heading	텍스트	
Entry Term	텍스트	

필드 속성

일반 | 조회

필드 크기	255
형식	
입력 시스템 모드	영문
입력 마스크	
캡션	
기본값	
유효성 검사 규칙	
유효성 검사 메시지	
필수	아니오
빈 문자열 허용	아니오
인덱스	아니오
유니코드 압축	예

그림 10. Entry term **테이블 구조 및 필드 설명**

바) Qualifier 테이블

Qualifier : 테이블

필드 이름	데이터 형식	설명
SH	텍스트	
QE	텍스트	
QA	텍스트	
QT	텍스트	
MS	메모	
GM	텍스트	
AN	텍스트	
OL	텍스트	

필드 속성

일반 | 조회

필드 크기	255
형식	
입력 시스템 모드	영문
입력 마스크	
캡션	
기본값	
유효성 검사 규칙	
유효성 검사 메시지	
필수	예
빈 문자열 허용	아니오
인덱스	예(중복 불가능)
유니코드 압축	예

필드 이름은 공백을 포함하여 64자까지 사용할 수 있습니다. 자세한 내용을 보려면 <F1> 키를 누르십시오.

그림 11. Qualifier **테이블 구조 및 필드 설명**

사) Check Tag 테이블

Check Tag : 테이블

필드 이름	데이터 형식	설명
CT	텍스트	

필드 속성

일반 | 조회

필드 크기	255
형식	
입력 시스템 모드	영문
입력 마스크	
캡션	
기본값	
유효성 검사 규칙	
유효성 검사 메시지	
필수	예
빈 문자열 허용	아니오
인덱스	예(중복 불가능)

그림 12. Check Tag **테이블 구조 및 필드 설명**

3) 일치도 코드 구성

일치도 코드 테이블에 저장된 개본적인 코드의 구성은 아래(표 10)와 같다.

표 10. 일치도 코드 및 설명

대분류	일치도 코드	일 치 도 설 명
완전일치	1	완전일치
부분일치	2-1	유형 1 Entry term 사용
	2-2	유형 2 단수/ 복수 오류
	2-3	유형 4 Subheadings 사용
	2-4	유형 4 전후방 일치(한단어 이상일치)
	2-5	유형 5 약어 사용
불일치	3	불일치 - 저자용어 사용

4) 입력 화면 구성

기본 화면 구성은 크게 두 부분으로 나뉘어진다. 개별 논문의 정보를 포함한 상위 부분과 각 논문별로 일정하게 정해지지 않은 저자용어 및 MeSH용어 정보를 담기위한 아래의 스프레드 쉬트가 그것이다(그림 13). 기본정보 입력화면에서 논문 정보를 입력한후 MeSH browser를 이용하여 그에 해당하는 MeSH용어를 변환한 정보를 입력한다.

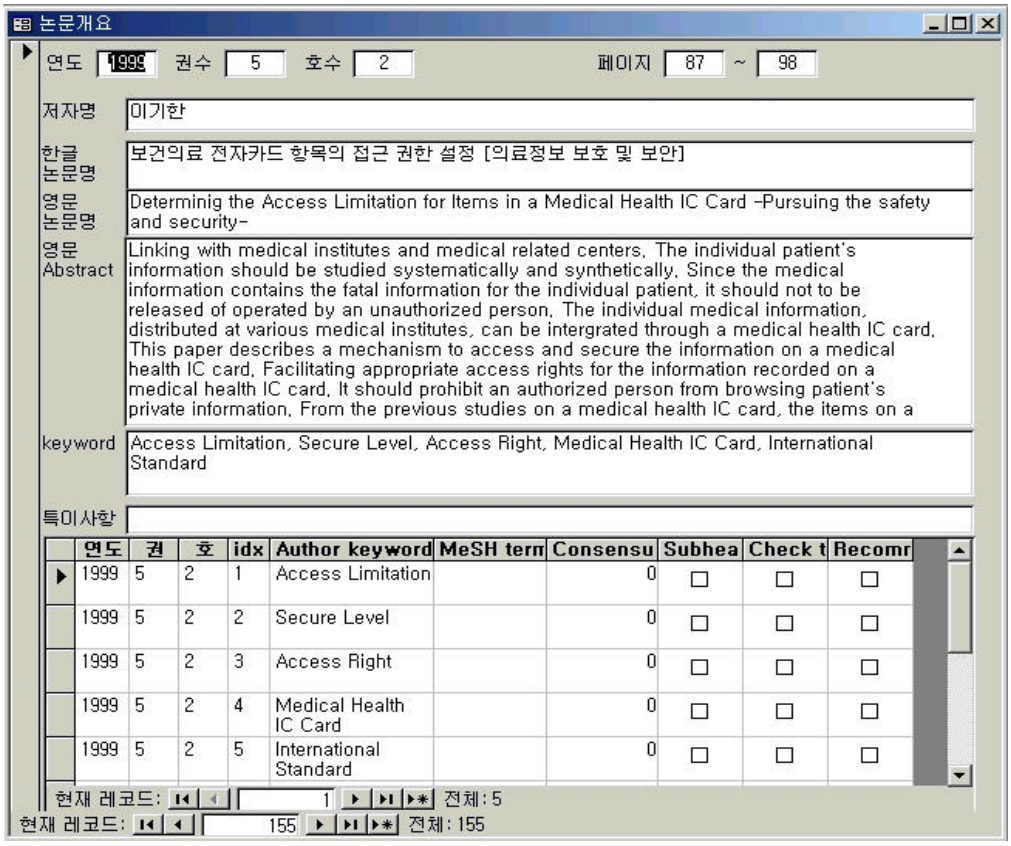

그림 13. 기본정보 입력화면

다. 저자용어를 MeSH 용어로 부여하기

현실적으로 indexer를 두어 모든 논문에 MeSH를 부여하는 것이 힘들기 때문에 저자용어를 MeSH로 부여하는 것이 차선책이라 할수 있다. 그러나 실제로 저자 용어들 중의 상당수는 MeSH 용어로 볼 수 없는 경우가 많다. 저자 용어를 MeSH 용어로 사용하려고 할 때는 몇 가지 사항에 대한 합의가 있어야 한다. 첫째는 check tag를 부여할 것인가이다. 앞에서 언급한 바와 같이 check tag는 연령, 성별, 실험동물처럼 특정 관심집단을 위하여 색인 하는 것이다. 이 경우 이런 check tag를 용어로 부여하는 것이 힘든 경우가 있기 때문에 이런 경우 어떻게

처리할 지에 대해서 합의가 있어야 한다. 둘째는 subheadings로 이것은 해당 주제어의 특정측면을 표시할 수 있는데 이런 subheadings를 주제어로 표시할 지를 결정해야 한다. 셋째는 publication type을 어떻게 처리할 지를 결정해야 한다. publication type는 검색 시에 아주 유용하게 쓰일 수 있기 때문에 색인이 필요한데 이를 주제어로 부여할 지에 대해 합의해야 한다.

가장 현실적인 방법은 저자가 부여한 주제어를 MeSH browser로 찾아서 적절한 MeSH 용어로 바꾸어 주는 것이다. MeSH browser는 PubMed, Internet grateful Med같은 인터넷상에서 검색이 가능하기 때문에 저자가 부여한 주제어로 검색을 해 보면 해당 MeSH 용어를 찾아낼 수 있다.

표 11. MeSH를 이용한 Indexing Criteria

- 주제어를 MeSH browser로 찾아서 적절한 MeSH 용어로 변환
- 변환된 MeSH 용어가 Check tag인지, Subheading인지를 표시함
- 주제어와 별도로 Check tag와 Subheading을 부여하지 않음
- Publication Types(PT)은 모두 Journal Article이므로 부여하지 않음

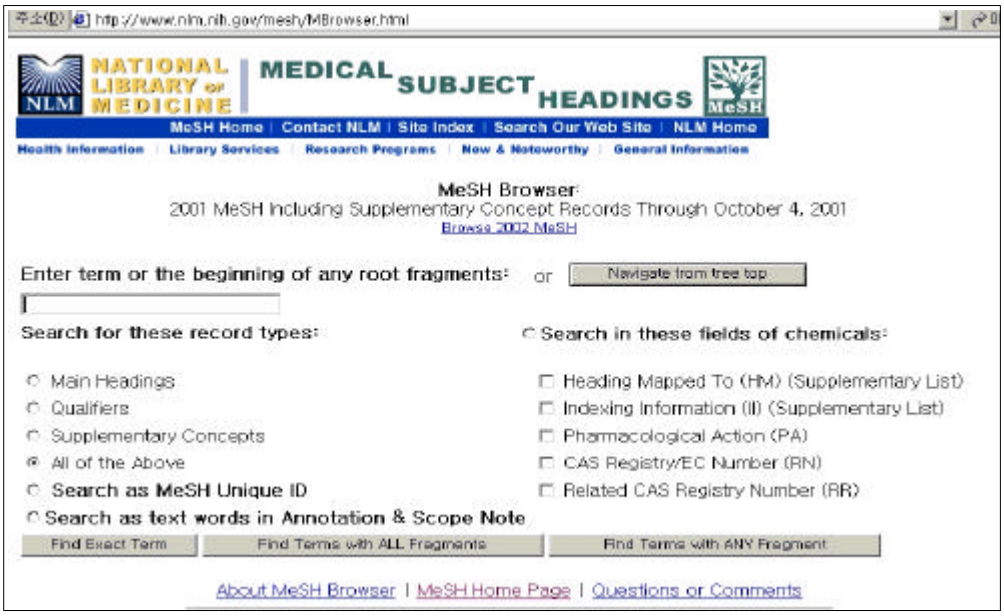

그림 14. MeSH Browser Home Page

그림 15. PubMed 제공 MeSH Browser

Ⅳ. 연구결과

1. 대한의료정보학회지 연도별 논문수 및 용어수 분석

데이터베이스에 수집된 자료를 기초로 대한의료정보학회지 1권 1호(1995년)부터 6권 4호(2000년)까지의 논문 및 저자용어를 분석한 결과는 표 12와 같다. 1995년에서 2000년까지의 총 논문수는 172편이었으며 이중 153편에 저자용어(Keyword)가 부여되었으며 19편이 저자용어가 부여되지 않았다. 이중 1995년에는 1편이 저자용어가 부여된 것에 비하여 2000년에 저자용어가 부여되지 않은 논문은 1편으로 논문의 형식면에서 많은 개선이 있었음을 알 수 있다.

표 12. 저자용어가 부여된 논문과 부여되지 않은 논문의 비교

연도	권	호	총 논문편수	저자용어가 부여된 논문	저자용어가 부여되지 않은 논문
1995	1	1	7	1	6
1996	2	1	8	8	0
1996	2	2	7	5	2
1997	3	1	8	7	1
1997	3	2	18	16	2
1998	4	1	16	16	0
1998	4	2	20	20	0
1999	5	1	15	15	0
1999	5	2	10	7	3
1999	5	3	21	17	4
2000	6	1	7	7	0
2000	6	2	8	8	0
2000	6	3	10	10	0
2000	6	4	17	16	1
Total			**172**	**153**	**19**

그림 16은 연도별 대한의료정보학회지 논문 수를 보여주고 있다. 1995년에 7편에서 시작한 논문이 지속적으로 증가하여 1999년 46편, 2000년 42편으로 양적으로도 많은 발전이 있었음을 보여준다.

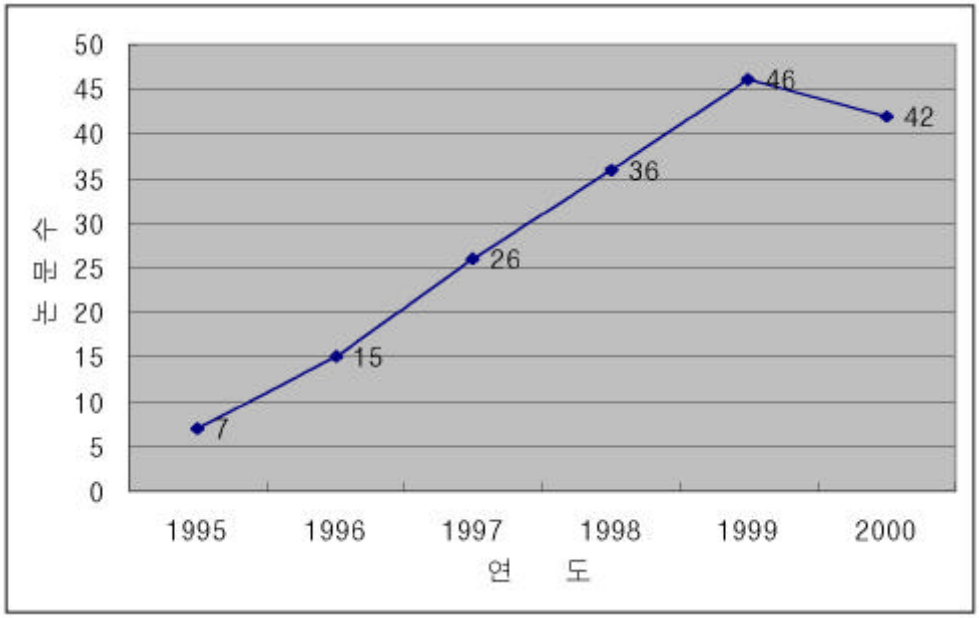

그림 16. 대한의료정보학회지 연도별 논문수

표 13 및 그림 17은 1995년에서부터 2000년까지의 대한의료정보학회지에 부여된 저자용어수 및 평균 저자용어수를 보여주고 있다. 대한의료정보학회지 부여된 저자용어수를 살펴보면 총 153편의 논문에 546개의 저자용어가 부여되어 있어, 평균 3.57개의 영문주제어를 부여하고 있다. 평균 저자용어수를 살펴보면 논문편수의 증감과 무관하게 3.25 5.00개의 저자용어가 비슷하게 부여되고 있는 것을 알 수 있다. 이는 MEDLINE이 부여하는 평균 MeSH 용어수인 10개에는 크게 못미치는 개수로 각 논문에 부여된 저자용어의 수와 MeSH 용어의 수는 큰 차이가 있음을 알 수 있다.

표 13. 대한의료정보학회지 연도별 논문수 및 부여된 저자용어수

연도	권	호	논문수	부여된 저자용어수	평균 저자용어수
1995	1	1	1	5	5.00
1996	2	1	8	26	3.25
1996	2	2	5	17	3.40
1997	3	1	7	22	3.14
1997	3	2	16	56	3.50
1998	4	1	16	60	3.75
1998	4	2	20	65	3.25
1999	5	1	15	55	3.67
1999	5	2	7	23	3.29
1999	5	3	17	61	3.59
2000	6	1	7	25	3.57
2000	6	2	8	32	4.00
2000	6	3	10	35	3.50
2000	6	4	16	64	4.00
	총 계		**153**	**546**	**3.57**

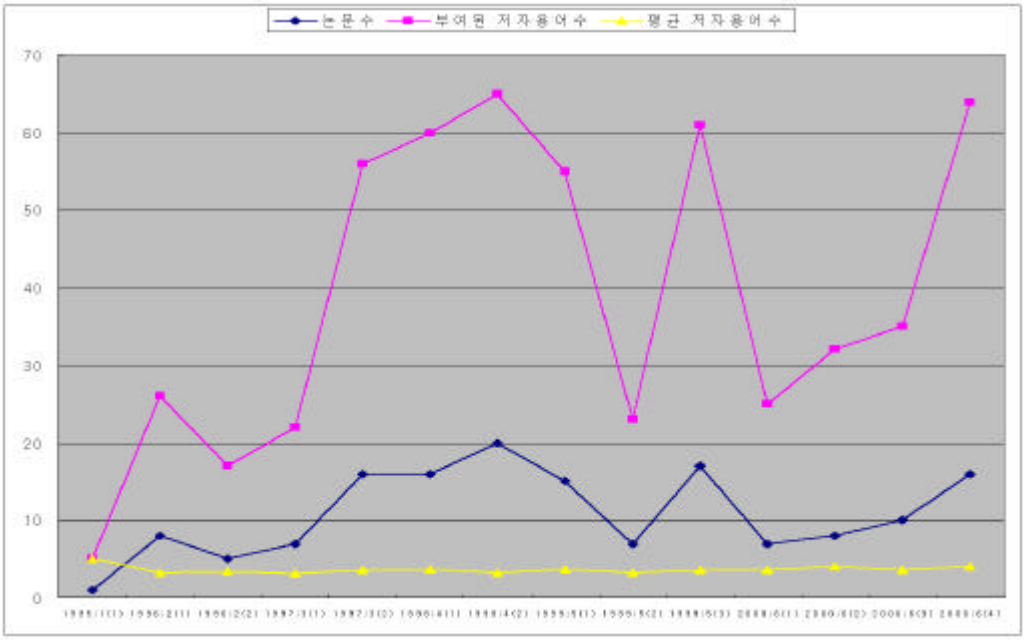

그림 17. 연도별(권호별) 논문수 및 부여된 저자용어수 추이

2. 저자용어와 MeSH 용어의 일치정도

분석대상 논문 153편의 저자용어와 MeSH용어를 비교, 분석하기 위하여 각 논문에 대하여 표 14와 같이 저자용어와 MeSH용어의 비교표를 만들고 이를 데이터베이스화 하였다. 대한의료정보학회지 논문에 부여된 저자용어는 보건연구정보센터(http:/ / www.richis.org, RICH)를 검색하여 논문별로 저장하거나 일부의 자료는 수작업을 통하여 직접 입력하였다. 동일한 논문에 부여된 MeSH용어는 저자용어를 기준으로 MeSH browser를 이용하여 검색한 후 논문단위별로 표 14의 요소들을 데이터베이스에 저장한 후 분석하여 항목별로 통계처리 하였다.

표 14. 저자용어와 MeSH 용어 비교 분석표
박준석, 민원기, 임경수, 임태환, 이정신, 곽원섭, 조한익. 1999. "웹 기반 임상병리검사정보 전달체계 구축". 대한의료정보학회지 5(1): 149-155.

저자 용어	MeSH 용어	일치도							주제어 추천
		완전 일치	부분일치[1]					불 일치	
			2-1	2-2	2-3	2-4	2-5		
Web	Internet							Y	N
Internet	Internet	Y							N
Laboratory Information System	Clinical Laboratory Information Systems		Y						N
Hospital Information System	Hospital Information Systems			**Y**					N

1) **부분일치 설명**: 2-1 Entry Term **사용**, 2-2 **단수/ 복수 오류**, 2-3 Subheadins **사용**, 2-4 **전후방 일치(한단어이상 일치)**, 2-5 **약어사용**

표 15, 그림 18 및 그림 19는 각 논문별로 저자용어와 MeSH 용어를 비교하여, 일치도에 따른 해당 용어수를 보여주고 있다. 일치율은 부분적으로 일치하는 경우 중 사소한 차이인 단/ 복수 오류(4.03%), entry term 사용(10.07%), 약어사용(0.73%)을 포함한다하더라도 25.56%에 불과하였다[2]. 저자가 MeSH 용어를 완전하게 사용하는 경우는 단지 11.72%에 불과함을 알 수 있다. 일치도 코드 2-4로 구분된 Subheading 사용의 경우 빈도가 많지 않으며 두 가지 유형에 동시에 해당하는 다중분석의 유형을 띄게 되므로 표 15에서는 따로 구분을 두지 않았으며, 본 논문에서는 Subheading에 대한 개별 통계는 사용치 않기로 한다. 대한의료정보학회지에서 나타난 Subheading 및 그 일치도 유형은 표 16과 같다.

표 15. 저자 용어와 MeSH 용어의 일치유형 분석

대분류	일치도 코드	빈도	%	일치도 설명
완전일치	1	64	11.72	완전일치
부분일치	2-1	55	10.07	유형 1 Entry term 사용
	2-2	22	4.03	유형 2 단수/ 복수 오류
	2-3	168	30.77	유형 4 전후방 일치(한단어 이상일치)
	2-5	4	0.73	유형 5 약어 사용
	소계	249	45.60%	
불일치	3	233	42.67	저자용어 사용
합계	-	546	100.00	

2) 부분일치중 전후방 일치(한단어 이상의 일치)의 경우 일치 유형으로 보기 어려운점이 많으므로 부분일치에서 제외하였다.

표 16. 대한의료정보학회지에 사용된 Subheading 및 일치도 유형

Subheading		일치도	빈도
classification	2-3	전후방 일치(한단어 이상일치)	1
classification	3	저자용어 사용	3
diagnosis	1	완전일치	1
drug therapy	2-2	단수/ 복수 오류	1
education	1	완전일치	1
methods	3	저자용어 사용	1
nursing	2-3	전후방 일치(한단어 이상일치)	2
rehabilitation	2-3	전후방 일치(한단어 이상일치)	1
계			11

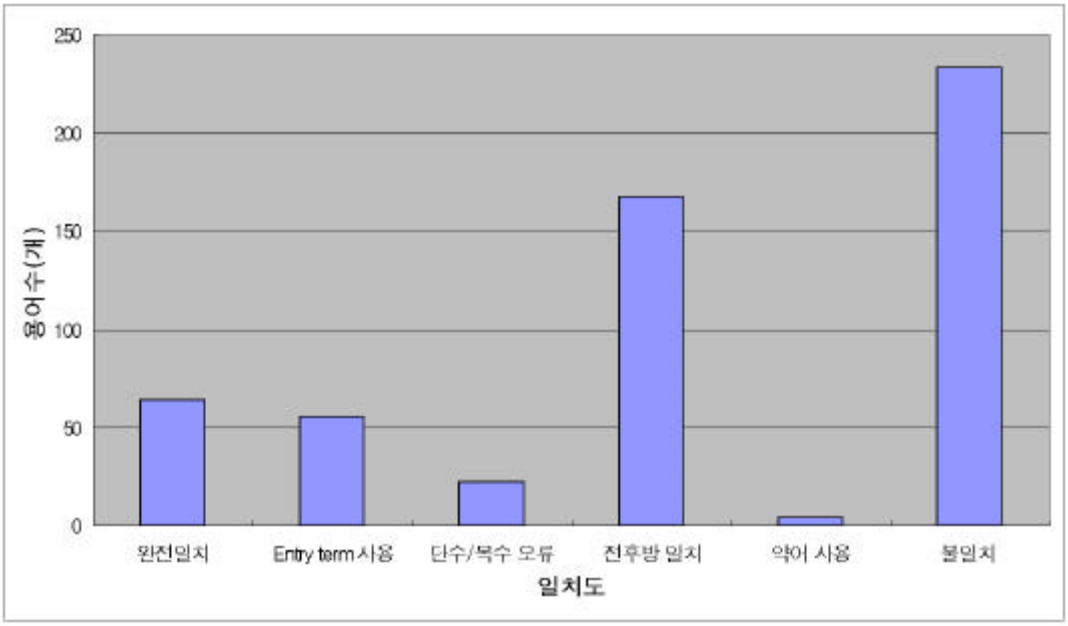

그림 18. 부여한 MeSH 용어의 일치정도(용어수)

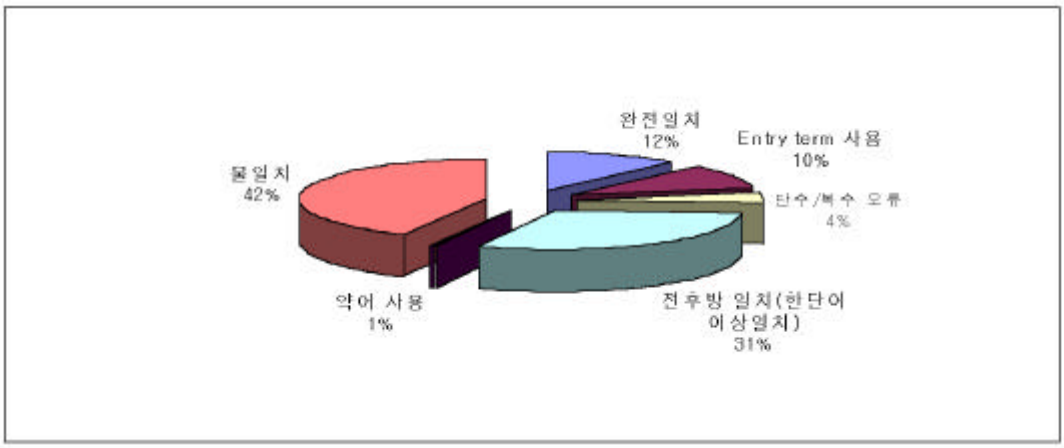

그림 19. 부여한 MeSH 용어의 일치정도(백분율)

표 17은 대한의료정보학회지 논문에 많이(4회 이상) 사용된 저자 용어를 보여주고 있다. 저자용어로 가장 많이 사용된 용어는 "Internet"이며 전체 저자용어중 2.28%를 차지하였다. 많이 사용된 용어중 "Hospital information system", "Medical Record"는 복수로 표기해야 옳은 MeSH 용어이나, 대한의료정보학회지 저자들은 대개의 경우 단수로 표기하였다. "Standardization"과 같은 용어는, 직접적인 MeSH 용어는 아니지만 그 어휘로 검색하면 MeSH 용어인 Reference Standards로 자동변환되어 검색되는 entry term이다. 저자가 키워드를 선택할 때 MeSH browser를 사용하여 확인한다면, MeSH 용어의 정확한 사용이 가능할 것이다. 세번째로 많이 사용되어진 용어인 Neural Network의 경우 Neural Network 자체만으로는 MeSH 용어도 entry term도 아니다. 올바른 MeSH 용어는 Neural Network(Computer)이다. 이처럼 다수의 저자가 사용한 용어중 MeSH 용어에 없는 용어(entry term 포함)는 Data mining, Security, Health information 등이 있다. 이는 대한의료정보학회지 저자들이 키워드 선정시 MeSH용어에 대한 기본적인 고려를 하지 않고 용어를 선택하

는 것을 보여주는 것이다. 이러한 문제점은 MeSH에 대한 이해 및 저자용어 선정시 MeSH browser를 통한 확인등을 통하여 해결이 가능할 것이다.

표 18은 논문에 부여된 저자용어를 기준으로 MeSH browser를 이용하여 검색한 후 이를 해당 MeSH용어로 변환 후, 대한의료정보학회지에 저자가 많이 사용한 MeSH 용어(5회이상)의 양상을 보여주고 있다. MeSH 용어로 가장 많이 사용된 용어 또한 저자 용어와 동일하게 "Internet"임을 볼 수 있으나 횟수가 현격히 증가하였음을 알 수 있다. 많이 사용된 용어의 수를 비교해보면 저자용어의 경우 4회이상 사용된 용어가 67개(12.27%)인데 반해, MeSH용어의 경우 188개(34.43%)로 저자용어에 비하여 용어의 통일화가 이루어져 있음을 알 수 있다. 즉, MeSH의 기본적인 통제어로서의 성질이 저자가 선택할 수 있는 용어수를 한정하므로 자연스러운 통일화가 이루어지게 되는 것이다.

표 17. 대한의료정보학회지에 많이 사용된 저자 용어(4회이상)

저자용어	횟수	%	누적 %	MeSH 용어
Internet	13	2.28	2.28	Yes
Hospital information system	8	1.47	3.85	Entry Term[3]
Neural Network	7	1.28	5.13	
Standardization	7	1.28	6.41	Entry Term[4]
Telemedicine	6	1.10	7.51	Yes
Multimedia	5	0.92	8.42	Yes
Data mining	5	0.92	9.34	
Medical record	4	0.73	10.07	Entry Term[5]
Security	4	0.73	10.81	
Health information	4	0.73	11.54	
Nursing Diagnosis	4	0.73	12.27	Yes
소 계	67	12.27	24.54	
All others(4회미만)	479	87.73	16.07	

3) MeSH Term "Hospital information systems"
4) MeSH Term "Reference Standards"
5) MeSH Term "Medical Records"

표 18. 대한의료정보학회지에 많이 사용된 MeSH 용어(5회이상)

MeSH 용어	횟수	%	누적 %
Internet	23	4.21	0.77
Reference Standards	19	3.48	4.25
Hospital Information Systems	12	2.20	6.45
Computer Communication Networks	12	2.20	8.65
Databases	11	2.01	10.66
Equipment and Supplies	11	2.01	12.68
Medical Informatics	11	2.01	14.69
Neural Networks (Computer)	9	1.65	16.34
Computer-Assisted Instruction	8	1.47	17.80
Multimedia	8	1.47	19.27
Hospital Communication Systems	7	1.28	20.55
Information Systems	7	1.28	21.83
Telemedicine	7	1.28	23.12
Nursing Diagnosis	6	1.10	24.21
Nursing Records	6	1.10	25.31
Information Storage and Retrieval	6	1.10	26.41
Medical Records Systems, Computerized	5	0.92	27.33
Diagnostic Imaging	5	0.92	28.24
Mining	5	0.92	29.16
Computer Security	5	0.92	30.08
Automatic Data Processing	5	0.92	30.99
소 계	188	34.43	65.42
All others(5회미만)	**358**	**65.57**	**65.57**

표 19는 논문에 주어진 저자 용어와 MeSH 용어가 완전히 일치한 경우를 보여주는 예이다.

표 19. 저자용어와 MeSH 용어가 완전히 일치한 경우의 예

저자용어 및 MeSH 용어	일치 횟수	%
Internet	13	2.38
Telemedicine	6	1.10
Multimedia	5	0.92
Nursing Diagnosis	4	0.73
Attitude	3	0.55
Spirometry	2	0.37
Fuzzy Logic	1	0.18
Job Satisfaction	1	0.18
Information Management	1	0.18
Infection Control	1	0.18
Hospice Care	1	0.18
Karyotyping	1	0.18
Hepatitis	1	0.18
Medical Informatics	1	0.18
Education	1	0.18
Diagnosis	1	0.18
Dementia	1	0.18
Confidentiality	1	0.18
Computers	1	0.18
Cadaver	1	0.18
Heterogeneous	1	0.18
Preventive Medicine	1	0.18
Teleradiology	1	0.18
Telemetry	1	0.18
Survival Analysis	1	0.18
Sleep	1	0.18
Quality Control	1	0.18
Learning	1	0.18
Privacy	1	0.18
Knowledge	1	0.18
Pattern Recognition	1	0.18
Nursing Research	1	0.18
Nursing Process	1	0.18
Needs Assessment	1	0.18
Urology	1	0.18
Mass Screening	1	0.18
Problem-Based Learning	1	0.18
계	**64**	**11.72**

3. MeSH내에서 보건의료정보분야의 범위 설정

가. 대한의료정보학회지에 부여된 MeSH 용어와 범주간 상관성 분석

표 20은 대한의료정보학회지에 부여된 MeSH 용어가 MeSH의 15범주 내에서 어느 곳에 분포하는지를 나타낸 표이다[6]. 해당범주별로 살펴보면, L. Information Science가 총 267건으로 가장 많이 분포했음을 알 수 있으며 이는 전체중 28.34%에 해당한다. 두번째는 N. Health Care인 265건(28.13%)으로 가장 많은수를 차지한 L. Information Science와의 차이가 거의 없는 것을 볼 수 있다. 세 번째의 경우는 114건(12.10%)으로 건수 및 비율이 현저히 줄어든 것을 알 수 있다.

표 20. 대한의료정보학회지에 부여된 MeSH 용어의 MeSH 15범주내에서의 분포도 및 다중분포 순위

	MeSH 15개 범주		횟수	순위
1	A.	Anatomy	10	9
2	B.	Organisms	4	14
3	C.	Diseases	22	8
4	D.	Chemicals and Drugs	9	10
5	E.	Analytical, Diagnostic and Therapeutic Techniques and Equipment	114	3
6	F.	Psychiatry and Psychology	39	6
7	G.	Biological Sciences	85	4
8	H.	Physical Sciences	67	5
9	I.	Anthropology, Education, Sociology and Social Phenomena	37	7
10	J.	Technology and Food and Beverages	9	11
11	K.	Humanities	7	13
12	L.	Information Science	267	1
13	M.	Persons	6	12
14	N.	Health Care	265	2
15	Z.	Geographical Locations	1	15
	총 계		942	

6) MeSH가 의미론적으로 hierarchically controlled tree structure이므로 (표 20)의 총계는 546건 보다 많다.

나. MeSH 중 L. Information Science 이외의 범주와 보건의료정보 관련분야 설정

표 21은 MeSH 중 L. Information Science 이외의 범주에서 보건의료정보관련 분야를 정의하기 위해 'inform' 관련 분야를 검색한후 보건의료정보와 관련된 분야를 재추출한 결과이다. 결과를 살펴보면 대한의료정보학회지를 기준으로 분석한 결과인 표 20과 유사하게 N. Health Care가 많은 부분을 차지하였으며, 그 외 I. Anthropology, Education, Sociology and Social Phenomena범주가 검색되었다. 검색되어진 결과를 보건의료정보관련 주제어로 제안하기 위하여 표 21에서 보여지는 세가지 방법을 통하여 비교하여 보았다. 첫째는 'I 완전 일치만을 포함'은 검색되어진 단어만을 포함하는 방법이다. 이 방법을 적용한 경우에는 추가되어지는 용어수가 25개에 불과한 협의의 방법이므로 본 논문에서는 선택치 않기로 한다. 둘째 및 셋째를 같이 살펴보면 'II 상위 두 단계까지 확장'과 'III 최상위 단계까지 확장'은 MeSH 기본구조가 hierarchically controlled tree structure라는 것을 활용하여 확장(explosion)을 적용한 것이다. 'II 상위 두 단계까지 확장'의 경우 상위 두 번째 단계(Tree Number 두자리, 예: N02.421)까지를 포함한 경우이며, 'III 최상위 단계까지 확장'은 최상위(Tree Number 두자리, 예: N02)까지 확장한 것이다. MeSH와 같은 구조적 통제어의 검색 개념중 정확도(precision)와 회귀도(recall)에 대한 고려가 'II 상위 두 단계까지 확장'과 'III 최상위 단계까지 확장' 중 어느 부분을 취할지에 대한 척도로 쓰이며, 정확도와 회귀도를 고려하기 위한 기초자료로는 표 20의 결과가 사용된다. 즉, 보건의료정보분야 주제어에 L. Information Science 이외의 주제어를 포함하기 위하여 각 범주별 관련성을 표 20에서 살펴보면 N. Health Care의 경우에는 L. Information Science에 근접한 밀접성을 보였으나, I. Anthropology, Education, Sociology and Social Phenomena

는 중간수위로 확장시 밀접성이 큰 경우(N. Health Care)는 broader(최상위 단계)로, 밀접성이 적은경우(I. Anthropology, Education, Sociology and Social Phenomena)는 narrower(상위 두단계)까지를 적용하기로 한다.

결국 본 논문에서는 보건의료정보분야에 포함시키는 주제어의 선택을 표 21 중 'II 상위 두 단계까지 확장'과 'III 최상위 단계까지 확장'을 혼용한 방식을 선택하되 밀접성이 큰 범주인 N. Health Care의 경우는 'III 최상위 단계까지 확장'하며 관련성이 떨어지는 I. Anthropology, Education, Sociology and Social Phenomena의 경우는 'II 상위 두 단계까지 확장'하는 방법을 취하기로 한다.

다. MeSH를 기준으로 보건의료정보분야 주제어집 제시

보건의료정보분야의 주제어집를 만들기 위해 MeSH를 바탕으로 Medical informatics 관련분야를 정의한다[7].

첫째, 15개의 MeSH 범주 중 "L. Information Science"는 전체를 포함하였다. "L. Information Science"의 MeSH 용어는 370개(표 22)이며, 370개에 대한 1,921개의 entry term(표 23)을 제공하고 있다.

둘째, MeSH 중 L. Information Science 이외의 범주에서 보건의료정보학회지를 기준으로 보건의료정보 관련분야를 재정의한 후 이를 포함하였다. 이때 추가되어지는 MeSH 용어의 수는 1,306(표 24)이며, entry term의 수는 7,892개(표 25)이다.

셋째, 저자가 사용한 용어중 MeSH 용어에 없으나, 필요하다고 인정되는 용어를 추가하였다.

---

7) 2002년 개정된 MeSH 기준임

표 21. MeSH의 L(Informatin Science)을 제외한 범주 중 의료정보관련 분야

Ⅰ 완전 일치만을 포함		Ⅱ 상위 두 단계까지 확장			Ⅲ 최상위 단계까지 확장		
MeSH Heading	Tree Number	MeSH Heading	Tree Number	개수	MeSH Heading	Tree Number	개수
National Health Planning Information Center	I01.409.275.300.650.790.525	**Government**	**I01.409**	**30**	Social Sciences	I01	376
Access to Information	I01.880.604.473.352.500.500	**Sociology**	**I01.880**	**201**			
Informed Consent	I01.880.604.473.650.718						
Informed Consent	I01.880.604.583.427						
Drug Information Services	N02.421.668.320	Health Services	N02.421	228	**Health Care Facilities, Manpower, and Services**	**N02**	**433**
Clinical Pharmacy Information Systems	N02.421.668.320.200						
National Health Planning Information Center	N03.540.427.300.650.790.525	Organizations	N03.540	74	**Health Care Economics and Organizations**	**N03**	**345**
Informed Consent	N03.706.437.650.312	Social Control, Formal	N03.706	82			
Informed Consent	N03.706.535.489						
Hospital Information Systems	N04.452.442.452	Organization and Administration	N04.452	221	**Health Services Administration**	**N04**	**297**
Clinical Pharmacy Information Systems	N04.452.442.452.200						
Operating Room Information Systems	N04.452.442.452.600						
Radiology Information Systems	N04.452.442.452.760						
Management Information Systems	N04.452.515						
Ambulatory Care Information Systems	N04.452.515.050						
Clinical Laboratory Information Systems	N04.452.515.080						
Clinical Pharmacy Information Systems	N04.452.515.095						
Hospital Information Systems	N04.452.515.360						
Ambulatory Care Information Systems	N04.452.515.360.050						
Clinical Laboratory Information Systems	N04.452.515.360.110						
Clinical Pharmacy Information Systems	N04.452.515.360.250						
Operating Room Information Systems	N04.452.515.360.555						
Radiology Information Systems	N04.452.515.360.750						
Personnel Staffing and Scheduling Information Systems	N04.452.515.800						
Radiology Information Systems	N04.452.515.825						
총 계(개)	**25**		**836**			**1,451**	

표 22. "L. Information Science"의 MeSH 용어의 예 (총 370개)

Tree Number	MeSH Heading	Depth
L01	Information Science	1
L01.040	Book Collecting	2
L01.080	Chronology	2
L01.100	Classification	2
L01.143	Communication	2
L01.143.050	Advertising	3
L01.143.230	Communication Barriers	3
L01.143.283	Cybernetics	3
L01.143.283.425	Feedback	4
L01.143.320	Diffusion of Innovation	3
L01.143.320.800	Technology Transfer	4
L01.143.380	Hotlines	3
L01.143.506	Language	3
L01.143.506.423	Language Arts	4
L01.143.506.423.348	Lipreading	5
L01.143.506.423.452	Multilingualism	5
L01.143.506.423.557	Reading	5
L01.143.506.423.676	Speech	5
L01.143.506.423.796	Translating	5
L01.143.506.423.906	Writing	5
L01.143.506.423.906.215	Authorship	6
L01.143.506.423.906.377	Correspondence	6
L01.143.506.423.906.539	Handwriting	6
L01.143.506.423.906.539.500	Paleography	7
L01.143.506.423.906.539.750	Shorthand	7
L01.143.506.598	Linguistics	4
L01.143.506.598.400	Terminology	5
L01.143.506.598.400.556	Names	6
L01.143.506.598.400.556.131	Abbreviations	7
L01.143.506.598.400.556.283	Anonyms and Pseudonyms	7
L01.143.506.598.400.556.536	Eponyms	7
L01.143.506.598.518	Phonetics	5
L01.143.506.598.628	Psycholinguistics	5
	.	
	.	
	계속	

표 23. "L. Information Science"의 Entry term의 예 (총 1,921개)

Tree Number	MeSH Heading	Depth	Entry Term
L01	Information Science	1	Science, Information
L01	Information Science	1	Sciences, Information
L01	Information Science	1	Information Sciences
L01.040	Book Collecting	2	Book Collectings
L01.040	Book Collecting	2	Collecting, Book
L01.040	Book Collecting	2	Collectings, Book
L01.080	Chronology	2	Dates
L01.100	Classification	2	Classifications
L01.100	Classification	2	Taxonomies
L01.143	Communication	2	Information Distribution
L01.143	Communication	2	Programs, Communication
L01.143	Communication	2	Program, Communication
L01.143	Communication	2	Personnel, Communications
L01.143	Communication	2	Misinformation0|abcdef
L01.143	Communication	2	Distribution, Information
L01.143	Communication	2	Communication Program
L01.143	Communication	2	Communications Personnel
L01.143	Communication	2	Communication, Personal
L01.143	Communication	2	Communication Programs
L01.143	Communication	2	Personal Communication
L01.143.050	Advertising	3	Advertisements
L01.143.050	Advertising	3	Promotion
L01.143.050	Advertising	3	Advertisement
L01.143.230	Communication Barriers	3	Barriers, Communication
L01.143.230	Communication Barriers	3	Communication Barrier
L01.143.230	Communication Barriers	3	Barrier, Communication
L01.143.283	Cybernetics	3	Cybernetic
L01.143.320	Diffusion of Innovation	3	Diffusion, Innovation
	:		
	계속		

표 24. "L. Information Science"**이외의 범주 중 보건의료정보관련** MeSH **용어의 예** (**총** 1,306**개**)

Tree Number	MeSH Heading	Depth
I01.409	Government	2
I01.409.275	Government Agencies	3
I01.409.275.125	United States Department of Agriculture	4
I01.409.275.250	United States Department of Veterans Affairs	4
I01.409.275.300	United States Dept. of Health and Human Services	4
I01.409.275.300.310	United States Centers for Medicare and Medicaid Services	5
I01.409.275.300.650	United States Public Health Service	5
I01.409.275.300.650.225	Centers for Disease Control and Prevention (U.S.)	6
I01.409.275.300.650.225.520	National Institute for Occupational Safety and Health	7
I01.409.275.300.650.400	National Center for Health Care Technology	6
I01.409.275.300.650.425	National Center for Health Statistics (U.S.)	6
I01.409.275.300.650.450	National Institutes of Health (U.S.)	6
I01.409.275.300.650.450.450	National Institute of Mental Health (U.S.)	7
I01.409.275.300.650.450.475	National Library of Medicine (U.S.)	7
I01.409.275.300.650.700	United States Agency for Healthcare Research and Quality	6
I01.409.275.300.650.760	United States Food and Drug Administration	6
I01.409.275.300.650.790	United States Health Resources and Services Administration	6
I01.409.275.300.650.790.525	National Health Planning Information Center	7
I01.409.275.300.650.825	United States Indian Health Service	6
I01.409.275.300.650.912	United States Office of Research Integrity	6
I01.409.275.300.650.920	United States Substance Abuse and Mental Health Services Administration	6
I01.409.275.340	United States Environmental Protection Agency	4
I01.409.275.380	United States Federal Trade Commission	4
I01.409.275.430	United States National Aeronautics and Space Administration	4
I01.409.275.480	United States Occupational Safety and Health Administration	4
I01.409.275.500	United States Office of Economic Opportunity	4
I01.409.275.550	United States Office of Technology Assessment	4
I01.409.275.775	United States Social Security Administration	4
I01.409.550	Local Government	3
I01.409.700	State Government	3
I01.880	Sociology	2
I01.880.143	Culture	3
I01.880.143.079	Acculturation	4
I01.880.143.257	Cross-Cultural Comparison	4
I01.880.143.329	Cultural Characteristics	4
I01.880.143.400	Cultural Deprivation	4
	:	
	계속	

표 25. "L. Information Science"**이외의 범주 중 보건의료정보관련** Entry term **의 예** (**총** 7,892**개**)

Tree Number	MeSH Heading	Depth	Entry Term
I01.409	Government	2	Governments
I01.409	Government	2	Federal Government
I01.409	Government	2	Federal Governments
I01.409	Government	2	Government Official
I01.409	Government	2	Government, Federal
I01.409	Government	2	Governments, Federal
I01.409	Government	2	Official, Government
I01.409	Government	2	Officials, Government
I01.409	Government	2	Government Officials|
I01.409.275	Government Agencies	3	Agency for International Development
I01.409.275	Government Agencies	3	Agencies, Government
I01.409.275	Government Agencies	3	Government Agency
I01.409.275	Government Agencies	3	Peace Corps
I01.409.275	Government Agencies	3	USAID
I01.409.275	Government Agencies	3	Agency, Government
I01.409.275.250	United States Department of Veterans Affairs	4	Veterans Administration
I01.409.275.250	United States Department of Veterans Affairs	4	United States Veterans' Administration
I01.409.275.300	United States Dept. of Health and Human Services	4	United States Department of Health, Education and Welfare
I01.409.275.300	United States Dept. of Health and Human Services	4	United States Department of Health, Human Services
I01.409.275.300	United States Dept. of Health and Human Services	4	United States Department of Health
I01.409.275.300	United States Dept. of Health and Human Services	4	Department of Health and Human Services
I01.409.275.300	United States Dept. of Health and Human Services	4	United States Department of Human Services
I01.409.275.300.310	United States Centers for Medicare and Medicaid Services	5	Centers for Medicare and Medicaid Services
I01.409.275.300.310	United States Centers for Medicare and Medicaid Services	5	Health Care Financing Administration
I01.409.275.300.650	United States Public Health Service	5	USPHS
I01.409.275.300.650	United States Public Health Service	5	Public Health Service
I01.409.275.300.650.225	Centers for Disease Control and Prevention (U.S.)	6	Center for Disease Control and Prevention
	:		

계속

Ⅴ. 고 찰

국내 보건의료분야 문헌의 DB 구축과 검색을 위해서는 통제된 용어의 사용이 절대적으로 필요하다. 용어의 통일성과 일관성을 유지하기 위해서는 논문을 발행할 때 키워드 선정시부터 통제된 용어의 사용이 이루어져야 한다. 이러한 인식 하에 보건의료분야 학술지, 특히 의학학술지의 대부분이 저자가 자신의 논문의 키워드를 선정할 때 MeSH 용어를 사용하도록 투고규정에 권장하고 있지만 아직까지 보건의료분야 학회중 단지 23개 학회만 용어집 혹은 색인집을 발표하고 있는 실정이다. 국내 의학주제목록을 구축하기 위해서는 우선 학회별로 지금까지 저자들이 사용한 저자 키워드를 참고로 MeSH와 어느 정도 일치하는지 파악하는 것이 중요하다. 그리고 MeSH와 일치하지 않는 용어의 경우 국내 표준용어집의 구축이 시급하다.

이에 본 연구는 대한의료정보학회지 1995년 1권 1호(창간호)에서 2000년 6권 4호까지의 전체논문을 대상으로 하여, 1995년에서 2000년까지 대한의료정보학회지의 논문편수 및 권호별 논문수 및 저자용어등의 변화추이등을 분석하였으며, 각각의 중심단어가 MeSH 용어와 갖는 차이점등을 분석하고, 분석결과를 기초로 2002년 MeSH 중 보건의료정보분야에 해당하는 용어를 설정하여 "한국보건의료정보분야 주제어집"을 제시하였다. 이에 대한 구체적인 결과를 살펴보면 첫째, 대한의료정보학회지 1995년에서 2000년까지의 총 논문수는 172편이었으며 이중 153편에 저자용어(Keyword)가 부여되었으며 19편이 저자용어가 부여되지 않았다. 논문수에서는 1995년에 7편에서 시작한 논문이 지속적으로 증가하여 1999년 46편, 2000년 42편으로 양적으로도 많은 발전이 있었음을 보여준다. 부여된 저자용어수를 살펴보면 총 153편의 논문에 546개의 저자용어가 부여되어 있어, 평균 3.57

개의 영문주제어를 부여하고 있으며 이는 논문수의 증가와 무관하게 개선되지 않는 양상을 보였다. 평균 저자용어수는 MEDLINE이 부여하는 평균 MeSH 용어수인 10개에는 크게 못미치는 개수로 앞으로 개선되어야 할 사항이다. 둘째, 저자가 설정한 영문주제어를 기준으로 MeSH browser를 통하여 MeSH 용어로 변환후 논문별로 저자용어와 MeSH 용어의 비교하였다. 일치도에 따른 해당 용어수를 살펴보면, 부분적으로 일치하는 경우 중 사소한 차이인 단/복수 오류(4.03%), entry term 사용(10.07%), 약어사용(0.73%)을 포함한다하더라도 25.56%에 불과하였다. 이 중 저자가 MeSH 용어를 완전하게 사용하는 경우는 단지 11.72%에 불과하며, 이는 대한의료정보학회지 저자들이 키워드 선정시 MeSH 용어에 대한 기본적인 고려조차하지 않은 채 키워드를 선정하는 것을 보여주는 것이라 할수 있다. 이에 대한 해결책으로는 MeSH에 대한 교육을 통한 이해 및 키워드 선정시 MeSH browser 통한 확인등으로 해결이 가능할 것이다. 대한의료정보학회지 저자들이 많이 사용한 용어를 살펴보면, 저자가 많이 사용한 용어를 4회이상으로 했을 때 11건이 나타났으며 이는 전체중 12.27%를 차지한다. 저자용어로 가장 많이 사용된 용어는 "Internet"이며 전체 저자용어중 2.28%를 차지하였다. 이중 완전일치한 용어는 4건 5.03%였으며, 부분일치한 경우(entry term 사용)가 3건 3.48%이었고, 불일치한 경우가 4건 3.76% 이었다. 저자가 키워드를 선택할 때 MeSH browser를 사용하여 확인한다면, MeSH 용어의 정확한 사용이 가능할 것이다. 대한의료정보학회지에 부여된 MeSH용어를 저자용어를 기준으로 MeSH browser를 이용하여 검색한 후 대한의료정보학회지에 저자가 많이 사용한 MeSH 용어(5회이상)를 살펴보면, MeSH 용어로 가장 많이 사용된 용어 또한 저자 용어와 동일하게 "Internet"임을 볼 수 있으나 횟수가 현격히 증가하였음을 알 수 있다. 많이 사용된 용어의 수를 비교해보면 저자용어의 경우 4회이상 사용된 용어

가 67개(12.27%)인데 반해, MeSH용어의 경우 188개(34.43%)로 전자에 비하여 용어의 통일화가 이루어져 있음을 알 수 있다. 즉, MeSH를 사용하는것 만으로도 MeSH 기본적인 통제어로서의 성질이 저자가 선택할 수 있는 용어수를 한정하여 주므로 자연스러운 통일화 및 그룹화가 이루어지게 되는 것이다. 셋째로 대한의료정보학회지에 부여된 MeSH 용어가 MeSH 15범주 내에서 어느 곳에 분포하는지를 살펴보면, L. Information Science가 총 267건으로 가장 많이 분포했음을 볼 수 있으며 이는 전체중 28.34%에 해당한다. 두 번째를 차지한 N. Health Care의 경우 265건(28.13%)으로 1위인 L. Information Science와의 차이는 거의 없는 것을 볼 수 있다. 세번째의 경우는 114건(12.10%)으로 건수 및 비율이 현저히 줄어든 것을 볼 수 있어, L. Information Science와 N. Health Care가 MeSH중 보건의료정보분야와 밀접한 상관성이 있음을 알수 있다. 이 결과를 바탕으로 본 논문에서는 보건의료정보분야에 포함시키는 주제어의 선택시 MeSH 기본구조가 hierarchically controlled tree structure라는 것을 활용하여 확장(explosion)을 적용하였으며 정확도(precision)와 회귀도(recall)에 대한 고려를 하였다. 즉, N. Health Care는 관련성이 많으므로 broader(최상위 단계)를 적용하였고, I. Anthropology, Education, Sociology and Social Phenomena는 중간수위에 머물렀으므로 narrower(상위 두단계)까지를 적용하였다. 넷째, 이상의 대한의료정보학회지의 분석결과를 바탕으로 설정된 MeSH 범주중 Medical informatics 관련분야를 설정한후 MeSH를 바탕으로 한 한국보건의료정보분야 주제어집을 제시하였다(부록 2. 참조).

이상의 결과를 바탕으로 MEDLINE에 게재되는 학술지와 대한의료정보학회지의 저자용어를 비교하여볼 때 학술지에서 요구하는 주제어 표기 방식과 수준은 MEDLINE 레코드에 적용되는 NLM의 MeSH 부여 방식이나 수준과 확실히 차이

가 있었다. MeSH 용어는 주제개념에 subheading을 연결하여 표기하고, 특히 "/" 를 이용하여 subheading의 구분을 하고 있다. 연구비 수혜 여부 및 연구비 지원 기관의 성격을 나타내는 check tag, 실험대상의 성격을 규명하는 human, animal 등의 check tags 등을 한 논문에 3-4개씩 부여하고 있었다. 또한 대한의료정보학회지에서는 전혀 사용하지 않는 major term을 표기하고 있다(Lowe 등, 1994). 그렇다면, 계속 출판되는 한국 의학학술논문에 적절한 MeSH 용어를 누가, 어떻게, 효율적으로 부여할 것인가? 우선, 저자용어 중 약어, 단수/ 복수의 차이, 어순이 바뀐 용어 등 MeSH용어에 근접한 용어는, 학술지의 편집자가 저자와 협의하여 MeSH 용어로 바꾸어 학술논문에 출판하도록 하여야 할 것이다. 학술지는 저자들에게 check tags의 사용을 권장하고, 필요하다면 투고규정에 10개 정도의 저자용어를 사용하라고 권장하여야 할 것이다(International Committee of Medical Journal Editors, 1997). 학술지의 편집자들은 저자가 키워드를 선정할 때 핵심개념의 주제어는 World Wide Web으로 접근 가능한 MeSH browser를 활용하여 선정하도록 유도하고, PubMed를 검색하여 비슷한 주제를 가진 다른 논문에 부여된 MeSH 용어를 참고하도록 권장하여야 할 것이다. 학술지가 요구하는 형식에 맞추기 위하여 subheading을 나열하는 MeSH 방식을 취하지는 않더라도, subheading에 해당하는 어휘들을 키워드로 부여하도록 권장하여야 할 것이다. 왜냐하면, 이러한 용어들을 검색에 사용한다면 post-coordination(후조합)이 가능하므로, 정확한 MeSH 용어 사용에 버금가는 효과를 얻을 수 있을 것이기 때문이다(Clarke 등, 1997). 대한기생충학회 집담회 참석자를 대상으로 한 설문조사(1999년 4월) 결과에 의하면, MeSH에 대하여 잘 모른다고 고백한 경우가 25%, MeSH가 계층구조를 갖고 있다는 것을 모른다는 저자가 70%로 나타났다. 그럼에도 불구하고, 77%의 응답자가 논문의 주제어 선정은 저자가 해야 한다고 주장

하였다(이춘실 등, 2000). 그러므로, 학회지 투고저자들을 대상으로 하는 "MeSH 기본 교육"이 매우 절실히 요구된다. 학술논문에 MeSH 용어를 부여하는 일은 저자에게만 맡겨서는 안되고, 각 학술지의 편집자와 전문인력이 학술지에 투고하는 저자들을 교육하고, 의학용어 사용에 있어서 중지를 모으는 등 총괄하도록 하여야 할 것이다(National Library of Medicine, 1997). 또한, MeSH 전문가 검색시스템에 입력되는 MeSH 용어를 일관성 있게 감수하거나 직접 부여하여야 할 것이다.

이러한 문제점을 극복하고 한국 저자들이 만족할 수 있는 검색시스템의 구축 방법론을 마련하자면, 저자용어는 저자용어대로 주제검색에 사용하면서 별도로 MeSH 용어를 검색에 활용하는 방안이 모색되어야 할 것이다(Celia 등, 2001). 그래야만 국내논문검색시스템 PubMed와 같은 방식으로 검색되어질 수 있기 때문이다. 현재 국내에서 사용중인 저자용어는 각 학회 및 개인적으로 사용하는 저자용어를 논문의 중심단어로 입력함으로써 동의어 및 유사어에 대한 데이터베이스의 혼동으로 인한 문제점을 안고 있다. 주제어에 대한 정의 및 통일화가 절실히 필요하며 이때 MeSH와 저자용어를 동시에 사용 가능토록 해주는 메타 용어집이 좋은 대안이 될 수 있을 것이다.

첨단화된 연구의 폭발적인 양적 증가와 새롭고 급변하는 보건의료정보는 보건의료과학과 정보기술 및 컴퓨터공학 분야의 관련 테크놀로지와 이론들(disciplines)을 결합한 학제간 협력분야임에도 불과하고 체계적으로 운영 관리하는데 그 동안 우리 나라 보건의료인들은 상대적으로 타 분야 전문가에 비해 무관심해 왔다. 그러나, 보건의료인들의 주요 세부 관심 분야인 교육, 연구, 진료의 부문에 있어서, 컴퓨터 프로그램은 문헌찾기, 데이터 저장, 통계적 분석 시스템 등의 결정적 도움을 받는다. 우리 나라의 경우, 기존 보건

의료 및 복지 분야의 도서관 환경으로는 급변하는 정보화 · 세계화시대에 당면한 문제해결연구가 어려우므로, 중장기적으로는 새로운 연구지원환경 구축 및 한국형 국립 보건의료과학기술도서관 설립 · 운영의 추진의 필요성이 제기되어 왔다(김병국 등, 1998). 이를 위해 지속적이고 체계적인 준비와 관련 전문인력양성, 통합 및 분산 연구 정보시스템 기반기술연구 등 정보인프라 구축이 요구된다. 본 논문에서는 이를 위한 작은 초석인 보건의료정보분야 주제어집(용어집)을 제시하였다. 이러한 다각적인 노력을 통하여서만 수준 있는 MeSH 용어 부여 작업이 가능하고, 국제적으로 유통되는 의학정보의 생산이 비로소 이루어지게 될 것이다.

Ⅵ. 결 론

본 연구는 대한의료정보학회지 창간호(1995년 1권 1호)에서 2000년 6권 4호까지의 전체논문을 대상으로 하였다. 1995년에서 2000년까지 대한의료정보학회지의 논문편수 및 권호별 논문수 및 저자용어등의 변화추이등을 분석하였으며, 각각의 중심단어가 MeSH 용어와 갖는 차이점등을 분석하고, 분석결과를 기초로 2002년 MeSH 중 보건의료정보분야에 해당하는 용어를 설정하여 "한국보건의료정보분야 주제어집"을 제시하였다. 첫째, 대한의료정보학회지 1995년에서 2000년까지의 총 논문수는 172편이었으며 이중 153편에 저자용어(Keyword)가 부여되었으며 19편이 저자용어가 부여되지 않았다. 1995년에 7편에서 시작한 논문이 지속적으로 증가하여 1999년 46편, 2000년 42편으로 양적으로도 많은 발전이 있었으나, 부여된 저자용어수는 평균 3.57개로 논문수의 증가와 무관하게 개선되고 있지 않은 양상을 보였다. 즉, 평균 저자용어수를 살펴볼 때 MEDLINE이 부여하는 평균 MeSH 용어수인 10개에는 크게 못미치고 있어 개선이 요구되고 있다. 둘째, 저자가 설정한 영문주제어를 기준으로 MeSH browser를 통하여 저자용어와 MeSH 용어의 일치도를 비교하였다. 일치도를 살펴보면, 부분적으로 일치하는 경우 중 사소한 차이인 단/복수 오류(4.03%), entry term 사용(10.07%), 약어사용(0.73%)을 포함한다하더라도 25.56%에 불과하였다 저자가 MeSH 용어를 완전하게 사용하는 경우는 단지 11.72%에 불과함을 보여주고 있다. 이는 대한의료정보학회지 저자들이 키워드 선정시 MeSH용어에 대한 기본적인 고려조차하지 않은 채 키워드를 선정하는 것을 보여주는 것으로 ,통일되고 일관성 용어의 사용이 왜 중요한지 교육을 하고

논문의 저자가 키워드를 선택할 때 표준용어집 혹은 MeSH용어를 참고하여 선택하도록 권고한다. 이렇게 되면 논문 검색시스템을 구축할 때 표준화된 용어 혹은 MeSH를 사용할 수 있을 것이다. 셋째, 대한의료정보학회지에 부여된 MeSH 용어가 MeSH 15범주 내에서 어느 곳에 분포하는지를 분석한 후 MeSH 중 보건의료정보분야에 포함시키는 주제어의 범위를 설정하였다. MeSH 확장시을 적용시 N. Health Care는 관련성이 많으므로 broader(최상위 단계)를 적용하였으며, I. Anthropology, Education, Sociology and Social Phenomena는 중간수위에 머물렀으므로 narrower(상위 두단계)까지를 적용하였다. 넷째, 대한의료정보학회지의 분석결과를 바탕으로 설정된 MeSH 범주중 Medical informatics 관련분야를 설정한후 MeSH를 바탕으로 한 한국 보건의료정보분야 주제어집을 제시하였다(부록 2 참조).

그 동안 우리나라 보건의료인들은 상대적으로 타 분야 전문가에 비해 무관심해 왔다. 이를 위해 지속적이고 체계적인 준비와 관련 전문인력양성, 통합 및 분산 연구 정보시스템 기반기술연구 등 정보인프라 구축이 요구된다. 국내 보건의료분야 문헌의 DB 구축과 검색을 위해서는 통제된 용어의 사용이 절대적으로 필요하다. 용어의 통일성과 일관성을 유지하기 위해서는 논문을 발행할 때 키워드 선정시부터 통제된 용어의 사용이 이루어져야 한다. 이러한 인식 하에 보건의료분야 학술지, 특히 의학학술지의 대부분이 저자가 자신의 논문의 키워드를 선정할 때 MeSH 용어를 사용하도록 투고규정에 권장하고 있지만 아직까지 보건의료분야 학회중 단지 23개 학회만 용어집 혹은 색인집을 발표하고 있는 실정이다. 국내 의학주제목록을 구축하기 위해서는 우선 학회별로 지금까지 저자들이 사용한 저자 키워드를 참고로 MeSH와 어느 정도 일치하는지 파악하는 것이 중요하며, 이중 MeSH와 일치하지 않

는 용어의 경우 국내 표준용어집의 구축이 시급하다. 이와 동시에 저자들과 학회지 편집인을 대상으로 통일되고 일관성 용어의 사용이 왜 중요한지 교육을 한다. 그리고 논문의 저자가 키워드를 선택할 때 표준용어집 혹은 MeSH용어를 참고하여 선택하도록 권고한다. 이렇게 되면 논문 검색시스템을 구축할 때 표준화된 용어 혹은 MeSH를 사용할 수 있을 것이다. 이렇게 하여 학회별 표준 용어집이 구축되면 학회별 표준 용어집과 국내 논문에서 활용된 MeSH 목록을 통합하여 국내 의학주제목록을 구축할 수 있을 것이다. 본 논문에서는 이를 위한 작은 초석인 보건의료정보분야 주제어집(용어집)을 제시하였다. 물론 이렇게 구축된 주제어집은 미국 국립의학도서관에서 개발된 MeSH와 국내에서 새로 추가되는 용어를 보완하는 계속적인 수정.보완이 이루어져야 할 것이다. 이러한 다각적인 노력을 통하여서만 수준 있는 MeSH 용어 부여 작업이 가능하고, 국제적으로 유통되는 의학정보의 생산이 비로소 이루어지게 될 것이다.

참 고 문 헌

김병국, 민승관, 송준용, 박선희, 윤봉자. 국내 의학도서관의 현황 분석. 한국의학교육, 1998; 9(2)

김병선, 김수영. 가정의학회지 논문의 영문 주제어 선택에 있어서 MeSH 용어 사용 여부와 선택 정확도. 대한가정의학회지, 1998; 19(7): 531-537

김수영. 근거중심의학과 메드라인 검색. 의료정보학술대회초록집, 1997; 85-92

대한 의학학술지 편집인 협의회[http://www.kamje.or.kr]

대한의료정보학회. 보건의료정보. 현문사, 1999

대한의료정보학회[http://www.kosmi.org/kosmi/index.html]

박형욱, 손명세, 김한중, 박은철, 유승흠. 한국표준의료행위 분류체계 개발. 예방의학회지, 1996; 29(4): 112-127

보건연구정보센터[http://www.richis.org]

서울대학교 의학도서관. PubMed Searching Guide. 극동디앤씨, 1999

이승하, 김호배. 의학주제어 표목표(영·한): 1995. 의학연구정보센터, 1996

이춘실, 문혜원. 한국의학학술 논문의 저자선정 주제어와 MeSH 용어의 비교분석. 정보관리학회지, 2000; 17(3): 109-123

Anita B, Olivier B. Methods for exploring the semantics of the relationships between co-occurring UMLS concepts. IMIA, 2001; 48: 171-175

Bodenreider O. Using UMLS semantics for classification purposes. Proc AMIA Symp, 2000; 17(20): 86-90

Celia B, Vicent B, Jean-Raoul S. HONselest: Multilingual assistant search

engine operated by a concept-based interface system to decentralized heterogeneous source. IMIA, 2001; 48: 309-313

Clarke ML, Greaves S, James S. MeSH terms must be used in MEDLINE searches. BMJ, 1997; 314(7088): 1203

Darmoni SJ, Thirion B, Leroyt JP, Douyere M, Lacoste B, Godard C, Rigolle I, Brisou M, Videau S, Goupyt E, Piott J, Quere M, Ouazir S, Abdulrab H. A search tool based on encapsulated MeSH thesaurus to retrieve quality health resources on the internet. Med Inform Internet Med. 2001; 26(3): 165-178

Fremer E. Understanding MeSH for literature searches. JAMA, 1995; 273(3): 184-185

Funk ME, Reid CA, McGoogan LS. Indexing consistency in MEDLINE. Bull Med Libr Assoc, 1983; 71: 176-183

International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. Ann Intern Med. 1997; 126: 36-47

KoreaMed[http://www.koreamed.org]

Lowe HJ, Barnett GO. Understanding and using the medical subject headings(MeSH) vocabulary to perform literature searches. JAMA, 1994; 271(14): 1103-1108

MeSH browser[http://www.ncbi.nlm.nih.gov/entrez/meshbrowser.cgi]

MeSH browser[http://www.nlm.nih.gov/mesh/meshhome.html]

Mitchell JA, Johnson ED, Hewett JE, Proud VK. Medical students using

Grateful Med: analysis of failed searches and a six month follow-up study. Comput Biomed RES, 1992; 25(1): 43-55

National Library of Medicine. MEDLAS training program-index training syllabus national library of medicine index section BSD. National Library of Medicine, 1997

National Library of Medicine. Medical Subject Headings(MeSH). http://www.nlm.nih.gov/mesh/meshhome.html, 2002

O'Rourke A, Booth A. Another fine MeSH: Clinical medicine meets information science. Journal of information science, 1999; 25(4): 275-281

PubMed[http://www.ncbi.nlm.nih.gov/PubMed/]

Ray JG, Vermeulen MJ. Mizspellin and MEDLINE. BMJ, 1996; 313: 1658-1659

부 록 1

저자용어의 MeSH 용어 변환 예시

Ⅰ. 저자용어의 MeSH 용어 변환

1. 각 논문별 변환과정

1) 개 요

○ **국문논문명**: 웹 기반 임상병리검사정보 전달체계 구축

○ **영문논문명**: Development of WebBased Laboratory Information Accessing System

○ **저자**: 박준석, 민원기, 임경수, 임태환, 이정신, 곽원섭, 조한익

○ **출처**: 대한의료정보학회지 [1999] 5(1) p.149-155

○ **저자용어**: Web, Internet, Laboratory Information System, Hospital Information System

○ MeSH **용어**: Internet, Internet, Clinical Laboratory, Information Systems, Hospital Information Systems

2) 저자용어와 MeSH 용어의 개별 변환관계표

박준석, 민원기, 임경수, 임태환, 이정신, 곽원섭, 조한익. 1999. "웹 기반 임상병리검사정보 전달체계 구축". 대한의료정보학회지 5(1): 149-155.

저자 용어	MeSH 용어	일치도							주제어 추천
		완전 일치	부분일치					불 일치	
			2-1	2-2	2-3	2-4	2-5		
Web	Internet							Y	N
Internet	Internet	Y							N
Laboratory Information System	Clinical Laboratory Information Systems		Y						N
Hospital Information System	Hospital Information Systems			**Y**					N

Ⅱ. MeSH 적용사례 세부적용 과정

1. 논문개요

○ **국문논문명**: 웹 기반 임상병리검사정보 전달체계 구축

○ **영문논문명**: Development of WebBased Laboratory-Information -Accessing System

○ **저자**: 박준석, 민원기, 임경수, 임태환, 이정신, 곽원섭, 조한익

○ **출처**: **대한의료정보학회지** [19990] 5(1) p.149-155

○ Abstract

There is still lack of a convinient system that connect referring physicians to the information system of referral hospitals, We proposed to develop the laboratory-information-accessing system(LISA) for physicians referred to AMC referral center with Web-based internet technology. We constructed the menu of AMC Referral Center as a part of Asan Medical Center(AMC) homepage. The information of the referred patients wee collected in the separate internal server and then transferred to the external network sever by a batch. Referring physicians was able to connect to AMC Referral Center via AMC homepage and then browse the list of their patients by putting their identification number and password. At the next, the physicians chose the patients and the test item to be displayed. In order that the referral system is established tin the medical part, the LISA world be more needed, in the society, which internet is getting popular. This LISA was introduced successfully now, and it could be a model of national standard for hospital information system.

○ **저자용어**: Web, Internet, Laboratory Information System, Hospital Information System

2. 저자용어와 MeSH 용어의 개별 변환 과정

1) Web

→ Internet

< MeSH Descriptor Data >

MeSH Heading Internet
Tree Number L01.700.568.080.110.500
Scope Note A loose confederation of computer communication networks around the world. The networks that make up the Internet are connected through several backbone networks. The Internet grew out of the US Government ARPAnet project and was designed to facilitate information exchange.
Entry Term World Wide Web
Allowable Qualifiers CL EC HI IS LJ OG SD SN ST TD UT
Previous Indexing Computer Communication Networks (1996-1998)
History Note 99; see COMPUTER COMMUNICATION NETWORKS 1996-98
Unique ID D020407

< MeSH Tree Structures >

Information Science [L01]
Medical Informatics [L01.700]
Medical Informatics Computing [L01.700.568]
Computer Systems [L01.700.568.080]
Computer Communication Networks [L01.700.568.080.110]
Internet [L01.700.568.080.110.500]
Local Area Networks [L01.700.568.080.110.600]

2) Internet

→ Internet

< MeSH Descriptor Data >

MeSH Heading Internet
Tree Number L01.700.568.080.110.500
Scope Note A loose confederation of computer communication networks around the world. The networks that make up the Internet are connected through several backbone networks. The Internet grew out of the US Government ARPAnet project and was designed to facilitate information exchange.
Entry Term World Wide Web
Allowable Qualifiers CL EC HI IS LJ OG SD SN ST TD UT

Previous Indexing Computer Communication Networks (1996-1998)
History Note 99; see COMPUTER COMMUNICATION NETWORKS 1996-98
Unique ID D020407

< MeSH Tree Structures >

Information Science [L01]

Medical Informatics [L01.700]

Medical Informatics Computing [L01.700.568]

Computer Systems [L01.700.568.080]

Computer Communication Networks [L01.700.568.080.110]

Internet [L01.700.568.080.110.500]

Local Area Networks [L01.700.568.080.110.600]

3) Laboratory Information System

→ Clinical Laboratory Information Systems

< MeSH Descriptor Data >

MeSH Heading Clinical Laboratory Information Systems

Tree Number L01.700.508.300.110

Tree Number L01.700.508.300.680.065

Tree Number N04.452.515.080

Tree Number N04.452.515.360.110

Annotation DF: CLIN LABORATORY INFORMATION SYSTEMS

Scope Note Information systems, usually computer-assisted, designed to store, manipulate, and retrieve information for planning, organizing, directing, and controlling administrative and clinical activities associated with the provision and utilization of clinical laboratory services.

Entry Term Laboratory Information Systems

Entry Term Information Systems, Clinical Laboratory

Allowable Qualifiers CL EC HI IS LJ MA OG SD SN ST TD UT

Entry Version CLIN LABORATORY INFORMATION SYSTEMS

History Note 91(87); was see under INFORMATION SYSTEMS 1987-90

Unique ID D002984

< MeSH Tree Structures >

Information Science [L01]

Medical Informatics [L01.700]
Medical Informatics Applications [L01.700.508]
Information Systems [L01.700.508.300]
Clinical Laboratory Information Systems [L01.700.508.300.110]
Community Networks [L01.700.508.300.184]
Decision Support Systems, Clinical [L01.700.508.300.190]
Databases [L01.700.508.300.221] +
Hospital Information Systems [L01.700.508.300.408]
Integrated Advanced Information Management Systems [L01.700.508.300.420]
Management Information Systems [L01.700.508.300.680] +
Medical Records Systems, Computerized [L01.700.508.300.695]
MEDLARS [L01.700.508.300.710] +
Online Systems [L01.700.508.300.742] +
Radiology Information Systems [L01.700.508.300.780]
Reminder Systems [L01.700.508.300.790]
Unified Medical Language System [L01.700.508.300.895]

Information Science [L01]
Medical Informatics [L01.700]
Medical Informatics Applications [L01.700.508]
Information Systems [L01.700.508.300]
Management Information Systems [L01.700.508.300.680]
Ambulatory Care Information Systems [L01.700.508.300.680.030]
Clinical Laboratory Information Systems [L01.700.508.300.680.065]
Clinical Pharmacy Information Systems [L01.700.508.300.680.085]
Database Management Systems [L01.700.508.300.680.110]
Decision Support Systems, Management [L01.700.508.300.680.135]
Hospital Information Systems [L01.700.508.300.680.360] +
Office Automation [L01.700.508.300.680.708] +
Personnel Staffing and Scheduling Information Systems [L01.700.508.300.680.800]
Radiology Information Systems [L01.700.508.300.680.850]

Health Services Administration [N04]
Organization and Administration [N04.452]
Management Information Systems [N04.452.515]

Ambulatory Care Information Systems [N04.452.515.050]
Clinical Laboratory Information Systems [N04.452.515.080]
Clinical Pharmacy Information Systems [N04.452.515.095]
Database Management Systems [N04.452.515.110]
Decision Support Systems, Management [N04.452.515.135]
Hospital Information Systems [N04.452.515.360] +
Office Automation [N04.452.515.708] +
Personnel Staffing and Scheduling Information Systems [N04.452.515.800]
Radiology Information Systems [N04.452.515.825] +

Health Services Administration [N04]
Organization and Administration [N04.452]
Management Information Systems [N04.452.515]
Hospital Information Systems [N04.452.515.360]
Ambulatory Care Information Systems [N04.452.515.360.050]
Clinical Laboratory Information Systems [N04.452.515.360.110]
Clinical Pharmacy Information Systems [N04.452.515.360.250]
Operating Room Information Systems [N04.452.515.360.555]
Point-of-Care Systems [N04.452.515.360.652]
Radiology Information Systems [N04.452.515.360.750]

4) Hospital Information System

→ Hospital Information Systems

< MeSH Descriptor Data >

MeSH Heading Hospital Information Systems
Tree Number L01.700.508.300.408
Tree Number L01.700.508.300.680.360
Tree Number N04.452.442.452
Tree Number N04.452.515.360
Annotation DF: HOSP INFORMATION SYSTEMS
Scope Note Integrated, computer-assisted systems designed to store, manipulate, and retrieve information concerned with the administrative and clinical aspects of providing medical services within the hospital.
Entry Term Multi-Hospital Information Systems
Entry Term Information System, Hospital

Entry Term Information System, Multihospital
Entry Term Information Systems, Hospital
Entry Term Information Systems, Multihospital
Entry Term Multihospital Information Systems
Allowable Qualifiers CL EC HI LJ OG SN ST TD UT
Entry Version HOSP INFORMATION SYSTEMS
History Note 87
Unique ID D006751

< **MeSH Tree Structures** >

Information Science [L01]
Medical Informatics [L01.700]
Medical Informatics Applications [L01.700.508]
Information Systems [L01.700.508.300]
Clinical Laboratory Information Systems [L01.700.508.300.110]
Community Networks [L01.700.508.300.184]
Decision Support Systems, Clinical [L01.700.508.300.190]
Databases [L01.700.508.300.221] +
Hospital Information Systems [L01.700.508.300.408]
Integrated Advanced Information Management Systems [L01.700.508.300.420]
Management Information Systems [L01.700.508.300.680] +
Medical Records Systems, Computerized [L01.700.508.300.695]
MEDLARS [L01.700.508.300.710] +
Online Systems [L01.700.508.300.742] +
Radiology Information Systems [L01.700.508.300.780]
Reminder Systems [L01.700.508.300.790]
Unified Medical Language System [L01.700.508.300.895]

Information Science [L01]
Medical Informatics [L01.700]
Medical Informatics Applications [L01.700.508]
Information Systems [L01.700.508.300]
Management Information Systems [L01.700.508.300.680]
Ambulatory Care Information Systems [L01.700.508.300.680.030]
Clinical Laboratory Information Systems [L01.700.508.300.680.065]

Clinical Pharmacy Information Systems
[L01.700.508.300.680.085]
Database Management Systems [L01.700.508.300.680.110]
Decision Support Systems, Management
[L01.700.508.300.680.135]
Hospital Information Systems [L01.700.508.300.680.360]
Operating Room Information Systems
[L01.700.508.300.680.360.500]
Office Automation [L01.700.508.300.680.708] +
Personnel Staffing and Scheduling Information Systems
[L01.700.508.300.680.800]
Radiology Information Systems [L01.700.508.300.680.850]

Health Services Administration [N04]
Organization and Administration [N04.452]
Hospital Administration [N04.452.442]
Ancillary Services, Hospital [N04.452.442.060]
Centralized Hospital Services [N04.452.442.110]
Financial Management, Hospital [N04.452.442.180]
Hospital Communication Systems [N04.452.442.322]
Hospital Departments [N04.452.442.422] +
Hospital Distribution Systems [N04.452.442.450]
Hospital Information Systems [N04.452.442.452]
Clinical Pharmacy Information Systems [N04.452.442.452.200]
Operating Room Information Systems [N04.452.442.452.600]
Point-of-Care Systems [N04.452.442.452.680]
Radiology Information Systems [N04.452.442.452.760]
Hospital-Patient Relations [N04.452.442.510]
Hospital-Physician Relations [N04.452.442.530] +
Hospital Restructuring [N04.452.442.550] +
Hospital Shared Services [N04.452.442.565]
Hospital Shops [N04.452.442.585]
Libraries, Hospital [N04.452.442.600]
Materials Management, Hospital [N04.452.442.650] +
Medication Systems, Hospital [N04.452.442.675]
Product Line Management [N04.452.442.700]

Health Services Administration [N04]

Organization and Administration [N04.452]
Management Information Systems [N04.452.515]
Ambulatory Care Information Systems [N04.452.515.050]
Clinical Laboratory Information Systems [N04.452.515.080]
Clinical Pharmacy Information Systems [N04.452.515.095]
Database Management Systems [N04.452.515.110]
Decision Support Systems, Management [N04.452.515.135]
Hospital Information Systems [N04.452.515.360]
Ambulatory Care Information Systems [N04.452.515.360.050]
Clinical Laboratory Information Systems [N04.452.515.360.110]
Clinical Pharmacy Information Systems [N04.452.515.360.250]
Operating Room Information Systems [N04.452.515.360.555]
Point-of-Care Systems [N04.452.515.360.652]
Radiology Information Systems [N04.452.515.360.750]
Office Automation [N04.452.515.708] +
Personnel Staffing and Scheduling Information Systems [N04.452.515.800]
Radiology Information Systems [N04.452.515.825] +

부 록 2

MeSH를 기본으로한
한국보건의료정보분야 주제어집

ABSTRACT

The Study on Subject Words of Korean Medical Informatics by Expanded MeSH

: Based on Journal of The Korean Society of Medical Informatics

Ae Kyung Kwon

Graduate School of

Health Science and Management

Yonsei University

(Directed by Professor Young Moon Chae, Ph. D.)

Health informatics has been characterized as an interdisciplinary field with various other fields such as informational technology, health science, and management. Despite of such interdisciplinary nature, professionals in this field have been neglected about establishing an infrastructure for supporting collaborative researches, constructing a scientific literature database, and training specialists in health informatics. In order to maintain a uniformity and consistency in terminology for constructing and searching a literature database, controlled vocabularies should be used as

key words in the journal.

While most of medical academic societies recommended MeSH be used as key words, only 23 societies published terminology or index book. Purposes of this study were to examine a current status of using MeSH in the Journal of Korea Society of Medical Informatics (KOSMI) and to develop a Korean terminology book on medical informatics. We reviewed 172 journals of KOSMI published during the period from 1995 to 2000 using MeSH brower. Only 11.72% of key words were completely consistent with MeSH terminolgy and 25.56% were partially consistent.

In the future, the terminology book from this study should be revised again using the new version of MeSH published by the National Library of Medicine in the United States and new terminologies introduced in other related fields in Korea.