

개념중심 용어테이블을 이용한 효능 중심 본초 검색방법 연구

대구의대학교 한의과대학·한국한의학회연구원·관문한의원
이병욱·엄동명·김정자·박지하*

Searching Methods of Studying Herb Efficacy by The use of Terminology Table

Dep. of Oriental Medical Classics, Daeguhaany University · Korea Institute of Oriental Medicine
· Gwan-Mun Oriental Medical Clinic
Lee, Byung-Wook · Kim, Jeong-Ja · Eom, Dong-Myung · Park, Ji-Ha

The purpose of this study is standardization of herbology terminology. For a long time, it was important problem that the Oriental Medicine didn't have standardization of terminology. For this reason, development of oriental medicine has been slow. We studied how we search methods of herb efficacy by the use of terminology table. In this way, we will suggest basic data and helpful methods of standardization of Oriental Medicine terminology.

Key Words: herb, terminology, searching method

I. 서론

현재 한의학 분야에서는 한의학의 발전과 국민의 보건수준 향상을 위한 다양한 연구가 진행되고 있다. 이러한 연구들은 약재로부터 유효성분을 밝혀내는 연구에서부터 특정 질환에 대한 치료방법의 유효성 및 안정성을 검증하는 임상시험에 이르기 까지 다양한 종류의 연구방법들이 시도되고 있다. 그러나 이러한 연구에 앞서서 한의학적 사고를 표현하는 용어에 대한 표준화가 이루어지지 않은 점 때문에 여러 가지 문제가 발생된다. 이러한 문제의 발단은 동일한 개념을 가진 여러 용어가 혼재되어 사용되고 또한 동일한

용어가 여러 가지 개념을 동시에 갖는 경우에서 비롯된다. 전자의 경우에는 연구자가 특정 개념을 검색하여 결과값을 얻고자 할 때 사용하는 용어에 대하여 표준화가 이루어지지 않았을 경우 결과에 대한 신뢰도가 낮아진다. 예를 들면, 濕邪를 제거하는 효능을 가진 약재를 찾는다고 가정할 때 사용자가 '去濕', '祛濕', '除濕', '去風濕' 등 찾아야할 용어가 너무 많다. 때문에 DB를 구축한 상황일지라도 만족할만한 결과값을 찾기가 어렵다. 게다가 동일 용어라 할지라도 용어 표기의 문제까지 고려해야하는 상황까지 생각한다면 표준화된 용어 시스템을 갖추지 못한 상황에서의 검색 결과는 신뢰도가 낮아질 수밖에 없다. 이에 본 연구에서는 용어 표기의 문제는 기존의 '송본 상한론의 한중 Code 비교¹⁾'라는 논문에서 제시

* 교신저자 : 박지하, 대구한의대학교 한의과대학 본초학교실,
053-770-2263, circle36@freechal.com

한 방법을 따른다는 가정하에 개념중심 용어의 선택이라는 관점만을 연구하고자 하였다. 다만 본 연구의 취지가 동일개념 용어들을 체계적으로 관리하여 효율적인 정보화를 추진하는 방법에 대한 기초연구에 있음을 밝힌다. 따라서 본 연구에서 동일개념 여부의 판단과 대표용어의 선정은 연구자들이 임의로 선택한 것이므로 한의학계 전문가들의 의견과 다를 수 있음을 밝힌다.

II. 연구방법

1. 연구대상의 선정

본 논문에서 『본초요략』을 선정하게 된 이유는 다음과 같다.

- 본초학 전문가의 저술이다.
- 저명한 저술을 근거로 정리된 본초서적이다.
- 효능용어 추출에 적합한 편집체계를 가지고 있다.
- 텍스트 파일을 연구에 활용할 수 있었다.

2. 연구수행 과정

본 연구는 Microsoft의 Access2003, MS-Word 2003, 한글2005로 진행하였다. 그리고 본초명, 효능용어 등에 대하여 각각의 ID를 부여하여 연구를 진행시키는 것이 원칙이나 소규모 실험적인 연구라는 것을 감안하여 별도의 ID체계를 구성하지 않고 문자열의 속성에 중복 불가능한 ID의 속성을 부여함으로써 ID체계를 대신하였다.

1) 본초요략 원문 테이블 구축

본초요략 원문을 하나의 필드에 모두 넣어 테이블을 구성하되, 본초명 필드를 추가하여 약재명으로 원문 전체를 볼 수 있도록 테이블을 만든다.

필드 이름	데이터 형식
본초번호	일련번호
본초명*	텍스트
전체원문	메모
완료여부	예/아니오

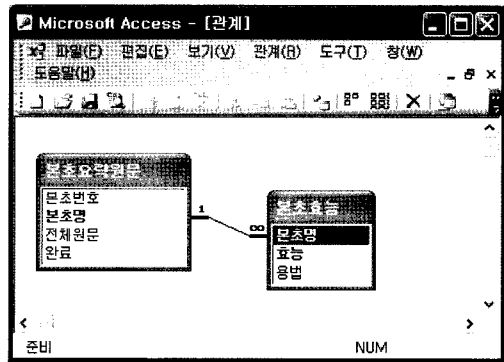
【표 1】 원문 테이블(*는 ID설정)

2) 문헌상 효능 입력

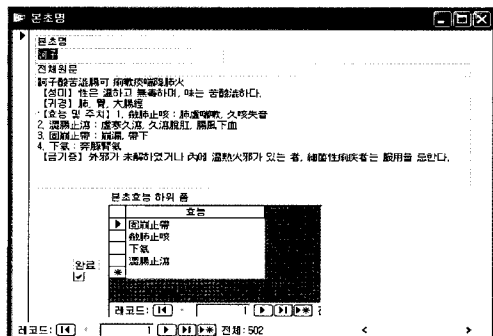
문헌상의 효능은 원문 테이블과 본초명으로 연결되는 효능 테이블을 만들어 원문 테이블과 1:∞ 관계를 형성한 뒤 입력폼을 만들어 문헌에 나타난 표현방식을 그대로 입력하였다.

필드 이름	데이터 형식
본초명*	텍스트
효능*	텍스트
용법	텍스트

【표 2】 효능 테이블(*는 ID설정)



【그림 1】 원문테이블과 효능테이블의 관계



【그림 2】 원문을 이용한 효능입력 폼

1) 이병욱 외 2인. 송본 상한론의 한증 Code 비교. 대한한의학 원전학회지. 2005. 18(4). pp.83-92.

3) 복합용어의 분리

본초별 효능용어를 살펴보면 단일 개념의 용어가 아닌 2개 이상의 개념이 합쳐진 형태로 구성된 경우가 많았다. 이러한 경우 'A', 'B', 'C' 세 개의 개념이 합쳐진 용어의 경우에는 ABC, ACB, BAC, BCA, CAB, CBA의 용어 조합이 가능하게 된다. 물론 관용적으로 사용되는 패턴이 있어서 모든 경우의 조합이 사용되지는 않는다. 그러나 컴퓨터를 이용한 정보화에서는 위와 같은 경우가 모두 다른 용어로 취급되기 때문에 정보화를 통한 한의학 지식관리의 효율성을 제고시키겠다는 취지에서 멀어지는 결과를 초래하게 될 것이다. 이러한 문제를 해결하기 위하여 복합용어를 분리하여 단일개념 용어로 분리하여 관리할 필요가 있다. 다만 원문에 가깝게 용어를 분리하는 과정에서 補中益氣와 같은 원문은 補中과 益氣로 분리하더라도 補中이라는 용어에 補脾와 健胃라는 개념이 동시에 있는 경우에는 補中과 益氣로 구분하지 않고 바로 補脾, 健胃, 益氣로 구분하여 저장하였다.

예) 散寒祛濕止痛→散寒, 祛濕, 止痛

4) 개념추출 및 대표용어 선정

단일개념으로 분리된 효능용어 역시 효능으로 약물을 검색할 경우 만족할 만한 검색결과를 가져다주지 못한다. 그 이유는 濕邪를 제거하는 의미를 가진 용어가 '去濕', '祛濕', '除濕' 등 여러 용어가 존재할 수 있다. 때문에 시스템 상에서 사용자가 원하는 결과를 제공하기 위해서는 복합개념으로 구성된 용어를 단일개념 용어로 구분하고 추가로 단일개념 용어 역시 대표용어를 중심으로 정리되어 활용되어야 한다. 뿐만 아니라 대표용어를 선정하여 사용할 경우에도 용어의 표기방법에 유의해야 한다.

예) 去濕, 祛濕, 除濕 → 去濕

5) 효능용어와 대표개념 연결

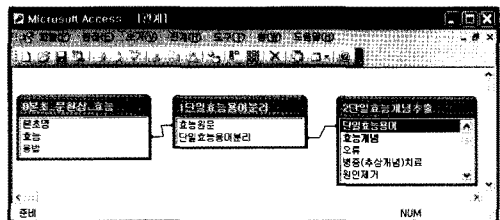
일반적으로 한의학 관련 산업에 종사하는 사람이거나 연구자일지라도 한의학계에서 선정한 표준용어를 모두 사용한다고 보장할 수 없다. 또한 그러한 상황을 개선하고자 한의학계에 종사하는 모든 사람들을 대상으로 표준용어를 추가적인 사업을 통하여 별도로

교육을 하는 방법 역시 효율적이지 못하다. 또한 입문자를 위한 교육에서도 표준용어만으로 한의학 지식을 전달하는데 한계가 있다. 한의학의 교육 특성상 古典文獻에 대한 의존도가 높아 古典文獻를 근거로 교육이 이루어지는 경우에는 표준용어만으로 교육을 진행시키기 어렵다. 따라서 이러한 경우에 표준용어와 비표준용어를 시스템적으로 연결해주는 방법을 통하여 古典文獻를 훼손시키지 아니하고 효과적으로 교육의 목적을 달성할 수 있을 것이다. 또한 고전문헌이 아니더라도 기존의 디지털화된 수많은 자원들을 활용할 때 표준용어를 모르는 상황에서도 효과적인 지식의 획득이 가능한 시스템을 구축해야 한다.

이에 일반적으로 사용하는 효능용어와 대표개념 용어를 연결하는 방법이 필요하다.

본 연구에서는 효능용어 목록을 단일개념 용어와 연결하고 다시 단일개념 대표용어와 연결하는 방법을 사용하였다.

예) 1단계) 去寒劫痰平喘 → 去寒, 劫痰, 平喘
 2단계) 去寒 → 去寒
 劫痰 → 去痰
 平喘 → 定喘



【그림 3】 효능용어와 대표개념용어의 관계

6) 효능용어의 한글표기 테이블 구축

한국에서는 일반적으로 사용자 입장에서 효능검색을 할 경우 한글음을 입력한 뒤 한자로 변환하는 과정을 거치게 된다. 이 과정에서 한글표기법 중 두음법칙 또는 同字異音漢字와 관련된 한자음 입력 오류가 많이 발생하며 이러한 경우 변환되는 한자 역시 별도의 유니코드값을 가지는 경우가 있기 때문에 정확한 검색조건을 입력하지 못하는 경우가 발생한다. 이에 두음법칙 적용의 오류를 고려하여 '祛風利濕'과

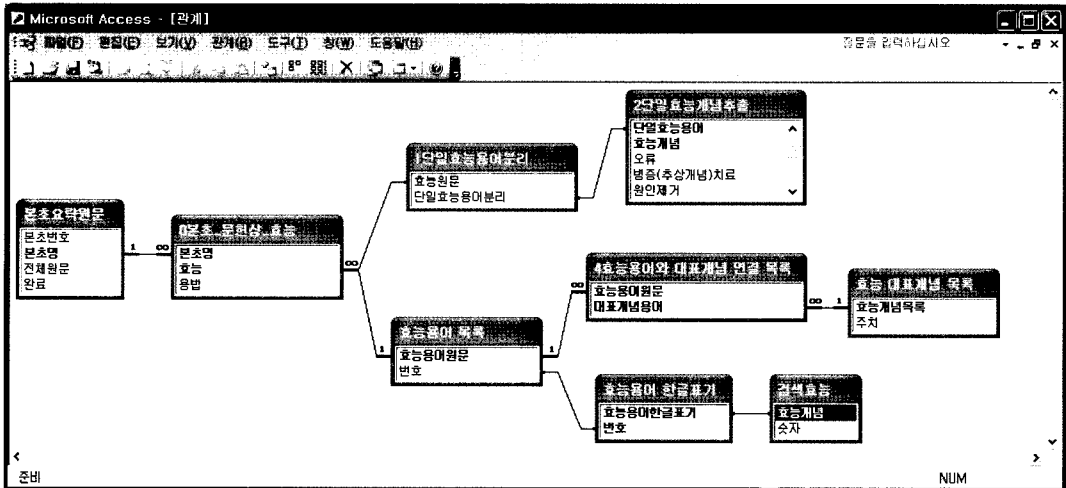
같은 용어의 경우 ‘거풍이습’과 ‘거풍리습’의 한글음 모두에 ‘祛風利濕’의 용어를 연결하고 ‘通便’의 경우는 ‘통변’과 ‘통편’ 모두 ‘通便’에 연결하여 검색하도록 테이블을 만들었다.

7) 효능을 이용한 본초 검색 폼

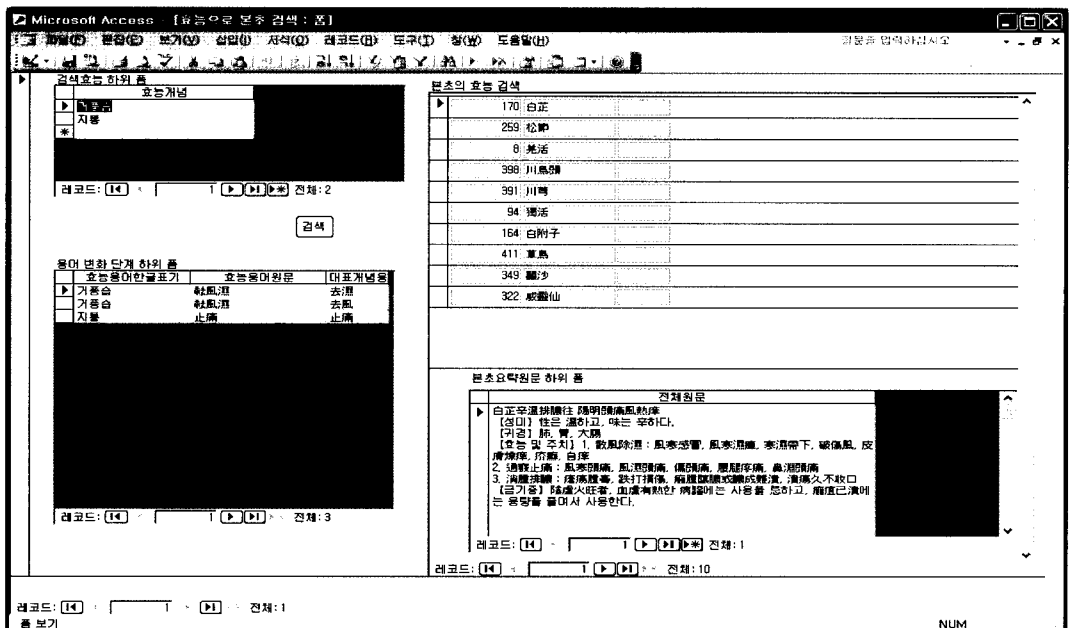
위의 단계를 거쳐 형성된 여러 테이블을 서로 연

결한 관계형 데이터베이스를 구성하면, 문헌에 기재된 효능을 그대로 옮긴 데이터 베이스보다 사용자에게 더욱 효과적인 사용환경을 제공하고 데이터베이스의 활용도가 높아질 것이다.

본 연구에서 작성한 테이블간의 관계도와 최종적으로 작성한 효능을 이용한 본초검색 폼은 다음과 같다.



【그림 4】 효능의 대표개념 용어를 이용한 본초검색 데이터베이스의 테이블간 관계



【그림 5】 효능의 대표개념 용어를 이용한 본초 검색 폼

III. 결 과

1. 기초정보

- 본초요약은 총 502개 약재의 효능을 표현하는데 1,626개의 용어를 사용하였다.
- 1,626개의 용어에는 858種의 용어가 사용되었다.
- 858種의 용어는 425種의 단일개념을 가진 용어의 조합으로 구성되어 있다.
- 425種의 단일개념 용어들은 232種의 개념이 서로 다른 용어로 표현된 것이다.

2. 한글표기-효능원문-단일개념-대표개념의 관계

아래의 도표는 효능을 한글로 입력한다는 가정하에 입력된 효능용어로부터 대표개념으로 연결되는 상황을 나타낸 것이다.

한글표기	효능원문	단일효능	대표개념
강근건골	强筋健骨	强筋	强筋
		健骨	强骨
강근골	强筋骨	强筋	强筋
		强骨	强骨
강기거담	降氣祛痰	降氣	降氣
		祛痰	去痰
강기거담지해	降氣去痰止咳	止咳	止咳
		降氣	降氣
		去痰	去痰
강기지구	降氣止嘔	降氣	降氣
		止嘔	止嘔
강기지해평천	降氣止咳平喘	平喘	定喘
		止咳	止咳
		降氣	降氣
강심	强心	强心	强心
강압	降壓	降血壓	降血壓
강역조중	降逆調中	降逆	降氣

강역지구	降逆止嘔	調中	和胃
		止嘔	止嘔
강역평천	降逆平喘	平喘	定喘
		降逆	降氣
강오슬	强腰膝	强膝	强膝
		强腰	强腰
강장근골	强壯筋骨	强筋	强筋
		强骨	强骨
강저혈압	降低血壓	降血壓	降血壓
강혈당	降血糖	降血糖	降血糖
강혈압	降血壓	降血壓	降血壓
강화리뇨소증	降火利尿消腫	消腫	消腫
		降火	降火
		利尿	利尿
강화이뇨소증	降火利尿消腫	消腫	消腫
		降火	降火
		利尿	利尿
중략		
화적살충	化積殺蟲	化積	化積
		殺蟲	驅蟲
화중개위	和中開胃	和中	和胃
		開胃	開胃
화중소식	和中消食	和中	和胃
		消食	消食
화중지구	和中止嘔	和中	和胃
		止嘔	止嘔
화해퇴열	和解退熱	和解	和解
		退熱	清熱
화혈산어	和血散瘀	散瘀	化瘀
		和血	利血
활락지통	活絡止痛	止痛	止痛
		活絡	通血脈
활장통편	滑腸通便	滑腸	滑腸
		通便	通便
활혈거어	活血祛瘀	活血	活血
		祛瘀	去瘀
활혈거어지통	活血祛瘀止痛	活血	活血
		祛瘀	去瘀

		止痛	止痛
활혈거풍	活血祛風	祛風	去風
		活血	活血
활혈거풍지통	活血去風止痛	活血	活血
		去風	去風
		止痛	止痛
활혈산어	活血散瘀	散瘀	化瘀
		活血	活血
활혈산어소증	活血散瘀消腫	活血	活血
		散瘀	化瘀
		消腫	消腫
활혈산풍정통	活血散風定痛	散風	去風
		活血	活血
		定痛	止痛
활혈소어	活血消瘀	消瘀	化瘀
		活血	活血
활혈소증	活血消腫	消腫	消腫
		消腫	消腫
활혈조혈	活血調血	調血	利血
		活血	活血
활혈지통	活血止痛	止痛	止痛
		活血	活血
활혈지혈	活血止血	活血	活血
		止血	止血
활혈통경	活血通經	通經	通經
		活血	活血
활혈통경활락	活血通經活絡	活絡	通血脈
		活血	活血
		通經	通經
활혈통맥	活血通脈	活血	活血
		通脈	通血脈
활혈통비	活血通痺	活血	活血
		通痺	除痺
활혈통체	活血通滯	通滯	導滯
		活血	活血
활혈파어	活血破瘀	活血	活血
		破瘀	去瘀
활혈해독	活血解毒	解毒	解毒
		活血	活血
활혈행경통락	活血行經通絡	活血	活血

		行經	通經
		通絡	通血脈
활혈행어	活血行瘀	活血	活血
		行瘀	行瘀
활혈화어	活血化瘀	活血	活血
		化瘀	化瘀
회양구역	回陽救逆	救逆	救逆
		回陽	補陽
회양통맥	回陽通脈	通脈	通血脈
		回陽	補陽

【표 3】 용어의 한글표기와 대표용어의 관계

3. 비교결과

去風(거풍), 利水(이수), 利水(리수), 補氣(보기)라는 용어를 사용하여 본초요약에 기재된 효능을 검색함에 있어서 본초요약의 원문을 그대로 입력한 경우와 개념을 중심으로 여러 용어의 관계를 정립한 시스템 사이의 결과를 비교하면 다음과 같다. 이때의 검색조건에서 원문을 그대로 사용한 경우에는 검색조건을 “*조건*”의 형태로 사용하였고, 대표개념 용어를 활용한 경우에는 한글로 ‘거풍’, ‘이수’, ‘보기’와 완전 일치하는 경우만을 검색하였다.

DB형태	원문 그대로 사용한 경우				대표개념 용어를 활용한 경우		
	去風	利水(이수)	利水(리수)	補氣	거풍	이수	보기
검색조건							
결과건수	17	1	54	3	84	63	18
결과	羌活, 龍葵, 蔓荊子, 防風, 防風, 白殭蠶, 浮萍, 沙參, 虎耳草, 荊芥, 椿白皮, 穿山龍, 天南星, 檉柳, 辛夷, 細辛, 蟬蛻	海藻	輕粉, 昆布, 郁李仁, 玉米鬚, 五加皮(南五加皮), 桑枝, 桑白皮, 商陸, 桔梗, 路路通, 大腹皮, 冬瓜子, 冬瓜皮, 冬葵子, 燈心草, 綠豆, 螻蛄, 鯉魚, 馬兜鈴, 馬鞭草, 麻黃, 文蛤, 半邊蓮, 半枝蓮, 防己, 白屈菜, 白石英, 白朮, 茯苓, 茯神, 萆麻子, 菝葜, 芫蔚子, 黑大豆, 黃芪, 琥珀, 香薷, 海帶, 通草, 澤漆, 澤瀉, 車前子, 車前, 秦艽, 地膚子, 知母, 齊菜, 田螺, 豬苓, 薏苡仁, 榆白皮, 鴨跖草, 蘇合香, 石葦	白朮, 白朮, 西洋參	降香, 羌活, 苦參, 檉寄生, 穀精草, 狗脊, 菊花, 金銀花, 老鸛草, 路路通, 凌霄花, 大風子, 獨活, 絡石藤, 露蜂房, 曼陀羅子, 蔓荊子, 木賊, 燕黃, 密蒙花, 薄荷, 胖大海, 防己, 防風, 白殭蠶, 白附子, 白鮮皮, 白芷, 白蒺藜, 白花蛇, 附子, 浮萍, 草薢, 沙參, 蛇床子, 蛇蛻, 桑寄生, 桑葉, 桑枝, 徐長卿, 旋覆花, 蟬蛻, 細辛, 松節, 辛夷, 夜交藤, 野菊花, 五加皮(南五加皮), 烏梢蛇, 龍葵, 牛蒡子, 威靈仙, 萎陵菜, 淫羊藿, 忍冬藤, 蠶沙, 前胡, 檉柳, 地膚子, 秦艽, 蒼耳子, 蒼朮, 川芎, 天南星, 天麻, 穿山龍, 川烏頭, 茜草根, 青葙子, 草烏, 椿白皮, 臭梧桐, 側柏葉, 巴戟天, 海桐皮, 海風藤, 香加皮, 荊芥, 虎骨, 虎耳草, 黑大豆, 稀莨草, 皂角刺, 菝葜	麻黃, 桑白皮, 商陸, 檳榔, 蔓麻子, 浮萍, 茯神, 茯苓, 番瀉葉, 白朮, 白石英, 白屈菜, 防己, 半枝蓮, 輕粉, 燈心草, 昆布, 桔梗, 路路通, 大腹皮, 冬瓜子, 半邊蓮, 冬葵子, 文蛤, 綠豆, 螻蛄, 鯉魚, 馬兜鈴, 馬鞭草, 旋覆花, 冬瓜皮, 海藻, 桑枝, 澤蘭, 澤瀉, 澤漆, 通草, 車前, 海帶, 秦艽, 香薷, 琥珀, 黃芪, 黑大豆, 芫蔚子, 菝葜, 巴豆, 薏苡仁, 葶藶子, 蘇合香, 鴨跖草, 五加皮(南五加皮), 玉米鬚, 車前子, 榆白皮, 石葦, 豬苓, 赤小豆, 田螺, 齊菜, 知母, 地膚子, 郁李仁	紫河車, 黨參, 大棗, 白朮, 浮小麥, 山藥, 西洋參, 狗肉, 人參, 甘草, 鍾乳石, 枳椇子, 陳倉米, 太子參, 黃芪, 黃精, 荔枝, 龍眼肉

【표 4】 대표 효능용어 사용 여부에 따른 검색 결과의 차이

IV. 결론

한의학의 발전과 국민보건에 기여하기 위한 여러 가지 연구 방법 중 한의학 지식의 정보화에 관한 연구 역시 중요한 부분을 차지한다. 그 중 용어표준화에 관한 연구는 한의학 지식의 체계화 및 대중화의 시발점이 될 수 있다.

본 연구에서는 한의학 문헌의 디지털화 작업만으로 얻을 수 있는 정보화의 성과와 용어표준화 단계를 거친 이후에 얻을 수 있는 성과를 비교하였다. 그 결과 디지털화 작업만으로 去風(거풍), 利水(이수), 利水(리수), 補氣(보기)라는 효능을 가진 약제를 검색했을 경우에는 각각 17건, 1건, 54건, 3건의 약제가 검색되었고 개념중심 용어표준화 과정을 거친 이후의 단계에서는 거풍, 이수, 보기라는 조건으로 약제를 검색했을 경우 각각 84건, 63건, 18건의 약제가 검색되어 검색효율에서 큰 차이를 보였다.

따라서 기존의 정보화 사업을 좀 더 발전시켜 개념중심 용어표준화 단계로 접어들어 선진화된 지식 체계를 구축하는 것이 필요함을 인지하게 되었다.

현재까지 한의학 정보화의 과정이 주로 한의학 전문 연구자들을 수혜자로 하는 한의학 문헌의 디지털화 사업에 치중되어 있었다면, 앞으로는 한의학 전문 연구자뿐만 아니라 한의학 관련 사업 종사자 및 주변학문 연구자와 일반인들의 요구까지 포괄할 수 있는 대중성 높은 정보화 사업이 추진되어야 할 것이다. 그리고 그 첫 단추는 용어표준화가 되어야 할 것이다.

<단행본>

1. 서부일. 본초요략. 대구. 벤엘기획. 2004.

參考文獻

<논문>

1. 이병욱, 신상우, 김은하. 송본 상한론의 한중 Code 비교. 대한한의학원전학회지. 2005. 18(4).
2. 김혜선. 통합의학언어 시스템(UMLS)의 구성 및 적용에 대한 고찰. 정보관리학회지. 2001. 32(2).