



Original Article

한국치위생학회지 게재논문의 피인용수에 영향을 미친 요인

전세정^{id}

대전과학기술대학 치위생과

Factors affecting the number of citations in papers published in the Journal of Korean Society of Dental Hygiene

Se-Jeong Jeon^{id}

Department of Dental Hygiene, Daejeon Institute of Science and Technology

Corresponding Author: Se-Jeong Jeon, Department of Dental Hygiene, Daejeon Institute of Science and Technology, 101 Hyecheon-ro, Seogu, Daejeon, 35408, Korea. Tel : +82-42-580-6351, Fax : +82-42-580-6301, E-mail : sejeong0084@gmail.com

ABSTRACT

Objectives: The purpose of this study was to analyze the factors that affected the number of citations for articles published in the Journal of Korean Society of Dental Hygiene based on previous studies. **Methods:** Information on papers including the number of citations was collected using a web crawling technique. The effect of the number of author keywords, the number of Medical Subject Headings (MeSH) keywords, MeSH match rate, abstract word count and keyword-abstract ratio on the number of citations was analyzed by multiple regression analysis. **Results:** The use of the MeSH keyword did not have a significant effect on the number of citations. Among the other factors, only the keyword-abstract ratio was statistically significant. **Conclusions:** Select a topic of constant interest in the field, write the title in detail using colons or asterisks if necessary, and do not repeat the words used in the title in keywords. Select specific keywords deeply related to the topic. In particular, choice words or phrases that are frequently used in the abstract. If the MeSH keyword selection contradicts the previous strategies, boldly give up the MeSH keyword.

Key Words: Citation counts, Impact factors, Journal of Korean Society of Dental Hygiene, Medical subject headings, Strategies for increasing citation

색인: 의학주제표목, 피인용수, 피인용지수, 한국치위생학회지

서론

학술지의 영향력을 평가하는데 있어서 Impact factor는 매우 주요하게 사용되는 지표이다. Impact factor는 1955년 Garfield에 의해 고안되었는데, 학술지에 실린 논문이 다른 논문에 인용된 횟수 즉 피인용수를 학술지에 실린 논문의 총 수로 나누어 매년 집계된다[1]. 높은 Impact factor가 반드시 논문의 질을 담보하는 것은 아니지만, 연구성과를 게재 학술지의 Impact factor로 평가받는 연구자들이 투고 대상 학술지를 선정함에 있어서 Impact factor를 우선적으로 고려하는 것은 분명하며, 이는 높은 Impact factor → 양질의 논문 투고 → 피인용수 증가 → Impact factor 상승으로 이어지는 선순환 고리의 시작점이 된다. 따라서 각 학술지의 편집위원회는 Impact factor를 높이기 위해 많은 노력을 기울이고 있다.

한국치위생학회지(Journal of Korean Society of Dental Hygiene)는 한국치위생학회에서 발행하는 공식 학술지로, 치위생 및 구강건강 관련 다양한 주제를 다루고 있으며 2011년 한국연구재단 등재지로 선정되어 치의학분야에서 가장 많은 논문이 게재되고 있다[2]. 한국치위생학회지 역시 Impact factor를 증가시키기 위하여 다양한 노력을 하고 있는데, 그 일환으로 2016년 12월 투고규정을 개정하여 저자 키워드(Author keywords)를 선정함에 있어 MeSH(Medical subject headings)에 수록된 용어를 사용하도록 요구하고 있다[3]. 저자 키워드는 논문

을 대표하는 핵심 주제어로 저자에 의해 선택된 몇 개의 단어 혹은 구를 의미하며[4], MeSH는 미국국립의학도서관(National Library of Medicine, NLM)에서 만든 의학 주제명 표목(Medical subject headings)을 말한다. 많은 의학분야 연구자들은 인용 자료를 검색하기 위하여 PubMed, Scopus, Web of Science 등의 데이터베이스 검색엔진을 이용하고 있는데[5], 이때 MeSH 키워드를 사용하면 보다 정확하고 효율적인 검색이 가능하다. 따라서 MeSH 키워드를 사용하였을 때, 논문이 검색될 가능성이 높아지고 이는 곧 피인용수의 증가로 이어질 수 있다[6].

이전 연구에서 김[7]은 2015년까지 한국치위생학회지에 게재된 논문의 저자 키워드를 대상으로 MeSH 일치율을 조사하여 보고한 바 있다. 그러나, MeSH 수록 용어의 사용이 피인용수에 미친 영향에 관한 고찰이 결여되었다. 또 2016년 투고규정 개정 이후에 MeSH 일치율이 어떻게 변화하였는지에 관한 추가 연구가 없었다. 따라서 본 연구에서는 한국치위생학회지에 최근 호(2021년 4월) 까지 게재된 논문을 대상으로 MeSH 일치율을 조사하고, 2016년 투고규정 개정이 MeSH 일치율에 미친 영향과 MeSH 일치율이 피인용수에 미친 영향을 조사해 보고자 하였다. 더불어 논문의 피인용수에 영향을 미치는 요인에 관한 기존 연구들[8-12]을 고찰하고, 몇 가지 요인을 선정하여 분석을 진행하였다.

연구방법

1. 연구대상 및 방법

연구에 필요한 데이터를 수집하기 위하여 웹 크롤링 기법(Web crawling techniques)이 사용되었다. 웹 크롤링 기법은 웹 페이지에서 자동으로 필요한 데이터를 빠르고 효율적으로 추출하는 기술이다[13]. 한국치위생학회지에 게재된 논문이 2015년 기준으로 이미 1,000여편에 달하였기 때문에 웹 크롤링 기법이 연구에 필요한 시간을 매우 단축시켜줄 것으로 예상되었다. 프로그램은 오픈 소스 기반 프로그래밍 언어인 파이썬(Python)을 이용하여 작성되었다.

한국학술지인용색인(Korea Citation Index, KCI)은 한국연구재단(National Research Foundation of Korea, NRF)에서 제공하는 인용색인이다. 국내 학술지 정보, 논문 정보(원문) 및 참고문헌을 데이터베이스화 하여 논문 간 인용관계를 분석 및 제공한다[14]. KCI에서 한국치위생학회지를 검색하고, 웹 크롤링 기법을 이용하여 2007년부터 2021년 4월까지 게재된 총 1,368편의 논문의 제목, 저자, 저자 키워드, 초록, KCI 피인용수를 수집하였다. 영문 저자 키워드가 없거나, 영문 초록이 없는 6편을 제외하고 총 1,362편의 논문이 연구 대상으로 최종 선정되었다.

MeSH 수록집은 미국국립의학도서관 웹 사이트(<https://nlm.nih.gov>)[15]에서 XML 파일 형태로 내려 받아 분석에 용이하도록 가공하였다. 2021년 최신판 기준으로 총 29,917개의 주제가 수록되어 있었다.

2. 연구도구

분석할 요인들은 기존 연구들을 바탕으로 하되, 본 연구에서는 직접 작성한 자동화 프로그램을 사용하였기에 수작업으로 진행되었던 기존 연구에 비하여 분석이 용이하였으므로 각 변수를 다시 설계하였다. 예를 들어, 김[7]의 연구를 포함한 MeSH 일치율을 조사한 대부분의 기존 연구에서는 전체 논문의 저자 키워드를 한꺼번에 나열하고 MeSH 키워드와 일치하는 수를 세어 학술지의 MeSH 일치율을 산출하였으나, 본 연구에서는 컴퓨터가 자동으로 각 논문마다의 일치율을 정확하게 계산해 주었기에 MeSH 키워드 사용이 KCI 피인용수에 미치는 영향을 통계적으로 분석할 수 있었다.

1) 저자 키워드의 개수와 초록의 단어 수

각 논문마다 영문 저자 키워드의 개수와 영문 초록의 단어 수가 추출되었다.

2) MeSH 일치율

각 논문의 영문 저자 키워드 중 MeSH 수록집에 포함된 키워드의 개수를 세어 MeSH 키워드 개수로 하고, 이를 전체 키워드에 대한 비율로 계산하여 MeSH 일치율로 하였다. 이때 대/소문자의 일치여부는 무시하였다.

$$\text{MeSH 일치율} = \text{MeSH 키워드 개수} / \text{전체 키워드 개수}$$

3) 키워드-초록 비율

각 논문의 영문 초록에서 키워드의 등장 빈도를 계산하였다. 이때 대/소문자의 일치여부는 무시하였으며, 키워드가 단어 단위로 정확히 일치할 경우에만 등장으로 처리하였다. 예를 들어 Child 키워드가 Children의 일부로 등장할 경우 이는 등장으로 처리하지 않았다. 이를 바탕으로 아래와 같이 키워드-초록 비율이 계산되었다.

$$\text{키워드-초록 비율} = \frac{\sum(\text{각 키워드의 등장 횟수} \times \text{키워드의 단어 수})}{\text{영문 초록의 전체 단어 수}}$$

예를 들어 100 단어로 이루어진 영문 초록에서 Oral health 키워드가 세 번, Dentures 키워드가 등장할 경우 키워드-초록 비율은 0.07로 계산되었다.

4) 논문 연령

논문이 오래될수록 피인용수는 누적하여 증가하기 때문에 논문 연령은 피인용수에 유의한 영향을 미칠 것으로 기대되었다. 통계분석에 투입하기 위하여 2021년을 기준으로 논문의 연령이 계산되었다.

$$\text{논문 연령} = 2021 - \text{논문의 게재 연도}$$

3. 자료분석

수집 및 가공된 데이터는 SPSS® Statistics 20(IBM®, New York, USA)을 이용하여 통계 분석하였다. 투고규정 전·후의 MeSH 일치율 변화를 확인하기 위하여 독립표본 t-검정(Independent t-test)를 수행하였고, 논문의 피인용수에 영향을 미치는 요인을 특정하기 위하여 다중회귀분석(Multiple regression analysis)을 실시하였다. 다중회귀분석에서 모든 변수는 동시에 투입되었으며, 유의 수준 $\alpha=0.05$ 로 하였다.

연구결과

1. MeSH 일치율 변화 추이와 2016년 투고규정 개정의 영향

연도에 따른 MeSH 일치율의 변화 추이를 <Fig. 1> 및 <Table 1>에 나타냈다.

2016년 12월 투고규정 개정 이전에는 MeSH 일치율에 큰 변화가 없었으나, 개정 이후에 MeSH 일치율이 0.16 ± 0.007 에서 0.30 ± 0.014 로 유의하게 증가한 것으로 나타났다($p < 0.001$)(<Fig. 2, Table 2>).

2. 피인용수에 영향을 미치는 요인

논문의 피인용수를 종속변수로 하고, 저자 키워드의 개수, 초록의 단어 수, MeSH 키워드 개수, MeSH 일치율, 키워드-초록 비율, 논문 연령을 독립변수로 하여 수행한 다중회귀분석을 수행하였을 때 F값은 46.142, F검정의 유의 확률은 0.001 미만으로 나타나 회귀모형이 적합한 것으로 나타났다. Durbin-Watson 값은 1.846으로 2에 가깝게 나타나 자기 상관 문제가 없었으며, 설명력을 의미하는 R^2 값은 0.170으로 나타났다. 분석 요인 중 키워드-초록 비율과 논문 연령의 영향력이 통계적으로 유의하였고($p < 0.001$), 다른 요인의 영향력은 유의하지 않은 것으로 나타났다($p > 0.05$)(<Table 3>).

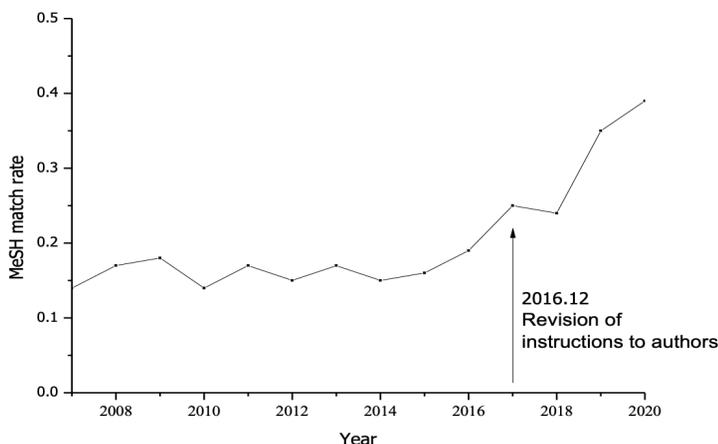


Fig. 1. Changes in MeSH match rate by year

Table 1. Changes in MeSH match rate by year

Year	Average MeSH match rate	Samples	Standard deviation
2007	0.14	34	0.205
2008	0.17	58	0.218
2009	0.18	83	0.244
2010	0.14	96	0.198
2011	0.17	93	0.228
2012	0.15	119	0.217
2013	0.17	122	0.229
2014	0.15	116	0.194
2015	0.16	126	0.190
2016	0.19	119	0.229
2017	0.25	103	0.249
2018	0.24	93	0.223
2019	0.35	94	0.295
2020	0.39	88	0.323
2021	0.19	18	0.178

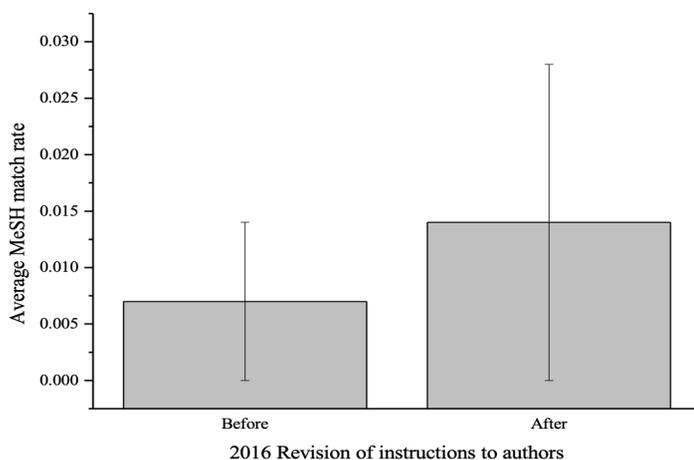


Fig. 2. Difference in match rate before and after 2016 revision of instructions to authors

Table 2. Difference in match rate before and after 2016 revision of instructions to authors

2016 revision of instructions to authors	N	Average	Standard deviation	t	<i>p</i> *
Before	966	0.16	0.007	-9.830	< 0.001
After	396	0.30	0.014		

*by independent t-test

Table 3. Multiple regression analysis results to analyze factors affecting the number of citations

Model	Unstandardized coefficients		Standardized coefficients	t	<i>p</i> *
	B	SE	β		
(Constant)	-0.132	0.840		-0.157	0.875
Number of author keywords	0.106	0.183	0.020	0.578	0.563
Number of MeSH keywords	-0.851	0.563	-0.158	-1.510	0.131
Mesh match rate	2.439	2.171	0.111	1.124	0.261
Abstract word count	-0.001	0.002	-0.011	-0.383	0.702
Keyword-abstract ratio	9.483	2.500	0.099	3.793	< 0.001
Article age	0.592	0.044	0.406	13.446	< 0.001

F = 46.142, *p* < 0.001, R² = 0.170, Adjusted R² = 0.166, Durbin-Watson = 1.846

*by multiple regression analysis

총괄 및 고안

Impact factor는 결국 게재논문의 평균 피인용수를 의미하기 때문에, 논문 각각의 피인용수를 늘리는 것이 가장 중요하다. 이에 따라 논문의 피인용수에 영향을 미치는 다양한 요인에 관한 연구들이 수행되었다.

Rostami 등[8]은 2014년 연구에서 논문 제목에 쌍점(Colon) 혹은 붙임표(Hyphen)를 사용하거나 저자 키워드로 제목에 없는 단어를 선정하는 것이 피인용수를 유의하게 증가시켰다고 보고하였다. Borsuk 등[9]은 저자의 성별, 제1언어, 저자 수 등 저자요인의 영향이 피인용수에 미친 영향을 분석한 결과, 저자 수만이 유의하게 나타났다고 하였다. 2013년 Falagas 등[10]은 일변량 분석(Univariate analysis)에서 게재 학술지의 Impact factor, 저자 수, 논문의 길이, 연구 디자인(중재/관찰연구 혹은 전향적/후향적 연구), 제목 및 초록의 단어 수, 참고문헌의 수가 피인용수와 모두 연관된 것으로 나타났으나 다변량 분석(Multivariate analysis)에서는 논문의 길이만이 통계적으로 유의하게 나타났다고 보고하였다. 2016년 Uddin과 Khan[11]은 저자 키워드와 관련된 요인들이 피인용수에 미친 영향을 중점적으로 조사하였는데, 키워드의 성장(분야 내에서의 증가 혹은 감소세), 키워드의 수, 중심성이 피인용수와 양의 상관관계를 보였으며, 새로운 키워드의 비율은 피인용수와 음의 상관관계를 보였다고 하였다. 즉, 해당 분야에서 꾸준히 관심 받는 주제의 중심어 위주로 다수의 키워드를 선정하는 것이 유리함을 시사하였다. 2017년 Sohrabi와 Iraj[12]는 키워드-초록 비율(Abstract ratio)에 관한 개념을 제시하고 이것이 피인용수와 양의 상관관계가 있음을 보고하였다. 따라서 연구자가 초록에서 키워드를 반복 사용함으로써 피인용수를 증가시킬 수 있다고 하였다. 이에 본 연구에서는 위와 같은 선행연구에서 제안된 몇 가지 요인들과 MeSH 일치율이 한국치위생학회지의 피인용수에 미친 영향을 조사하고자 하였다.

연구 결과를 보면, MeSH 일치율은 2016년 투고규정 개정 이후에 급격하게 증가하였으며 2020년에는 약 40%에 달하였으나, 이전 연구에서 보고된 대한응급의학회지의[16] 50.7%나 대한간호학회지의[17] 51.6%에 비하면 저조한 수준이다. 그러나 투고규정 개정 이후에도 MeSH 용어 사용이 전혀 증가하지 않았던 사상체질의학회지[18]에 비하면 나름의 성과를 거두고 있는 것으로 보인다. 투고규정 개정 이후에 아직까지는 많은 시간이 흐르지 않았기에, 후속 연구에서 보다 장기간의 추이를 조사할 필요가 있을 것으로 생각된다.

그런데, 본 연구에서 MeSH의 일치율이나 MeSH 키워드 개수가 피인용수에 유의한 영향을 미치지 못한 것으로 나타났다. 저자는 이를 투고규정을 준수하기 위하여 연구 주제와 연관성이 떨어지거나 너무 포괄적인 MeSH 키워드가 선정되었기 때문으로 추정하였는데, 예를 들어 'Dental hygienist', 'Oral health' 등이 관습적으로 사용되고 있었다. 그러나 이는 저자의 단순 추정일 뿐 후속 연구가 더 필요하다.

추가로 선정된 요인들 중에서는, 논문 연령과 키워드-초록 비율만이 유의한 것으로 나타났으며, 선행연구[10,12]의 결과와 일치하였다.

그러나 본 연구에서 사용된 피인용수 데이터가 KCI에서 제공하는 것으로 국내 학술지에서의 피인용 횟수만을 의미하기에, MeSH 키워드 사용 전략의 주 타겟이 되는 해외 학술지에 얼마나 인용이 되었는지는 알 수 없었으며, 이는 본 연구의 한계이기도 하다. 따라서 추후 해외 학술지 인용 횟수를 조사하여 추가적인 연구가 수행되어야 할 필요가 있다.

마지막으로, 선행연구와 본연구의 결과를 종합하여 다음과 같이 제언하는 바이다. 첫째, 제목은 필요하다면 쌍점 혹은 붙임표를 사용하여 구체적으로 작성하는 것이 좋다. 둘째, 해당 분야에서 꾸준히 관심 받는 주제를 선정하는 것이 유리하다. 셋째, 키워드에서 제목에 사용된 단어를 반복하기 보다는 대체할 수 있는 용어를 사용할 것을 권장한다. 넷째, 주제와 깊이 연관된 구체적인 키워드를 선정하고, 특히 초록에서 많이 사용되는 단어 혹은 구를 선택하는 것이 좋다. 다섯째, 만약 MeSH 키워드 선정이 앞의 전략들과 배치된다면 과감하게 MeSH 키워드를 포기할 것을 권장한다.

결론

본 연구는 한국치위생학회지 게재논문을 대상으로 논문의 피인용수에 영향을 미치는 요인을 분석하고, 피인용 가능성을 높이기 위한 전략을 제시하는 데에 그 목적을 두었다. 선행연구를 바탕으로 후보 요인들을 선정하고, 다중회귀분석을 수행한 결과 다음과 같은 결과를 얻었다.

1. 저자 키워드의 개수와 초록의 단어 수, MeSH 키워드의 개수나 일치율은 피인용수에 유의한 영향을 미치지 못하였다.
2. 논문의 연령과 키워드-초록 비율은 피인용수에 유의한 영향을 미쳤다.

따라서 논문의 피인용수를 높이기 위하여, 앞서 제시한 5가지 전략을 반영하도록 이후의 연구자들에 적극 권고하는 바이다.

Conflicts of Interest

The author declared no conflict of interest.

References

1. Garfield E. Citation indexes for science. A new dimension in documentation through association of ideas. *Science* 1955;122(3159):108-11. <https://doi.org/10.1126/science.122.3159.108>
2. Jang JH, Won BY, Jang GW, Kim SA, Oh SH, Kim YJ, et al. Trend analysis of research in the journal of Korean society dental hygiene from 2001 to 2015. *J Korean Soc Dent Hyg* 2017;17(4):693-704. <https://doi.org/10.13065/jksdh.2017.17.04.693>
3. Korean Society of Dental Hygiene. 2016 JKSDH submission guidelines revision announcement [Internet]. *J Korean Soc Dent Hyg*. [cited 2021 Jun 11]. Available from: http://www.ksdh.or.kr/bbs/board.php?bo_table=sub_06_01&wr_id=150.
4. Kwon SY. Characteristics of interdisciplinary research in author keywords appearing in Korean journals. *Malaysian J Libr Inf Sci* 2018;23(2):77-93. <https://doi.org/10.22452/mjlis.vol23no2.5>
5. Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *FASEB J* 2008;22(2):338-42. <https://doi.org/10.1096/fj.07-9492LSF>
6. Dhammi IK, Kumar S. Medical subject headings (MeSH) terms. *Indian J Orthop* 2014;48(5):443-4. <https://doi.org/10.4103/0019-5413.139827>
7. Kim YJ. Comparison of author key words and medical subject heading terms in the journal of Korean society dental hygiene from 2001 to 2015. *J Korean Soc Dent Hyg* 2018;18(6):1047-55. <https://doi.org/10.13065/jksdh.20180090>
8. Rostami F, Mohammadpoorasl A, Hajizadeh M. The effect of characteristics of title on citation rates of articles. *Scientometrics* 2014;98(3):2007-10. <https://doi.org/10.1007/s11192-013-1118-1>
9. Borsuk RM, Budden AE, Leimu R, Aarssen LW, Lortie CJ. The influence of author gender, national language and number of authors on citation rate in ecology. *Open J Ecol* 2009;2(1):25-8. <https://doi.org/10.2174/1874213000902010025>
10. Falagas ME, Zarkali A, Karageorgopoulos DE, Bardakas V, Mavros MN. The impact of article length on the number of future citations: a bibliometric analysis of general medicine journals. *PLoS One* 2013;8(2):e49476. <https://doi.org/10.1371/journal.pone.0049476>
11. Uddin S, Khan A. The impact of author-selected keywords on citation counts. *J Informetr* 2016;10(4):1166-77. <https://doi.org/10.1016/j.joi.2016.10.004>
12. Sohrabi B, Iraj H. The effect of keyword repetition in abstract and keyword frequency per journal in predicting citation counts. *Scientometrics* 2017;110(1):243-51. <https://doi.org/10.1007/s11192-016-2161-5>
13. Olston C, Najork M. *Web crawling*. Boston: Now Publishers Inc; 2010: 1.
14. National research foundation of Korea. Introduction of Korea citation index [Internet]. National research foundation of Korea. [cited 2021 Jun 11]. Available from: <http://www.kci.go.kr/kciportal/aboutKci.kci>.
15. National library of medicine. Download MeSH data [Internet]. National Library of Medicine. [cited 2021 July 01]. Available from: <http://www.nlm.nih.gov/databases/download/mesh.html>.
16. Cho JS, Lee MJ. Coincidence analysis of key words and MeSH terms in the journal of the Korean society of emergency medicine. *J Korean Soc Emerg Med* 2009;20(6):722-8.
17. Jeong GH, Ahn YM, Cho DS. Coincidence analysis of keywords of the journal of Korean academy of nursing with MeSH. *J Korean Acad Nurs Adm* 2005;35(7):1420-5.
18. Kim YY, Park HJ, Lee SW, Yoo JH. Comparison of keywords of the journal of Sasang constitutional medicine with MeSH terms. *J Sasang Constit Med* 2013;25(1):34-42.