

## Chapter 11

# Bibliometrics

제 11 장 / 계량서지학

### 1. 계량서지학이란? Definition of bibliometrics

우리나라의 연구비 지원 기관이 1990년대 초부터 SCIE 인용도지수 높은 학술지에 논문을 게재한 학자에게 연구비 지원에서 우선권을 주기 시작하면서부터 의학계에서도 이런 물결을 피해갈 수 없었다. 정부의 연구비 지원액 급증과 이러한 학술지 정책으로 인하여 1990년대 중반부터 의학계에서 SCIE 등재 논문수가 급속하게 증가하였다. 정부가 계량서지학을 학자 평가에 반영하기 시작하였기 때문이다. 이 계량서지학은 어느 서지 정보가 얼마나 유효하게 유통되고 있고 다른 저자나 다른 서적과 얼마나 관계를 맺고 있는지, 또한 그 내용은 어떠한 것을 다루는 지를 분석, 측정하는 학문이다. 문헌과 정보를 연구하고 측정하는 방법을 총괄하는 학문분야이다.<sup>1</sup> 우리 의학계에서 계량서지학 연구를 보면 1996-1997년의 우리나라 의학 분야 SCIE 논문 생

산량이 전 세계의 1.02%를 차지한다는 보고가 있었고,<sup>2</sup> 1988년부터 1999년 사이 Medline에 실린 우리나라 학술지를 분석하여 60%의 논문이 상위 5개 대학에서 나왔고, 영상의학이 가장 국제화된 분야이며 내과학이 가장 빠르게 Medline 논문수가 증가하고 있다고 하였다.<sup>3</sup> 2001년부터 시작한 우리나라 의학학술지 인용색인사업으로 KoMCI 결과가 나오고, 이어서 Web of Science에 해당하는 KoMCI Web 및 SCIE JCR (Journal Citation Reports)에 해당하는 KoMCI Journal Web을 통하여 우리나라 의학학술지에 국내 유통에 대한 정확한 정보를 파악할 수 있게 되었다.<sup>4,5</sup> 이 KoMCI Journal Web은 국제적으로도 Thomson Reuters의 JCR, Elsevier사의 SCOPUS를 분석한 SCImago, 또한 Eigen Factor 등 국제적인 상업회사나 대학에서 다루고 있는 것과 비교하여 기능에 손색이 없어 유용하게 활용하고 있다.<sup>6,7</sup>

## 2. 계량서지학의 분야 Area of bibliometrics

크게 참고문헌 인용분석, 서지사항, 저자 인용 분석과 같은 인용분석이 있고, 내용 분석 크게 두 가지로 나눈다. 내용 분석은 MeSH와 같은 주제어를 붙여서 어느 분야에 해당하는 지 파악하고 내용과 유사성을 찾아가는 것으로 최근에는 알고리즘의 발달로 PubMed에서 관련 논문 보기 기능이 있다. 그 외 사회 네트워크 분석(social network analysis)과 같은 기법으로 관련도를 분석하는 경우도 있다.

### 3. 산출 방법 Calculating methods

여기서는 주로 인용 분석을 다루고 직접 자료를 가지고 측정하여 본다. KoMCI Web을 이용한 KoMCI Journal Web의 인용도지수 분석, Web of Science를 이용한 인용도지수 분석, Google Scholar를 이용한 개인의 h-index 산출, SCImago의 SJR 및 h-index<sup>9</sup> 결과에 대하여 알아보도록 한다.

#### 1) KoMCI Journal Web 및 KoMCI Web 이용한 학술지 인용도지수 산출

- KoMCI.org에 들어간다.
- KoMCI journal web을 선택한다.
- 2006년도 인용도지수 자료를 title로 검색한다.
- 자신의 학술지를 찾아 간다.
- 다음 자료를 찾아본다.
  - ▶ Bibliographic Data
  - ▶ IF: Impact Factor
  - ▶ Journal Cited-Half Life
  - ▶ Journal Citing Half-Life
  - ▶ Citation Information
  - ▶ ZIF: Z Impact Factor (Impact factor에서 자기 인용도 제외)
  - ▶ Cited Journal Graph
  - ▶ Citing Journal Graph

• 2006년도 Korean J Parasitol 인용도지수 계산법

|   |   |                                  |   |
|---|---|----------------------------------|---|
| Cites in 2006 to articles published in: | 2005 = 5<br><u>2004 = 6</u><br>Sum = 11 | Number of articles published in: | 2005 = 25<br><u>2004 = 29</u><br>Sum = 54 |
| Calculation: Cites to recent articles   |   | 11                               |   |
| Number of recent articles               |   | 54                               |   |

• 이 자료를 KoMCI Web 및 KoreaMed를 이용하여 계산하여 본다.

① 우선 KoreaMed에서 2004년 및 2005년에 발표한 Korean J Parasitol 논문을 센다. 이 값을 적는다.

**2004년도 발표한 논문수:**

**2005년도 발표한 논문수:**

**합:**

② KoMCI Web을 방문하여 Cited Reference Search로 들어가서 검색 창에서 Cited Journal에 자신의 잡지 약어를 찾아서 입력하고 Cited Year에 '2004 or 2005' 를 입력하고 Source Year에 '2006' 을 입력하고 오른쪽 아래 Search 단추를 누른다. 그러면 2006년도 KoreaMed 학술지에서 인용한 2004, 2005년도 발표 Korean J Parasitology 논문 수 및 논문 목록이 나온다. 그 목록 수를 적는다.

③ ①에서 구한 값을 분모로 2에서 구한 값을 분자로 하여 계산하면 인용도지수가 나온다. 이 구한 값과 KoMCI Journal Web에서 나온 값과 차이가 나는지 살펴보자. 차이가 난다면 왜 날까 생각하여 본다.

| 학술지명                           |  |
|--------------------------------|--|
| 2006년도 KoMCI Journal Web 인용도지수 |  |

## 2) SCIE JCR 및 Web of Science 이용한 학술지 인용지수 산출

이것은 SCIE 잡지인 경우에는 JCR에 값이 나오나 그렇지 않은 대부분의 학술지는 따로 계산을 하여야 한다.

- 우선 2004, 2005년도 발행 학술지 논문 수는 KoreaMed에서 나온 값을 그대로 사용한다(1).
- 도서관 주소로 들어간다. Web of Science 구독하고 있어야 한다. Web of Science를 누른다.
- 다음 바로 가기를 누르고 화면에서 다시 Select a Database를 누른 뒤 Web of Science를 누른다.
- 화면에서 다시 Cited Reference Search를 눌러서 Cited Work에 학술지명을 입력하는데 이 데이터베이스에 없으므로 자신의 학술지의 영문 약어명을 \* 와 함께 쓴다. 예를 들면 Korean J Pediatr 같으면 Kor\* J\* Ped\* 이렇게 친다. 그 이유는 SCIE 잡지가 아닌 경우는 학술지명의 전거통제를 이 database에서 하지 않기 때문에 저자들이 각각 다르게 표기하는 그대로 입력하기 때문이다. Cited year는 앞에서와 마찬가지로 2004 or 2005를 입력한다. 아니면 2004-2005와 같이 하이픈으로 연결하여도 된다. Source Year를 2006으로 설정하고 Search를 누르면 결과 화면이 나오고 거기서 다시 Select All을 선택하여 Search를 누르면 각각의 논문이 나온다. 왼쪽 차림표에서 Publication Year를 누르면 2006년도에 인용한 source 논문의 수가 나온다. 그 값을 적어 분자로 쓴다(2).
- 계산: 앞에서와 같이 (1)에서 구한 값을 분모로 (2)에서 구한 값을 분자로 하면 그 값이 2006년도 인용도지수이다. 비록 SCIE 학술지가 아니라도 이런 방법으로 자기 학술지의 인용도지수를 구할 수 있다.

### 3) h-Index 산출

자신의 논문을 인용이 많이 된 순으로 나열하였을 때 번호가 인용된 논문수 이상일 때 그 번호를 Hirsch-index라고 한다. Web of Science에서도 가능하나 Google Scholar에서도 구할 수 있다. Google Scholar에서 구하는 법은 다음과 같다.<sup>10</sup>

- Google Scholar에서 학술고급 검색을 누른다(<http://scholar.google.com>).
- 저자에 자신의 영문을 그리고 출판물에 논문이 실린 학술지명을 약어로 입력한다. 그리고 오른쪽 위 화면의 학술검색을 누른다.
- 그러면 자신의 논문이 죽 나오면서 인용회수가 나온다.
- 인용회수별로 죽 나열하고 순번을 매긴다.
- 인용회수와 순번이 일치하는 경우 그 순번이 h-index이며 일치하지 않으면 순번이 인용회수보다 적은 수 중 최대값이 h-index이다.
- 자신의 논문이 인용된 것을 찾을 수 있는 곳은 Web of Science, KoMCI Web 및 Google Scholar가 있으므로 정확히 찾아 보고 싶으면 세 곳을 다 찾아 보아야 한다.
- Table 11-1의 경우 h-index는 80이다.

**Table 11-1.** Number of cited per paper and the rank for a researcher

| No. cited | Rank |
|-----------|------|
| 34        | 1    |
| 21        | 2    |
| 18        | 3    |
| 16        | 4    |
| 12        | 5    |
| 10        | 6    |
| 9         | 7    |
| 8         | 8    |
| 7         | 9    |
| 6         | 10   |
| 6         | 11   |
| 5         | 12   |
| 5         | 13   |
| 5         | 14   |
| 4         | 15   |
| 4         | 16   |
| 4         | 17   |
| 3         | 18   |
| 3         | 19   |
| 3         | 20   |

#### 4) SCImago Journal Rank (SJR)

SCOPUS는 Elsevier라는 학술지 시장에서 가장 큰 다국적 상업회사에서 운영하는 데이터베이스로 자신 회사의 학술지 이외에도 SCI, Chemical Abstract, Biological Abstract, Excerpta Medical, Medline 학술지, 그 외 open access 학술지 등 13,000여 종을 다루는 학술문헌 데이터베이스이다. Web of Science의 7,000종보다 두 배 정도 많다. 지금 찾아 볼 수 있는 모든 데이터베이스 중에서 Google Scholar를 제외하고는 가장 큰 규모의 데이터베이스인데, 이곳에서는 기존의 Thomson이 장악하고 있는 학술지 평가 도구 및 개인 연구자 평가 도구인 인용도지수 및 총피인용수와 조금 다른 개념의 지표를 만들어 평가하고 있다. 즉 Thomson과 맞설 수 있는 더 큰 규모의 데이터베이스를 통하여 조금 더 많은 연구자가 공감하고 객관으로 평가 가능하다고 여기는 지표를 개발하였다. 그 중 하나가 SJR이고 다른 하나는 이미 앞에서 언급한 h-index를 학술지 단위로 측정하는 것이다. h-Index는 워낙 개인 연구자의 업적 평가를 위하여 개발되었지만 학문 분야별로 평가하는 h-b-index도 사용할 수 있고 이렇게 SCImago에서처럼 학술지 단위로도 가능하다<sup>11</sup>

SCIE 인용도지수의 단점은 어느 학술지에서 인용하더라도 같은 영향력을 미친다고 보는 것이다. SJR에서는 인용도지수가 낮은 학술지에서 인용하는 것보다 Nature, Science, Lancet 및 New England Journal of Medicine 같은 인용도지수가 매우 높은 학술지에서 인용하는 것을 더 높이 평가하여 다른 학술지 논문을 인용하면서 그 학술지의 명성을 보내주는 것이다.<sup>12</sup> 즉 우리가 보통 자신의 논문이 유명 학술지에서 인용되는 것을 높이 평가하는 것을 측정에 반영하여 현실에 더 맞는 지표로 삼으려는 것이다. 명성의 이전 (transfer of prestige)이라는 개념은 우리 일상 생활에서 늘 볼 수 있다. 친



구를 보면 그 사람을 판단할 수 있다고 여기는 것과 비슷한 개념이다. 측정은 Google에서 사용하는 PageRank 알고리즘을 사용하여 한다.<sup>13</sup> SJR이 인용도지수와 차이가 나는 점은 크게 4가지인데 이런 알고리즘의 차이가 있으며, 데이터베이스의 양이 SCIE보다 2배 크고, 인용 논문 분석하는 시점을 지난 3년 치로 확장하여 조금 더 안정된 자료를 얻을 수 있고 마지막으로 누구나 사용할 수 있는 무료 정보라는 점이다. 이 SJR은 높은 피인용회수를 보이면서도 SCIE에 등재되지 않은 open access 학술지 평가에 특히 유용하게 쓰인다. 또한 SJR 값과 인용도지수 값은 늘 상관관계가 높지는 않다. 측정 방법이 다르기 때문이다.

SJR을 계산하는 PageRank 알고리즘은 웹에서 어떤 페이지의 순위를 정할 때 사용하는데, 단순히 hyperlink를 많이 받는 것이 아니라 그 받는 페이지의 hyperlink를 고려하여 결정하고 hyperlink로 연결된 페이지 사이의 연결도 고려한다. 꽤 복잡한 알고리즘이지만 이런 알고리즘을 학술문헌의 인용 분석에 사용하여 현실에 조금 더 적절한 평가를 하려는 시도이다. 이 지표가 앞으로 얼마나 현실을 잘 반영하고 유용하게 쓰일지는 조금 더 두고 보아야 한다. 그러나 무료로 제공한다는 점, 데이터베이스 양이 전 세계의 모든 인용 색인데이터베이스에서 가장 많다는 점이 이 알고리즘의 가치를 더욱 높이 평가할 가능성이 있다. 우리나라 의학학술지 중 SCOPUS에 등재된 것은 이 값이 나온다.

## 5) Eigen factor<sup>7</sup>

이것은 2006년도 미국의 University of Washington의 생물학과 Bergstrom lab에서 지원하는 과제로 비영리 학술 과제로 진행하고 있다. Google의 PageRank와 같은 알고리즘을 사용하는 데 문헌은 SCIE 학술지의 참고문

현 지난 5년 치를 다루었다. 이 Eigen factor는 전체 과학 사회에서 학술지의 총 영향력을 표시하려는 측정도구이고, 모든 학술지의 Eigen factor의 합을 100으로 정하였다. Nature는 Eigen factor가 2.1이고 상위 1,000개의 학술지는 0.01을 넘었다. 특이하게 이곳에서는 학술지의 구독료의 값어치를 계산하여 학술지가 구독료만큼의 값어치가 있는지 판단하여 도서관에서 구독할 때 참고 자료로 삼을 수 있도록 하였다.

계량서지학 분야는 비록 의학자가 전공하지는 않지만 그 결과가 우리의 학계 연구 평가에 영향을 미치므로 어떤 지표가 어떤 장단점을 가지고 있는지 알 필요가 있다. 우리나라에서는 국내 문헌의 지표를 산출하는 작업을 이미 시도하여 적절한 데이터베이스가 있어서 활용할 수 있다. 우리나라 의학 학술지가 이런 국제적인 데이터베이스에서 평가지표가 나오려면 SCIE 등재 또는 SCOPUS 등재와 같은 색인데이터베이스에 등재가 필수이다. 앞으로 이런 분야에 대하여 편집인이 관심을 갖고 꾸준히 노력하면 계량서지학 결과를 학술지 발전의 한 평가 지표로 삼을 수 있을 것이다.

## REFERENCES

1. Bibliometrics [Internet]. Wikipedia; [cited 2010 Apr 19]. Available from: <http://en.wikipedia.org/wiki/Bibliometrics/>.
2. Lee CS. Productivity of SCI Korean medical papers: 1996-1997. J Korean Med Sci 1999;14:351-8.
3. Han MC, Lee CS. Scientific publication productivity of Korean medical colleges: an analysis of 1988-1999 MEDLINE papers. J Korean Med Sci 2000;15:3-12.
4. KoMCI [Internet]. Seoul: KoMCI; [cited 2010 Apr 19]. Available from:<http://komci.org/>.
5. ISI Web of knowledge [Internet]. London: Thomson Reuters; [cited 2010 Apr 19]. Available from: <http://apps.isiknowledge.com/>.
6. SCImago Journal & Country Rank [Internet]. Granada: SCImago Research Group; c2007-2010 [cited 2010 Apr 19]. Available from:<http://www.scimagojr.com/>.
7. eigenfactor.org [Internet]. Washington: Eigenfactor; [cited 2010 Apr 19]. Available from:<http://eigenfactor.org/>.
8. PubMed [Internet]. Bethesda (MD): US National Library of Medicine; [cited 2010 Apr 19]. Available from: <http://pubmed.org/>.
9. Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci USA 2005;102:16569-72.
10. Google. Google Scholar [Internet]. Google; [cited 2010 Apr 19]. Available from: <http://scholar.google.com/>.
11. H-b-index [Internet]. Wikipedia; [cited 2010 Apr 19]. Available from: [http://en.wikipedia.org/wiki/H-b\\_index](http://en.wikipedia.org/wiki/H-b_index).
12. BioMed Central [Internet]. London: BioMed Central; [cited 2010 Apr 19]. Available from: [http://blogs.openaccesscentral.com/blogs/bmcblog/entry/scimago\\_a\\_new\\_source\\_of/](http://blogs.openaccesscentral.com/blogs/bmcblog/entry/scimago_a_new_source_of/).
13. PageRank [Internet]. Wikipedia; [cited 2010 Apr 19]. Available from: <http://en.wikipedia.org/wiki/PageRank>.