Journal Metrics 개요

이 재 윤

(명지대학교)

Journal Metrics

- Impact Factor
  - E. Garfield (1955)가 제안
  - 1964년 Science Citation Index와 함께 발표되기 시작
  - 현재는 Thomson Reuters의 Web of Science (WoS)에 포함된 Journal Citation Reports에 매년 발표
- SCI는 평가 도구가 아니라 정보검색 도구로 개발됨
  - IF는 원래 저널 구독을 고려하는데 참고하기 위한 지표로 고안됨
  - SCI와 IF가 연구 평가 도구로 활용되면서 다양한 비판이 제기됨
  - Garfield (2004) “the tail of scientometrics is wagging the dog of information retrieval”
- Journal Metrics의 역사는 평가 지표로서 IF에 대한 문제점 지적과 대응의 역사
  - Web of Science는 새로운 지표를 꾸준히 추가
  - Scopus는 독자 지표 개발하고 꾸준히 개선
  - KCI에서도 지표의 토착화를 지속적으로 시도

IF

Journal Impact Factor

Journal Impact Factor (JCR-WoS)

저널 영향력 지수 IF

- 직전 2년 논문의 Y년 평균 인용빈도

$$IF(Y) = \frac{CIT_{Y-1}(Y) + CIT_{Y-2}(Y)}{PUB(Y-1) + PUB(Y-2)}$$

- IF 2014는 2012년, 2013년 논문이 2014년에 인용된 평균 횟수

$$IF_2(2014) = \frac{CIT(2014 \rightarrow 2012) + CIT(2014 \rightarrow 2013)}{PUB(2012) + PUB(2013)}$$

Journal Impact Factor

Publication-Citation Matrix

	발행년						합계
	2010	2011	2012	2013	2014	2015	
논문 수	62	62	67	60	61	57	369
2010인용횟수	1						1
2011인용횟수	20	6					26
2012인용횟수	41	31	20				92
2013인용횟수	24	44	36	14			118
2014인용횟수	38	49	37	54	21		199
2015인용횟수	24	47	37	55	57	9	229
인용합계	148	177	130	123	78	9	

Journal Impact Factor

Publication-Citation Matrix

Impact Factor, 2014년

	발행년						합계
	2010	2011	2012	2013	2014	2015	
논문 수	62	62	67	60	61	57	369
2010인용횟수	1						1
2011인용횟수	20	6					26
2012인용횟수	41	31	20				92
2013인용횟수	24	44	36	14			118
2014인용횟수	38	49	**37**	**54**	21		199
2015인용횟수	24	47	37	55	57	9	229
인용합계	148	177	130	123	78	9	

$$IF = \frac{37+54}{67+60} = 717$$

?

Journal Impact Factor

즉시성 지수 Immediacy Index

- IF 산출에서 누락되는 출판 당해년도의 인용을 별도로 고려
- Y년에 발행된 논문이 Y년에 인용된 평균 횟수

$$II(Y) = \frac{CIT(Y, Y)}{PUB(Y)}$$

- II 2014는 2014년 논문이 2014년에 인용된 평균 횟수

$$II(2014) = \frac{CIT(2014 \rightarrow 2014)}{PUB(2014)}$$

Journal Impact Factor

Publication-Citation Matrix

Immediacy Index, 2014년

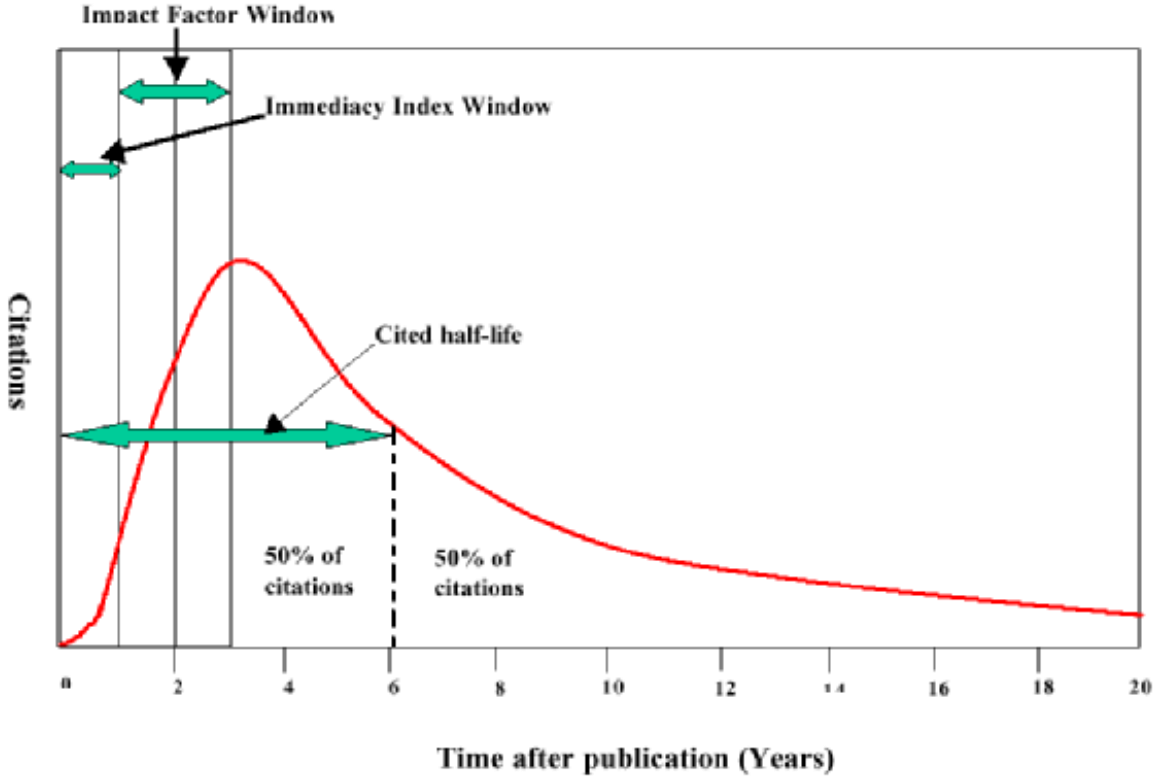
	발행년						합계
	2010	2011	2012	2013	2014	2015	
논문 수	62	62	67	60	61	57	369
2010인용횟수	1						1
2011인용횟수	20	6					26
2012인용횟수	41	31	20				92
2013인용횟수	24	44	36	14			118
2014인용횟수	38	49	36	14	21		199
2015인용횟수	24	47	37	55	57	9	229
인용합계	148	177	130	123	78	9	

$$II = \frac{21}{61} = 0.344$$

Journal Impact Factor

JCR의 초기 핵심 지표

- Impact Factor, Immediacy Index, Cited half-life



IF에 대응하는 Scopus 지표 - IPP

- Impact Per Publication

Journal Impact Factor

IF에 대한 비판 - 1

- 분야별 특성을 고려하지 않았으므로 학문 분야 간 직접 비교가 어렵다(Pinski & Narin, 1976)
- 타 분야 학술지로부터의 인용도 반영하므로 특정 분야 내에서의 중요성을 나타내지 못한다(Hirst, 1978)
- 각 학술지의 특성을 고려하지 않고 일률적으로 최근 2년의 데이터만 반영하였으므로 수치와 순위가 매년 크게 변화하는 경향이 있다(Glänzel & Moed, 2002)
- 리뷰 논문처럼 긴 논문이 실리는 학술지에 유리하다(Pinski & Narin, 1976; Hecht et al., 1998)
- 인용하는 학술지의 중요도를 반영하지 않으므로 핵심 학술지와 비핵심 학술지에 의한 인용이 모두 동일하게 반영된다(Buela-Casal, 2004)

Journal Impact Factor

IF에 대한 비판 - 2

- 사실상 '학술지' 단위가 아닌 평균 '논문' 단위의 지수이다(Hecht et al., 1998)
- 각 학술지가 지닌 다양한 특성을 하나의 단일 척도로 나타내는 것은 무리가 있다(Glänzel & Moed, 2002)
- 인용 측정 범위가 너무 짧으며 10년 정도 장기간의 데이터에 근거하여 영향력 지수를 산출하는 것이 바람직하다(Rousseau, 2002)
- 산출 공식의 분자/분모에서 대상 논문의 유형 규정이 다르다(분자에서는 모든 수록 논문에 대한 인용 포함; 분모에서는 정규논문, 노트, 리뷰만 포함)
- 인용은 논문/저널의 가치 중 일부분만 반영한다
- 영어권 저널에 유리하다.

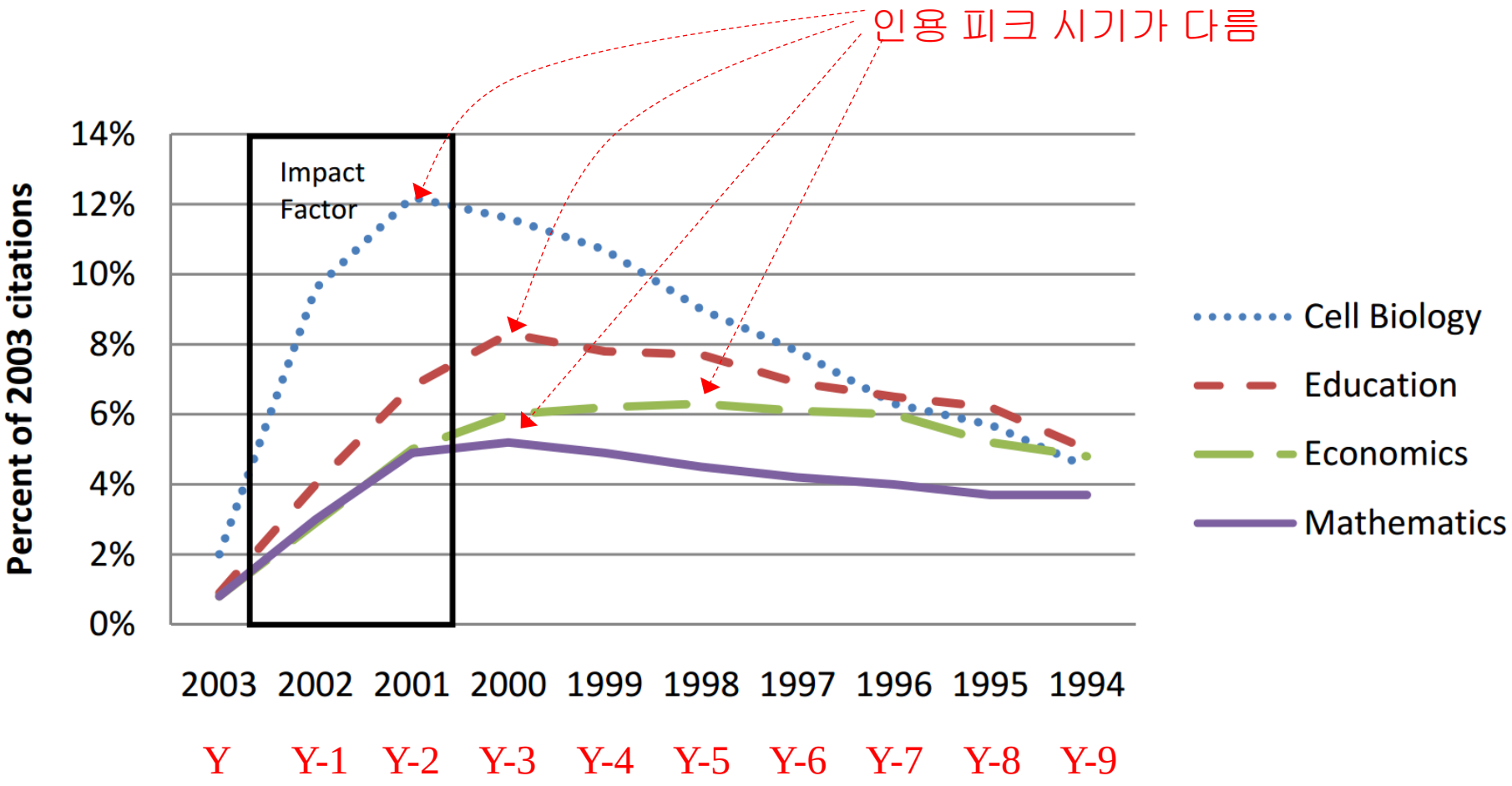
인용 기간과 문헌 유형

5-Year Impact Factor (5-Y IF)

Impact Per Publication (IPP)

Journal Impact Factor – 인용 기간

citation curves - 분야별

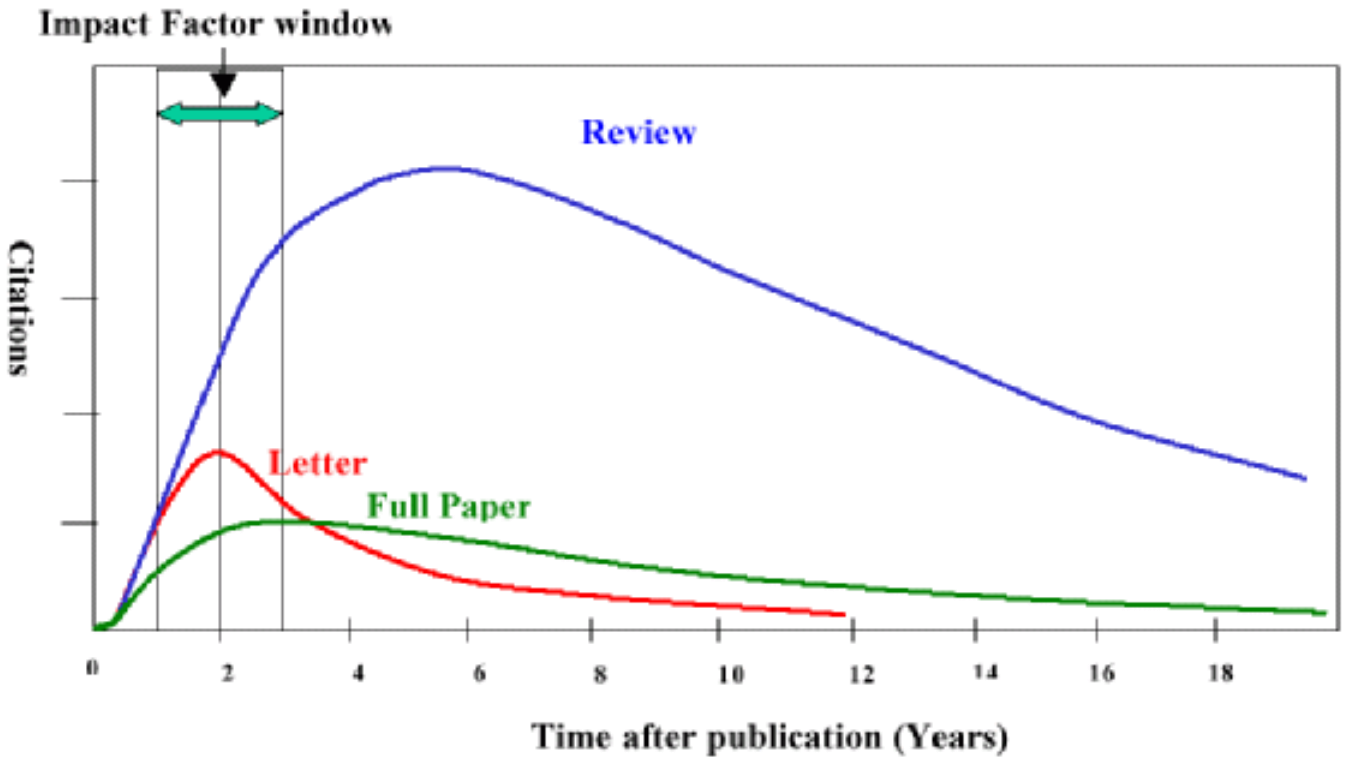


출처: Adler, Ewing, & Taylor (2008)

Journal Impact Factor – 인용 기간

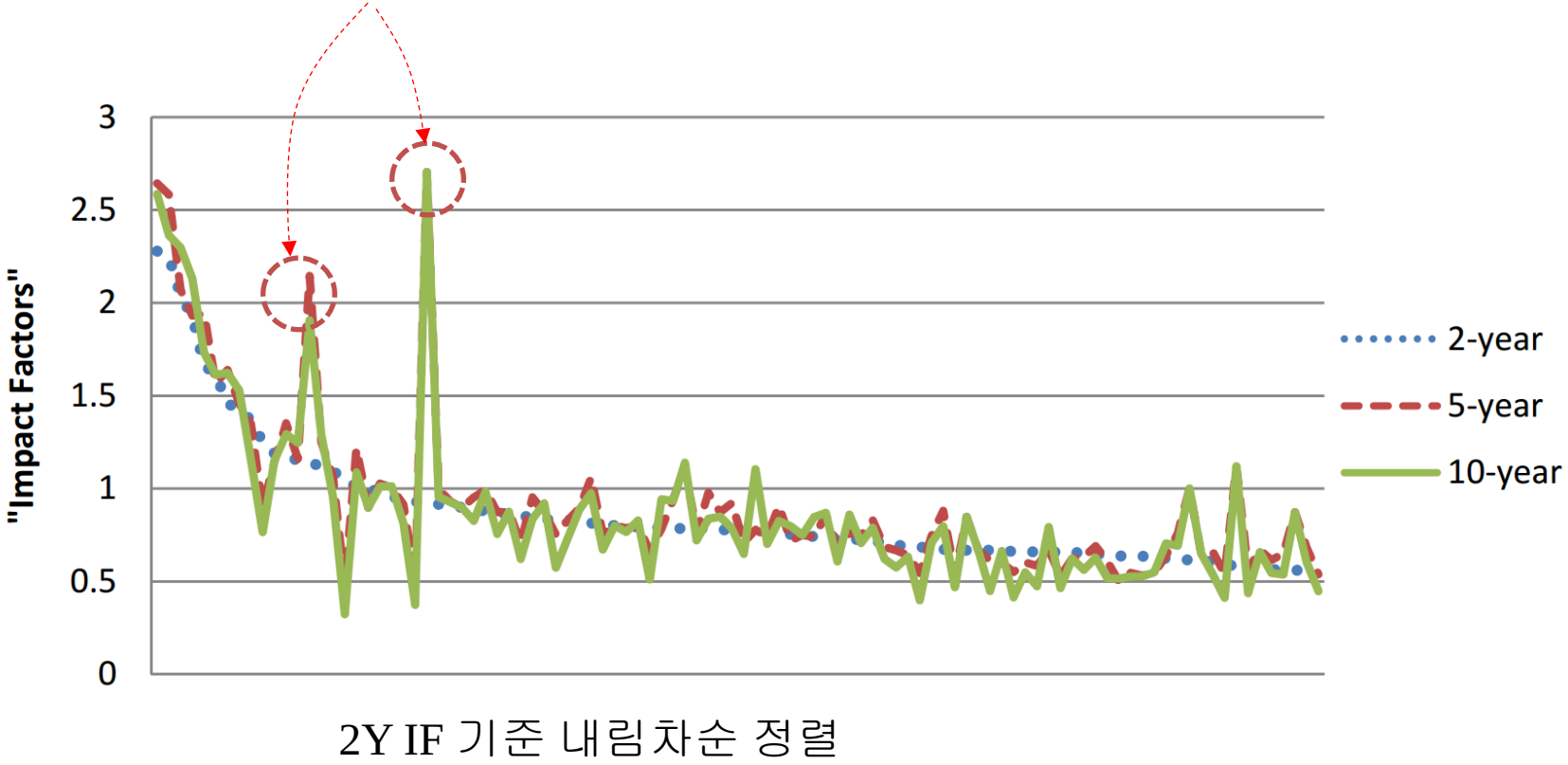
citation curves - 유형별

- 긴 논문이나 리뷰 논문을 싣는 학술지에 유리
  - 리뷰 논문은 일반 논문에 비해서 인용빈도가 세 배 정도로 추정됨



Journal Impact Factor – 인용 기간

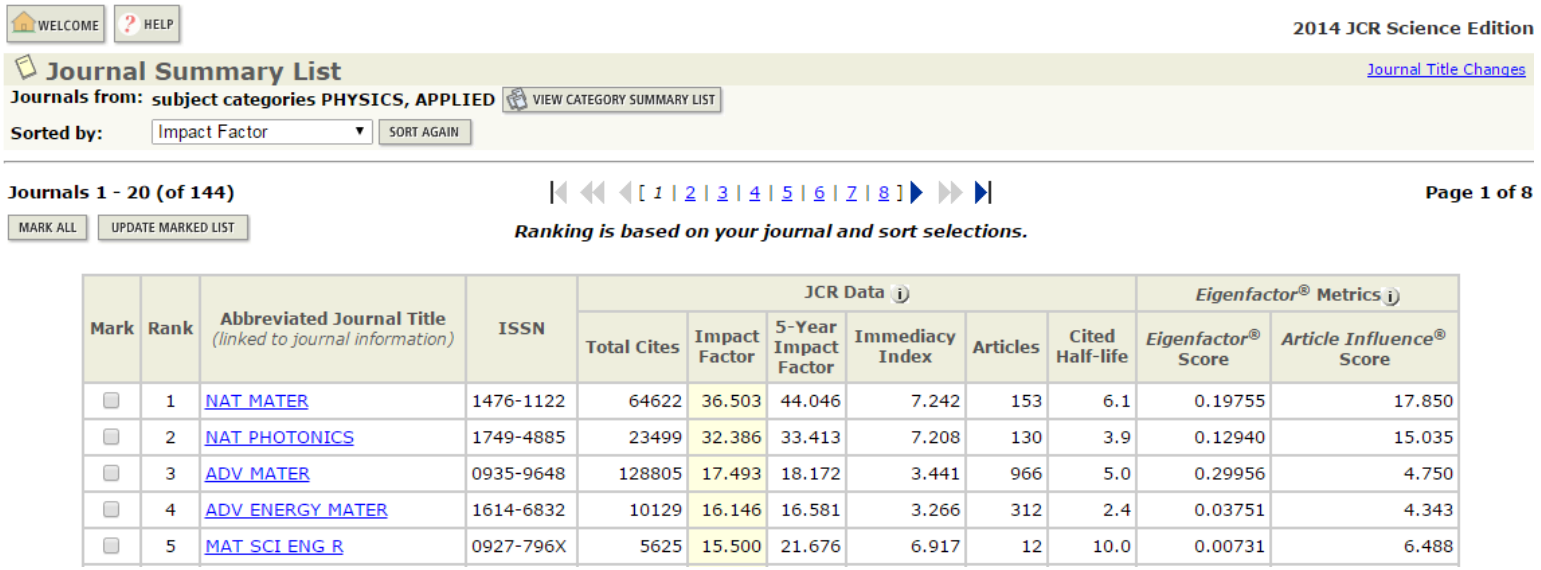
- **WoS 수학 분야 상위 저널 100종의 측정 기간별 IF 비교**
  - 2-Y IF, 5-Y IF, 10-Y IF는 대략 비례함 (r > 0.9)
  - 일부 저널의 순위가 예외적으로 크게 달라짐



출처: Adler, Ewing, & Taylor (2008)

5-Year Impact Factor (WoS)

- 인용 측정 기간 2년에 대한 대안
  - 직전 5년 내에 발간된 논문의 인용을 반영하는 지표 산출



WELCOME HELP

2014 JCR Science Edition

Journal Summary List

Journal Title Changes

Journals from: subject categories PHYSICS, APPLIED VIEW CATEGORY SUMMARY LIST

Sorted by: Impact Factor SORT AGAIN

Journals 1 - 20 (of 144) [1 | 2 | 3 | 4 | 5 | 6 | 7 | 8] Page 1 of 8

MARK ALL UPDATE MARKED LIST

Ranking is based on your journal and sort selections.

Mark	Rank	Abbreviated Journal Title (linked to journal information)	ISSN	JCR Data						*Eigenfactor*® Metrics	
				Total Cites	Impact Factor	5-Year Impact Factor	Immediacy Index	Articles	Cited Half-life	*Eigenfactor*® Score	*Article Influence*® Score
☐	1	NAT MATER	1476-1122	64622	36.503	44.046	7.242	153	6.1	0.19755	17.850
☐	2	NAT PHOTONICS	1749-4885	23499	32.386	33.413	7.208	130	3.9	0.12940	15.035
☐	3	ADV MATER	0935-9648	128805	17.493	18.172	3.441	966	5.0	0.29956	4.750
☐	4	ADV ENERGY MATER	1614-6832	10129	16.146	16.581	3.266	312	2.4	0.03751	4.343
☐	5	MAT SCI ENG R	0927-796X	5625	15.500	21.676	6.917	12	10.0	0.00731	6.488

  - Scopus의 IPP는 3년 단위로 측정

Journal Impact Factor – 논문 유형

대상 논문 유형 범위의 문제

- IF 산출 공식의 분자/분모에서 대상 논문의 유형 규정 불일치

$$IF(Y) = \frac{CIT_{Y-1}(Y) + CIT_{Y-2}(Y)}{PUB(Y-1) + PUB(Y-2)}$$

- 분자에서는 유형 불문 모든 수록 논문에 대한 인용 포함
- 분모에서는 정규논문(article), 노트, 리뷰만 포함하여 논문 수 산출

Scopus의 IPP는 유형을 일치시킴

- 대상: article, conference paper, review

Impact Per Publication (Scopus)

IPP

- WoS의 IF에 대응하는 Scopus의 논문 당 평균 인용빈도 지표
- 네덜란드 Leiden 대학의 CWTS(Centre for Science and Technology Studies)에서 개발
- 동일한 단순 평균 지표이지만, IF에 대해서 지적된 비판을 일부 고려하여 개발됨

IF와 IPP의 차이점

	IF	IPP
source DB	Web of Science	Scopus
citation window	직전 2년	직전 3년
cited document	모든 유형	article, conference paper, review로 제한
citing document	모든 유형	article, conference paper, review로 제한
number of publications	article, note, review로 제한	article, conference paper, review로 제한

Impact Per Publication (Scopus)

IPP의 안정성 문제

- 논문 수가 적은 소규모 저널은 IPP의 연간 변화가 상대적으로 크게 나타나고 지표값의 신뢰 구간이 넓을 수밖에 없음.
- CWTS에서는 연간 50건 미만의 논문이 게재되는 소규모 저널의 IPP는 신뢰도가 낮은 것으로 간주하여 기본 목록에서는 제시하지 않음

Median IF

유동적인 인용 기간 설정

$$MIF(Y) = \frac{CIT(Y)/2}{CPUB(Y - X_M, Y)}$$

$$= \frac{Y\text{년의 인용빈도의 절반}}{Y\text{년부터 인용반감기간까지의 누적 논문수}}$$

- 학술지별 인용빈도 감소 특성을 고려한 영향력 측정
- Sombatsompop et al.(2004) 제안, Rousseau(2005)가 정형화
- ISI IF는 인용반감기가 짧은 분야 학술지(의학, 공학)일수록 높게 나타나지만, MIF는 이와 무관함.
- MIF를 일반화한 Percentile IF도 제안됨

인용 출처의 차별화

Scimago Journal Rank (SJR - Scopus)

Article Influence Score (AIS - WoS)

Weighted Impact Factor (WIF)

중심성 지수 **(KJR - KCI)**

인용 출처의 차별화 문제

- **IF나 IPP는 인용의 경중을 반영하지 않음**
  - 핵심 학술지와 비핵심 학술지에 의한 인용이 모두 동등하게 빈도로만 반영됨
  - Science, Nature 등에서 인용되는 것과 무명의 학술지에서 인용되는 것은 차별적으로 반영해야 한다는 주장
  - 대안 : Weighted IF (WIF), SJR, AIS, KJR

인용 출처의 차별화 문제

Weighted Impact Factor (WIF)

$$WIF = \frac{MIFJC + IF}{2}$$

$MIFJC =$ 지난 2년간 학술지 A에 게재된 논문을 올 해에 인용한 학술지의 IF를 학술지 A를 인용한횟수에 따라 가중평균한 값

- Buela-Casal(2004)이 주요 학술지로부터의 인용을 더 중요하게 반영하는 저널 지표로 제안함
- 학술지 A를 인용한 학술지들의 IF를 가중 평균하여, 학술지 A의 IF와 더해서 2로 나눔
- MIFJC는 평균으로의 회귀 경향이 있기 때문에 자신의 IF와 더해서 보완함
- 수학적인 엄정성이 결여되어 있기 때문에 널리 사용되지 못함

인용 출처의 차별화 문제

Journal PageRank

$$JPR_i = \frac{(1-d)}{N} + d \cdot \sum_{j=1}^{N} \frac{JPR_j \cdot C_{ji}}{C_j}$$

여기서,

JPR_i : 학술지 i의 페이지랭크

C_{ji} : 학술지 j로부터 학술지 i에 대한 인용 빈도

C_j : 학술지 j의 참고문헌 수(나가는 인용 빈도)

d : 상수, 보통은 0.85

N : 학술지 종수

- Google의 PageRank 알고리즘을 응용하여 인용한 저널의 중요도를 인용 비중과 곱하여 합산함 (Bollen, Rodriguez, & Van de Sompel, 2006)
- IF는 popularity를 측정하는데 반해서 JPR은 prestige를 측정

인용 출처의 차별화 문제

Journal PageRank

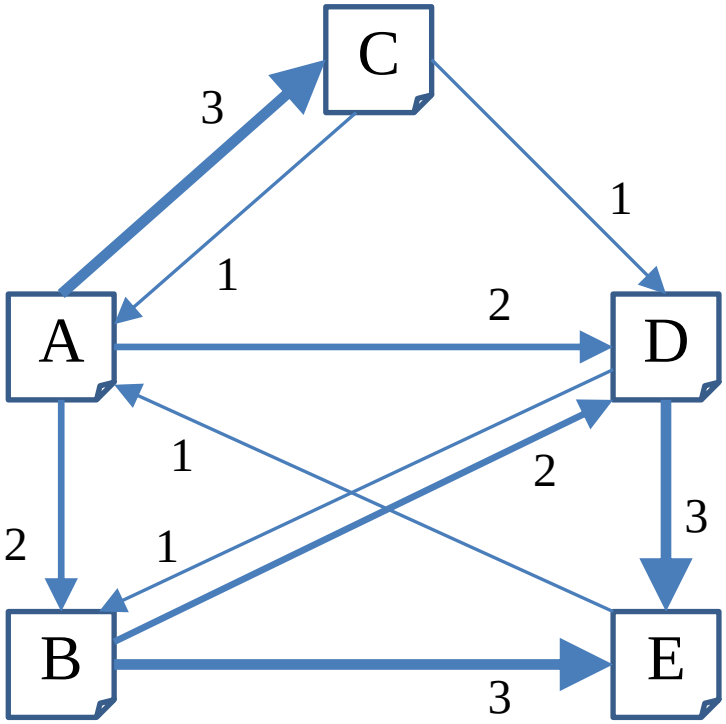
$$JPR_i = \frac{(1-d)}{N} + d \cdot \sum_{j=1}^{N} \frac{JPR_j \cdot C_{ji}}{C_j}$$

- 공식의 앞 부분은 각 저널이 가지는 최소 중요도
  - 이 값 덕분에 인용빈도가 0인 저널이 받은 중요도가 없어도 다른 저널에게 인용을 통해 중요도를 넘겨줄 수 있음
- 공식의 뒷 부분은 각 저널이 가진 중요도를 상대방 저널을 인용한 비중만큼 보내주는 역할
- 중요도 전체를 보내지 않고 d(damping factor)만큼 할인하여 보내줌. 나머지 1-d는 각 저널에게 고르게 배분됨.
- 처음에는 각 저널의 중요도를 1/N로 초기화하여 계산 실시
- 여러 차례 반복 계산하여 값의 변화가 일정 수준 이하가 되도록 수렴시키는 방식으로 최종 JPR 산출

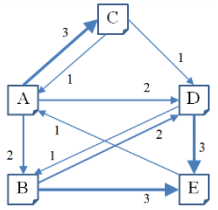
인용 출처의 차별화 문제

Journal PageRank 계산

- 저널 A의 초기 JPR
  - 1/5
- 저널 A가 1차 계산에서 받는 값
  - C로부터 1/2 × 1/5
  - E로부터 1/1 × 1/5
  - 합계 = 3/10
- 각 저널의 JPR 합계는 항상 1.0



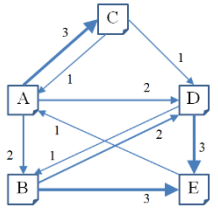
인용 출처의 차별화 문제



Journal PageRank 계산시 고려할 점

1. Dangling node
   - 내보내는 인용이 없는 저널의 중요도는 어디로?
   - 이를 고려하지 않으면 값이 새어나가 반복 계산할수록 합계가 줄어듬
2. 저널 자기 인용
   - 자기 인용은 중요도를 나누지 않고 받기만 하여 자신에게 누적시킴
   - 자기 인용이 많을수록 JPR의 비정상적인 인플레이션이 발생함
3. 불활성 참고문헌
   - 활성 참고문헌은 평가 대상 기간에 속하는 인용DB 논문에 대한 인용을 의미함
   - DB 외부(미등재) 저널에 대한 인용과 대상 기간 외 논문에 대한 인용이 많은 경우에 이를 무시하는 경우와 포함하는 경우의 결과가 달라짐
4. 논문 수
   - JPR은 저널 단위로 측정되므로 논문이 많은 저널에게 유리함
   - JPR을 논문 수로 나누어서 비교해야 함
5. damping factor를 통한 전체 균등 배분 시 저널 규모 고려
   - 각 저널이 가진 중요도의 1-d 만큼이 저널의 규모와 상관없이 균등하게 배분되기 때문에 논문 당 받는 중요도는 오히려 큰 저널이 역차별받음

Scimago Journal Rank (Scopus)



SJR

- Journal PageRank의 실제 적용 시 발생하는 문제점을 세심하게 고려한 지표 (SCImago Research Group, 2007).
  1. Dangling 저널의 중요도는 전체 저널에게 논문 수에 따라 차등 배분
  2. 저널 자기 인용을 전체 인용의 1/3 이내가 되도록 제한 (타 저널 인용의 절반까지만 저널 자기 인용을 허용, 나머지는 버림)
  3. 활성 참고문헌과 불활성 참고문헌을 내보내는 인용의 총량에 모두 포함하되, 불활성 참고문헌의 비율만큼 중요도를 전체 저널에 규모에 비례하여 고르게 배분
  4. 저널 단위로 산출한 중요도인 PSJR을 논문 수로 나누어서 최종 SJR을 산출
  5. 1-d의 비율만큼 전체 학술지에 고르게 배분되던 가중치를 학술지의 규모에 비례하도록 차등 배분
- IPP와 마찬가지로 직전 3년 동안 발간된 지정 유형의 논문에 대한 인용만 고려함
- Scopus 데이터에 대해서 SJR과 IPP의 순위 상관계수는 0.93인 것으로 보고됨 (González-Pereira, Guerrero-Bote, & Moya-Anegón 2010).

Scimago Journal Rank

SJR 공식

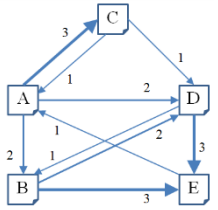
$$SJR_i = c \cdot \frac{PSJR_i}{Art_i}$$

$$PSJR_i = \frac{(1-d-e)}{N} + e \bullet \frac{Art_i}{\sum_{j=1}^{N} Art_j}$$

$$+ d \bullet \sum_{j=1}^{N} \frac{C_{ji} \bullet PSJR_j}{C_j} \bullet \frac{1 - \left(\sum_{k \epsilon (Dangling-nodes)} PSJR_k \right)}{\sum_{h=1}^{N} \sum_{k=1}^{N} \frac{C_{kh} \bullet PSJR_k}{C_k}}$$

$$+ d \bullet \left[\sum_{k \epsilon (Dangling-nodes)} PSJR_k \right] \bullet \frac{Art_i}{\sum_{j=1}^{N} Art_j}$$

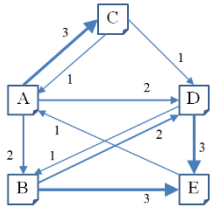
Article Influence Score (WoS)



AIS

- Journal PageRank방식으로 WoS에서 측정한 Eigenfactor를 논문 수로 나누어 정규화한 지표 (Bergstrom, 2007).
- Scopus의 SJR과의 차이점
  1. 저널 자기 인용을 모두 제거
  2. 활성 참고문헌만 내보내는 인용으로 고려하고 불활성 참고문헌은 배제
  3. 직전 5년 동안 발간된 논문에 대한 인용만 고려
  4. 전체 AIS의 평균이 1이 되도록 조정

Korean Journal Rank (KCI)



■중심성 지수 KJR

- Scopus의 SJR을 국내 상황에 맞도록 수정하여 KCI에 적용한 지표
- Scopus의 SJR과의 차이점
  1. 저널 자기 인용을 전체 저널의 자기인용율 평균+표준편차까지 허용
  2. 비활성 참고문헌 중에서 일부를 제외

		참고문헌의 KCI 등재 여부	
		KCI 논문	비KCI 논문
y년에 발행된 논문의 참고문헌의 발행 연도	y+1년 ~ y+k년	a	b
	y+0년, y+k+1년 이후	c	d

r1 = a + b + c + d ; 사용된 모든 참고문헌 집합

r2 = a + b ; 발행년이 직전 1년~직전 k년 이내인 참고문헌 집합

r3 = a ; 발행년이 직전 1년~직전 k년 이내이면서 KCI 수록 논문인 참고문헌 집합

- r2와 r3를 각각 내보내는 인용집합으로 간주하여 측정한 두 값의 기하 평균을 최종 지표값으로 사용함

분야별 차이의 반영 문제

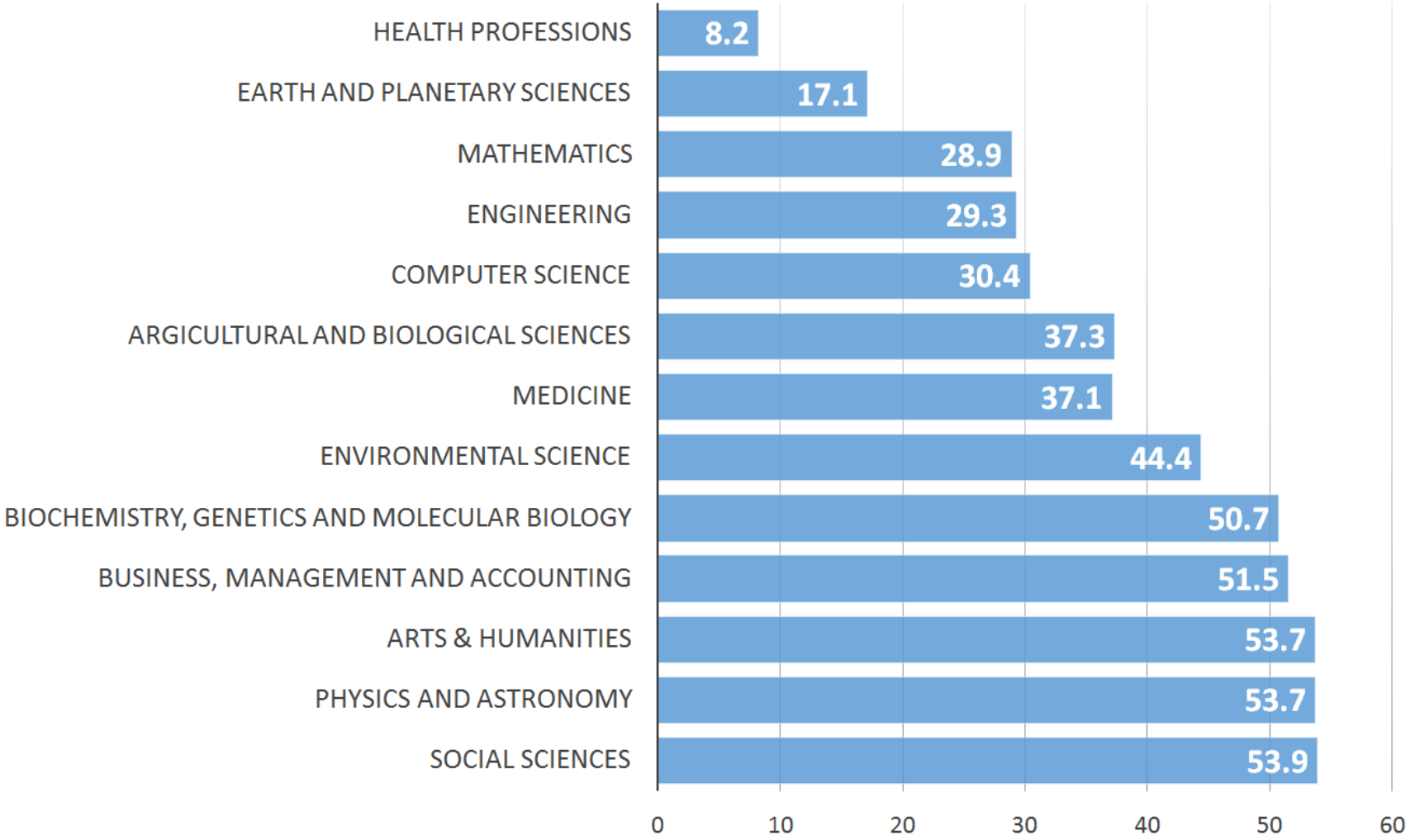
Relative Citation Impact Index (RCII)

Source Normalized Impact per Paper (SNIP - Scopus)

분야별 논문 당 참고문헌 수

JCR 주제 분야 일부 표본 분석 결과

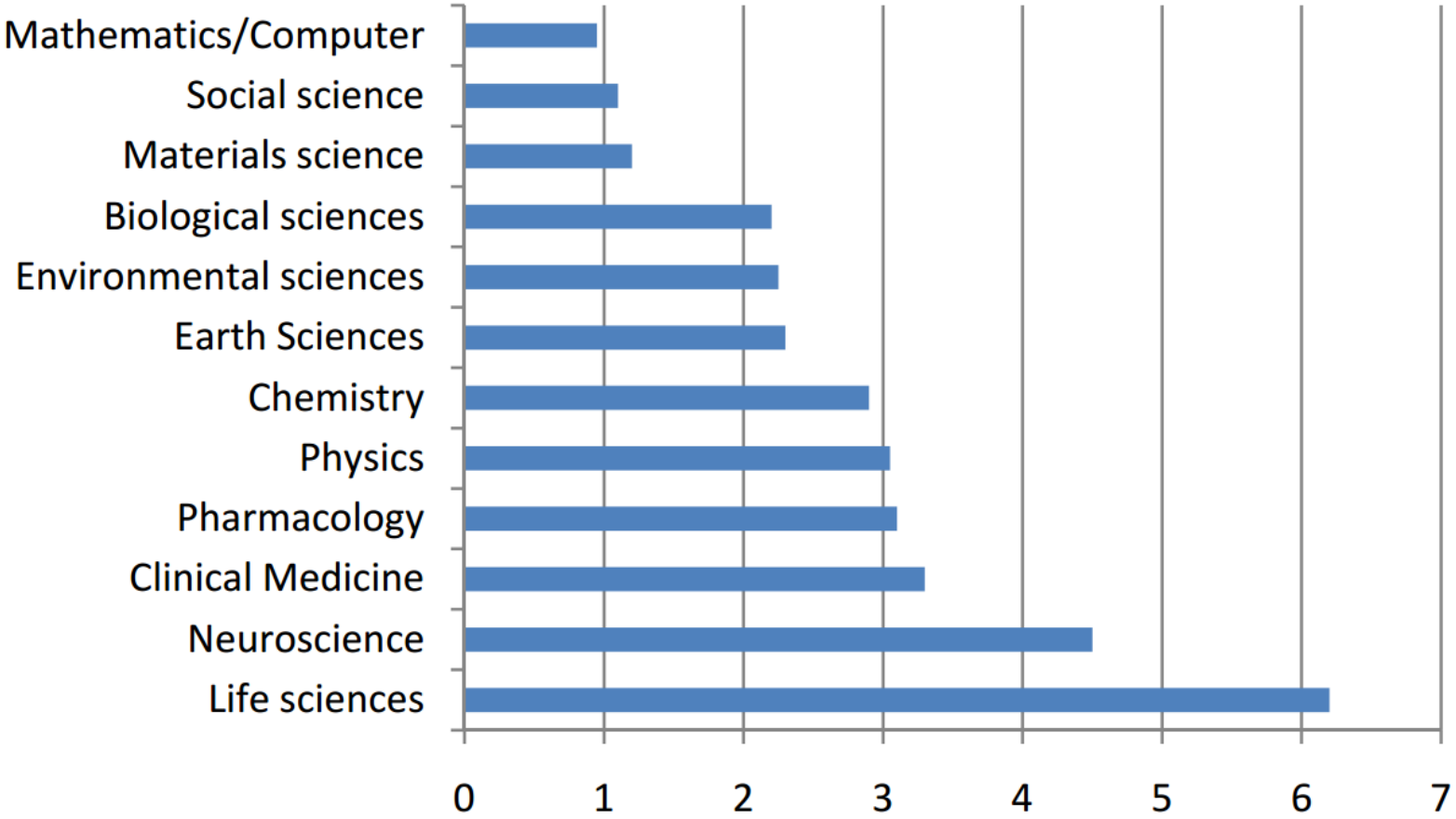
출처: Halevi (2013)



14개 분야 표본 저널 63종 분석 결과

분야별 논문 당 평균 인용빈도

- **JCR 분석 결과**



출처: Adler, Ewing, & Taylor (2008)

분야별 차이 보정 – Normalization

Normalization

- 분야별 인용 관행의 차이를 보정하는 방법
  - 물리학 분야와 같이 논문 당 참고문헌이 많은 경우는 평균 인용이 높음
  - 수학 분야와 같이 논문 당 참고문헌이 적은 경우는 평균 인용이 낮음

Normalization 방식 구분 – 인용측/피인용측

- Cited side normalization
- Citing side normalization

Normalization 방식 구분 – 고정 범주/맞춤 범주

- Fixed subject category
- Journal specific subject category

분야별 차이 고려

Relative Citation Impact Index (RCII)

$$RCII = \frac{RCR - 1}{RCR + 1}$$

$$RCR = \frac{MOCR}{MECR} = \frac{\text{실제 평균 인용율}}{\text{기대 평균 인용율}}$$

$$MECR = \text{특정 집단의 평균 인용율}$$

- 피인용측 정규화, 고정 범주 정규화 (CWTS)
- Relative Citation Rate (RCR) 값을 정규화(-1에서 1사이)한 지수
- RCR은 인용지수의 기대값 대비 실제값
- 인용지수의 기대값은 일정 집단의 평균값
  - 일정 집단 : JCR 저널 소속 주제 저널 집단
- 인용율은 IF나 IPP를 비롯한 기존 인용지수 사용 가능
- 0 이하면 특정 집단에서 평균 이하의 인용율, 이상이면 평균 이상의 인용율을 나타냄
- -1은 전혀 인용되지 않는 경우, 1은 이론상 인용빈도가 무한대인 경우

Source Normalized Impact per Paper (Scopus)

SNIP

$$SNIP = \frac{RIP}{RDCP}$$

RIP : Raw Impact per Paper
RDCP : Relative Database Citation Potential

- 인용측 정규화, 맞춤 범주 정규화 (Moed, 2010)
- 분자는 논문 당 인용빈도
- 분모는 인용한 논문의 참고문헌 규모의 상대적인 수준
- 참고문헌이 많은 논문으로부터 인용된 경우보다 참고문헌이 적은 논문으로부터 인용된 경우를 더 높게 반영함

결론

- 저널 인용 지표는 단순 평균 방식 IF에서 출발
- 측정 기간의 문제점은 3년(IPP), 5년(5-Y IF) 등으로 개선됨
- 인용 출처의 차등 고려 문제는 SJR, AIS 등의 지표에서 반영됨
- 분야별 차이 문제는 SNIP에서 반영됨
- 그러나, 본질적인 한계는 여전히 남아있음
  - 저널 평가는 다면적인 지표를 활용해야 함
  - 인용으로 반영되지 않는 측면을 고려해야 함
  - 저널 평가를 연구 성과 평가의 수단으로 직접 활용하는 것은 무리