

공저자 수를 고려한 공저 네트워크 중심성과 연구성과의 연관성 분석

Comparative Analysis on the Relationships between the Centralities in Co-authorship Networks and Research Performance Considering the Number of Co-authors

이재윤 (Jae Yun Lee)*

초 록

국내 문헌정보학 분야에서 10년간 발표된 논문의 저자와 인용빈도를 대상으로 공저 네트워크에서의 중심성과 연구성과 지수 사이의 관계를 분석하였다. 특히 공저를 고려하지 않고 연구성과 지수를 산출하는 경우와 공저를 고려하여 연구성과 지수를 산출하는 경우로 나누어서 분석하였다. 또한 저자 집단을 논문 수에 따라 다르게 설정하여 지수 사이의 상관관계를 분석한 결과, 연구자의 인용지수와 연구자 중심성 사이의 상관관계에 대한 선행 연구의 일관성없는 결과를 설명해낼 수 있었다. 전체적으로 공저 활동의 정도는 연구성과와 상관관계가 유의하지 않았으며 일부에서는 오히려 부정적인 상관관계를 가진 것으로 나타났다. 중심성과 연구성과 사이의 관계는 통계적으로 유의한 긍정적인 상관관계가 나타났으나 상위 저자 30명만을 대상으로 분석한 결과에서는 상관관계가 유의하지 않았다.

ABSTRACT

We analyzed the relationships between the co-authorship network centralities and the research performance indicators with the authors and the number of citations of the papers published for 10 years in Korean library and information science journals. In particular, the research performance indicators were calculated with normal counting and with fractional counting also. As a result of correlation analysis between the variables by setting the different ranges of the author groups to be analyzed according to the number of articles, it was possible to explain the inconsistent results of the previous studies on the correlations between the researchers' citation indicators and their co-authorship network centralities. Overall, the degree of co-authorship activities measured by collaboration coefficient showed no or negatively correlated with research performance. There were statistically significant positive correlations between the centralities and the research performance indicators, but the correlation was not significant in the analysis of the top 30 authors by number of articles.

키워드: 공저 네트워크, 연구성과, 네트워크 중심성, 공저자, 공동연구
co-authorship network, research performance, network centrality, co-authors,
research collaborations

* 명지대학교 문헌정보학과 부교수(memexlee@mju.ac.kr)

■ 논문접수일자: 2016년 11월 21일 ■ 최초심사일자: 2016년 12월 14일 ■ 게재확정일자: 2016년 12월 16일
■ 정보관리학회지, 33(4), 175-199, 2016. [http://dx.doi.org/10.3743/KOSIM.2016.33.4.175]

1. 서론

공동연구가 활성화되면서 최근 발표되는 학술 논문에도 공동저술 논문의 비율이 지속적으로 증가하고 있다. 1900년부터 2011년까지 100년이 넘는 기간 동안의 공저 증가 추세를 분석한 Larivière, Gingras, Sugimoto, Tsou(2015)는 단독저술 논문의 비중이 1900년에는 자연과학 및 의학 분야가 87%였고 인문사회과학 분야가 97%였는데, 2011년에는 각각 7%와 38%로 크게 감소했다고 보고하였다. Katz와 Martin(1997)은 이런 공동연구의 증가 원인으로 연구에 소요되는 비용의 급증, 연구 협력에 소요되는 비용(커뮤니케이션 비용)의 급감, 협력을 통한 과학 발전의 가속 효과, 과학의 세분화에 따라 각자의 전문성을 살리는 협력의 필요성 증가, 학제적 분야의 중요성 증가를 제기한 바 있다. 이와 같이 공동연구의 증가는 비용 문제와 같은 환경적인 요인 이외에 협력을 통해 연구 성과를 더 향상시킬 수 있다는 가정이 전제되어 있다.

그러나 기대만큼 연구 협력이 긍정적인 것만은 아니라는 지적도 제기되고 있다. Lee와 Bozeman(2005)은 연구 협력이 연구 생산성을 높일 것이라는 기대가 널리 퍼져 있지만 둘 사이의 관계는 뚜렷하지 않다고 주장하였다. 이들은 협력이 연구 생산성에 오히려 저해가 될 수 있는 여러 가지 상황을 지적하였다. 공동연구에 소모되는 논의 및 대기 시간문제, 일부 참여자의 역할 미이행으로 인한 프로젝트의 부진, 학생이나 신참 연구원을 지도하는 중견 연구자의 시간 소모 문제 등이 그것이다.

이처럼 공동연구의 효과에 대한 논란이 있음

에 따라서 공동연구를 통한 공동저술 논문 발표와 연구 성과 사이의 관계에 대해서 탐구한 다수의 논문이 발표되었다. 이들 연구에서는 공저와 연구 성과라는 두 가지 요인에 대해서 각각 다양한 변수를 사용하여 측정하고 상관관계를 비교하였다. 선행 연구에서 공저에 대해서 측정된 주요 변수로는 협력한 공저자의 수(Lee & Bozeman, 2005; Li, Liao, & Yen, 2013), 단독저술 비율(박치성, 2012), 공저 네트워크에서의 중심성(김은주, 남태우, 2015; 박치성, 2012; 이수상, 2013; 임병학, 2013; Abbasi, Altmann, & Hossain, 2011; Lee & Bozeman, 2005; Li, Liao, & Yen, 2013) 등이 활용되었다. 연구 성과에 대해서 측정된 주요 변수로는 발표 논문 수(박치성, 2012; 이수상, 2013; 임병학, 2013; Abbasi, Altmann, & Hossain, 2011; Lee & Bozeman, 2005), 총 인용빈도(김은주, 남태우, 2015; 이수상, 2013; Abbasi, Altmann, & Hossain, 2011; Li, Liao, & Yen, 2013), 평균인용빈도(김은주, 남태우, 2015), h-지수(이수상, 2013; Abbasi, Altmann, & Hossain, 2011) 등이다. 이외에도 Bordons, Aparicio, González-Albo, Díaz-Faes(2015)의 연구도 있으나, 이들은 단 3년 분량의 논문 집합을 대상으로 공저 네트워크를 구축하였으므로 공저 관계가 충분히 포착되지 못한 결과라고 판단하여 분석에서 제외하였다.

연구성과 변수는 크게 양적 성과 변수와 질적 성과 변수로 구분할 수 있다. 일반적으로 논문 수를 양적 성과 지표로, 평균인용빈도를 질적 성과 지표로 다루며 h-지수(Hirsch, 2005)나 총 인용빈도는 양적 성과와 질적 성과를 함께 반영하는 혼합 지표로 구분한다. 선행 연구

의 결과를 비교해보면 공저 관련 변수와 연구 성과 변수 사이의 상관관계에 대해서 부분적으로 불일치하는 경우가 있었다. 특히 공저 네트워크에서 연구자의 입지를 나타내는 중심성과, 질적 성과 지표인 인용빈도 사이의 상관관계에 대해서 상반된 연구결과가 제시된 바 있다. 예를 들어 Abbasi, Altmann, Hossain(2011)은 공저 네트워크에서의 연결정도중심성과 매개 중심성이 연구자의 인용빈도와 긍정적인 상관관계가 통계적으로 유의하게 나타났다고 보고한 반면, 이수상(2013)은 동일한 변수 사이의 상관관계가 유의하지 않았다고 보고하였다.

이 논문에서는 이와 같이 선행 연구에서 다루어진 연구성과 변수와 공저 관련 변수 사이의 상관관계를 정리해본 후, 두 가지 주요 변수를 통제하면서 상관관계를 분석해보고자 한다. 통제하려는 첫 번째 변수는 논문의 저자 수이다. 이수상(2013)은 논문 수를 단순 건수로 측정하지 않고 각 논문의 저자 수를 고려하여 논문 수를 분수 계상(fractional counting)하는 방식으로 공저 보정 논문 수를 산출하였다. 이 연구에서는 논문 수뿐만 아니라 인용빈도와 h-지수, 평균인용빈도 등의 질적 연구성과 변수도 분수 계상함으로써 모든 연구성과 산출시 공저자 수를 통제해보았다. 그리고 공저자 수를 보정한 경우와 보정하지 않은 경우의 연구성과 변수를 별도로 분석하여 공저자 수의 통제 여부에 따라 공저 관련 변수와 연구성과 변수 사이의 상관관계가 달라지는지를 확인해보았다. 통제 대상인 두 번째 변수는 저자의 논문 수이다. 발표한 논문 수가 한 두 편인 연구자에 비하면 수십 편의 논문을 발표한 연구자는 질적 연구성과가 높을 가능성이 크고 공저 관

계도 더 활발할 여지가 있다. 따라서 논문 수가 매우 적은 연구자도 포함한 경우와 논문 수가 많은 연구자만 포함한 경우를 구분하여 변수 간 상관관계를 검증해보고자 한다. 이를 통해서 선행 연구에서 일관되지 않은 결과가 얻어진 공저와 연구성과 사이의 상관관계를 면밀하게 파악하는 것이 이 연구의 목적이다.

2. 선행연구 분석

공저 관련 변수와 연구성과 변수 사이의 상관관계를 다룬 주요 연구를 살펴보면 다음과 같다. Lee와 Bozeman(2005)은 미국 연구 기관에 속한 이공계 분야 연구자 443명에 대해서 조사한 결과 협력한 공저자의 수가 많은 연구자일수록 발표한 논문 수는 많아지지만, 공저자 수로 나누어서 산출한 논문 수는 협력한 공저자 수와 상관이 없다는 결과를 보고하였다.

Abbasi, Altmann, Hossain(2011)은 5개 iSchool에서 생산된 논문의 저자 1,809명을 대상으로 네트워크 분석 지표인 연결정도, 근접중심성, 매개중심성, 아이겐벡터 중심성, 평균 연결 강도와 연구성과 지표인 논문 수, 인용빈도, h-지수, g-지수 사이의 스피어맨 순위 상관계수를 산출해본 결과 연결정도중심성, 매개중심성, 평균연결강도가 모든 성과 지표와 뚜렷하게 유의한 상관관계를 보이는 것으로 나타났다.

박치성(2012)은 행정학 분야에 대한 분석에서 단독저술 비율, 연결정도중심성, 매개중심성이 발표논문 수에 긍정적인 영향력을 끼친다고 보고하였다.

임병학(2013)은 로고스경영연구의 2003년부

터 2011년까지의 122개의 논문과 89명의 저자를 대상으로 네트워크 분석 지표인 연결정도, 근접중심성, 매개중심성, 구조적 공백과 논문 생산성이 모두 통계적으로 유의한 상관관계를 보인다고 보고하였다.

이수상(2013)은 국내 기록관리학 분야 2개 학술지에 논문을 발표한 258명 중에서 연결정도, 근접중심성, 매개중심성 상위 30위 이내에 포함된 저자들을 대상으로 투고논문수, 인용빈도, h-지수와 의 상관관계를 산출해보았다. 그 결과 세 가지 중심성 모두 투고논문수 및 h-지수와는 유의한 상관관계를 보였지만, 인용빈도는 연결정도 및 매개중심성과는 상관관계가 없고 근접중심성과의 상관관계만 유의하게 나타났다.

Li, Liao, Yen(2013)은 SSCI 데이터베이스에서 정보시스템 분야 137명 연구자 사이의 공

저 네트워크 중심성과 공저자 수 등이 인용빈도와 가지는 상관관계를 검증해본 결과 매개중심성을 비롯한 중심성 지수는 긍정적 상관관계를 보였지만 공저자 수는 그렇지 않았다고 보고하였다.

김은주와 남태우(2015)는 정신의학 분야의 국가 간 국제공동연구 네트워크를 분석하였는데 중심성과 논문 수, 총 피인용빈도 사이에는 통계적으로 유의한 상관관계가 나타났지만, 중심성과 평균 피인용도 사이의 상관관계는 유의하지 않았다. 이들의 분석 단위는 개인이 아닌 국가 단위였다.

이상의 선행 연구에서 다루어진 공저 관련 변수와 연구성과 변수 및 그 사이의 상관관계 검증 결과를 정리해보면 <표 1>과 같다. 연구성과 변수 중에서 확실한 양적 성과에 해당하는 것은 논문 수와 공저 보정 논문수이며 가장 질적

<표 1> 선행 연구에서 검증된 연구성과 변수와 공저 관련 변수 사이의 상관관계

비교 대상		상관관계 검증 결과	
공저 관련 변수	연구성과 변수	긍정적 상관관계 유의	상관관계 유의하지 않음 또는 부정적 상관관계 유의
단독저술 비율	논문 수	박치성(2012)	
	논문 수	Lee와 Bozeman(2005)	
협력한 공저자 수	공저 보정 논문 수		Lee와 Bozeman(2005)
	인용빈도		Li, Liao, Yen(2013)
중심성*	논문 수	김은주와 남태우(2015); 박치성(2012); 임병학(2013); Abbasi, Altmann, Hossain(2011)	
	공저 보정 논문 수	이수상(2013)	
	h-지수	이수상(2013); Abbasi, Altmann, Hossain(2011)	
	인용빈도	김은주와 남태우(2015); Abbasi, Altmann, Hossain(2011); Li, Liao, Yen(2013)	이수상(2013)
	평균인용빈도		김은주와 남태우(2015)

* 여기서 중심성은 여러 연구에서 공통적으로 검토된 연결정도중심성과 매개중심성을 의미함.

성과에 가까운 것은 평균인용빈도이다. 양적 성과와 질적 성과가 혼합되어 있는 것은 h-지수이며 (총) 인용빈도도 논문 수에 따라 늘어날 수 있으므로 질적 성과만이 아닌 양적 성과가 다소 혼합된 변수이다. 공저 관련 변수 중에서 중심성은 공저 네트워크에서 각 저자의 중심성을 측정하는 것으로서, 일관성이 떨어지게 나타난 근접중심성을 제외하고 연결정도중심성과 매개중심성을 변수로 하여 공저 관련 변수와의 상관관계를 검증한 결과만 <표 1>에 포함하였다.

<표 1>의 상관관계 검증 결과를 보면 명확하게 불일치한 결과가 나타난 경우는 연구성과 변수 중 인용빈도와 공저 관련 변수 중 중심성 사이의 상관관계 검증 결과이다. 그중에서도 대조적인 결과는 Abbasi, Altmann, Hossain (2011)과 이수상(2013)의 연구에서 나타났다. 중심성과 인용빈도 사이에 통계적으로 유의한 상관관계가 나타난 Abbasi, Altmann, Hossain (2011)의 연구에서는 5개 iSchool에서 생산된 논문의 저자 1,809명 전체를 대상으로 분석한 반면, 통계적으로 유의한 상관관계가 나타나지 않은 이수상(2013)은 기록관리학 분야 258명의 저자 중 중심성 상위 30명을 대상으로 분석하였다. 즉, Abbasi, Altmann, Hossain(2011)의 연구는 전체 저자를 대상으로 분석한 결과

이고 이수상(2013)의 연구는 상위 저자만을 대상으로 분석한 결과인데 상반된 결과가 나타난 것이다. Abbasi, Altmann, Hossain(2011)의 연구와 일치한 결과를 얻은 Li, Liao, Yen(2013)의 연구도 조사된 연구자 전체를 대상으로 분석한 결과이다. 따라서 분석 대상 집단이 전체인지 상위권 일부인지에 따라서 중심성 지수와 인용빈도 사이의 상관관계 유의성 판단이 달라질 수 있다는 가정을 해볼 수 있다.

이런 가정의 타당성을 검토해보기 위해서 Abbasi, Altmann, Hossain(2011)의 논문에서 제시된 g-지수(Egghe, 2006) 상위 30명의 데이터(p.602, table 2)를 사용하여 이 연구에서 직접 g-지수와 중심성 지수 사이의 상관관계를 검증해보았다. 그 결과 전체 1,809명을 대상으로 분석한 원래 논문의 결과에서는 통계적으로 유의하게 나타났던 g-지수와 중심성 사이의 상관관계가, 상위 30명을 대상으로 이 연구에서 직접 분석한 결과인 <표 2>에서는 통계적으로 유의하지 않은 것으로 나타났다. 원 논문에 g-지수 이외의 연구성과 지수 데이터가 공개되지 않았으므로 이 이상의 분석은 할 수 없었으나, 전체 중 상위 연구자만을 대상으로 지수 간 상관관계를 분석하면 전체를 대상으로 분석했을 때와 다른 결과가 나타날 수 있다는 가능성은 확인되었다.

<표 2> Abbasi, Altmann, Hossain(2011)에 제시된 g-지수 상위 30명을 대상으로 g-지수와 중심성 지수 사이의 스피어맨 상관관계를 검증한 결과

		연결정도중심성	근접중심성	매개중심성
g-지수	스피어맨 상관계수	.332	.020	.270
	유의확률(양측)	.073	.917	.149
	N	30	30	30

3. 연구 설계

3.1 연구 데이터

이 연구에서 사용한 데이터는 한국학술지인용색인 KCI에 2004년부터 2013년까지 10년 동안 <표 3>과 같이 문헌정보학 분야 등재지 9종 학술지에 발표된 논문 3,661편의 공저 데이터와 각 논문이 2013년까지 인용된 횟수 데이터이다.

<표 3> KCI로부터 수집된 학술지별 논문 수 (2004-2013)

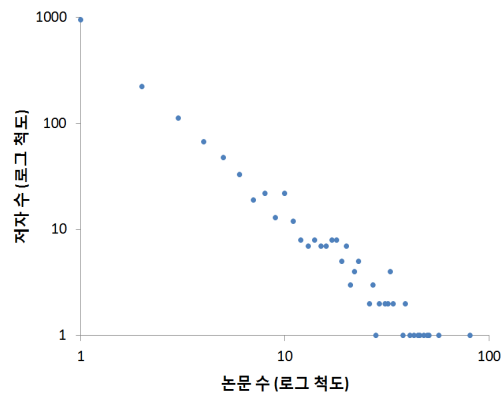
학술지	논문 수
기록학연구	206
서지학보	81
서지학연구	401
정보관리연구	335
정보관리학회지	599
한국기록관리학회지	162
한국도서관·정보학회지	775
한국문헌정보학회지	675
한국비블리아학회지	427
합계	3,661

각 논문의 저자 수를 파악해본 결과는 <표 4>와 같다. 3,661편 중에서 저자가 1명인 논문이 2,290편으로 62.55%를 차지하며, 2명이 공저한 논문이 976편(26.66%), 3명 이상이 공저한 논문이 395편으로 10.79%를 차지하였다.

<표 4> 문헌정보학 분야 논문 3,661편의 저자 수별 논문 수

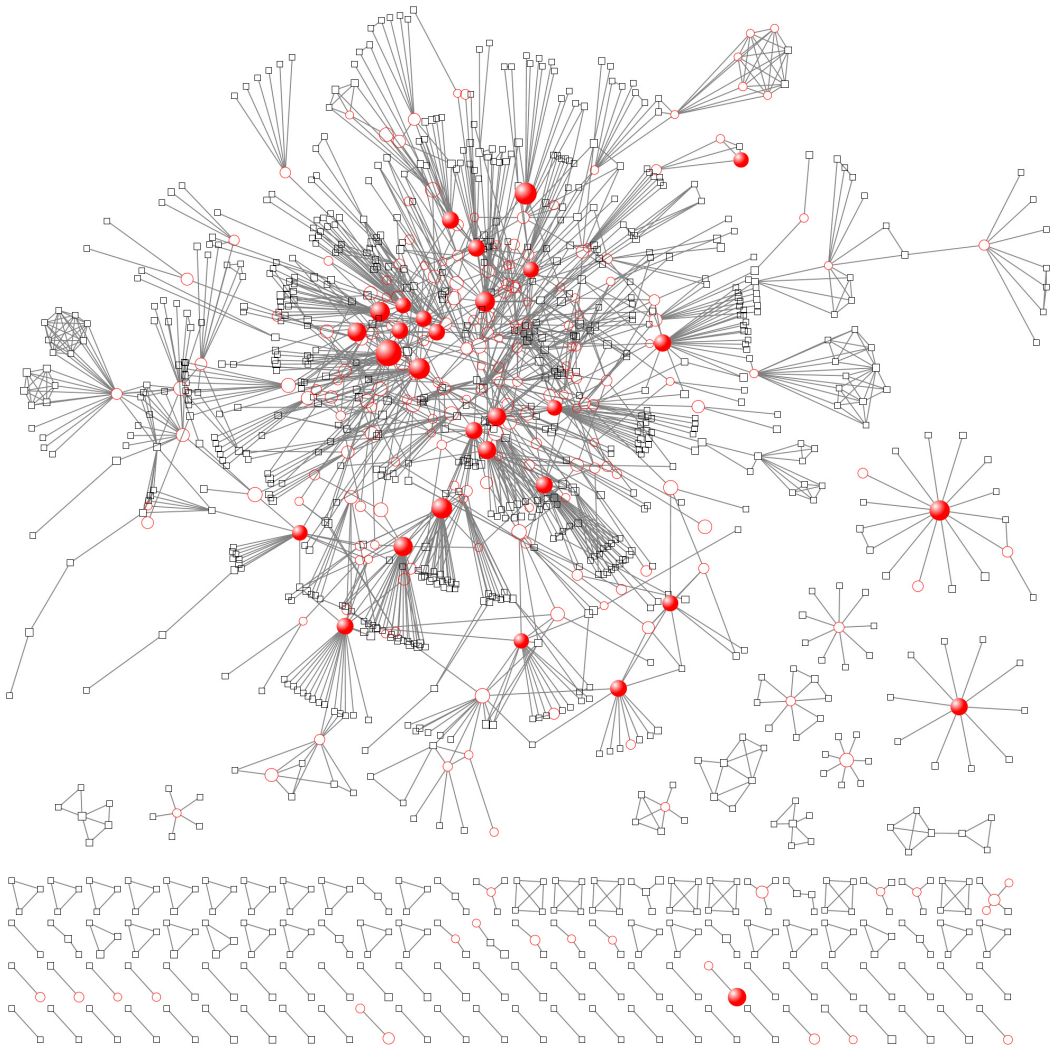
저자 수	1	2	3	4	5	6	7	8	9	합계
논문 수	2,290	976	252	95	30	9	7	1	1	3,661
비율	62.55%	26.66%	6.88%	2.59%	0.82%	0.25%	0.19%	0.03%	0.03%	100.0%

저자 소속과 이메일 주소 등을 이용해서 수작업으로 저자를 식별한 결과 10년 동안 3,661편에 논문을 한 편 이상 투고한 저자는 1,638명으로 확인되었다. 발표 논문 수별 저자 수는 <그림 1>과 같다. 가장 많은 논문을 발표한 저자의 논문 수는 81편이었으며, 그 다음으로 57편, 51편, 50편, 48편을 발표한 저자 순으로 5위 이내에 포함되었다. 논문 수 30위가 26편, 100위가 12편을 발표했으며 5편 이상을 발표한 저자가 276명이었다. 5편 미만의 경우에는 4편 발표 저자가 67명, 3편 발표 저자가 113명, 2편 발표 저자가 223명, 1편 발표 저자가 959명이었다.



<그림 1> 발표 논문 수별 해당 저자 수

전체 저자 1,638명 중에서 단독 저술만 발표한 저자 364명을 제외하고 나머지 저자 1,274명의 공저 네트워크를 구축한 결과를 NodeXL



〈그림 2〉 2004-2013 문헌정보학 분야 9종 학술지 논문 저자들의 공저 네트워크(원의 크기는 발표논문 수와 비례; 단독 저술만 발표한 저자 364명 제외; 작은 사각형은 10년 간 발표한 논문이 5건 미만인 저자, 원은 발표한 논문이 5건 이상 26건 미만인 저자, 채운 원은 발표한 논문이 26건 이상인 저자)

프로그램을 이용해서 〈그림 2〉와 같이 제시하였다. 1,274명 사이의 공저 관계는 2,120개였고 네트워크 밀도는 0.0026으로 매우 낮았다. 이 그림에서 원의 크기는 발표논문 수와 비례하도록 설정하였다. 전체 저자를 발표 논문 수에 따라

서 세 집단으로 구분하였는데, 일단 통계적인 검증의 최소 기준이 되는 5건에 못 미치는 논문을 발표한 저자들은 논문 수 하위 집단으로 사각형으로 표시하였다. 그리고 표본 집단의 최소 규모로 30 이상이 권장되는 것을 고려하여 발표

한 논문이 많은 상위 30명(26건 이상 발표)을 논문 수 상위 집단으로 구분하여 안을 채운 원으로 표시하였다. 논문 수가 5건에서 25건 사이인 저자들은 논문 수 중위 집단으로 구분하여 붉은색 빈 원으로 표시하였다. 대부분의 논문 수 상위 저자는 다수의 저자가 연결된 최대 컴포넌트에 포함되어 있지만, 그림의 오른쪽과 아래에 별도로 독립된 작은 컴포넌트를 이끌고 있는 논문 수 상위 저자가 3명 발견된다.

앞에서 분석 대상 집단이 전체인지 아니면 상위권 일부인지에 따라서 중심성 지수와 인용빈도 사이의 상관관계 유의성 판단이 달라질 수 있다는 가정을 하였다. 이에 따라 이후의 분석에서는 대상 집단의 범위를 저자가 발표한 논문 수를 기준으로 세 가지로 다르게 설정하였다. 전체 1,638명에 대한 검증과 함께 논문 수 5건 이상인 276명에 대한 검증, 그리고 논문 수 상위 30명에 대해서만 검증도 별도로 수행하였다. 이를 통해 선행 연구에서 상이한 결과가 나타났던 원인으로 추정되는 분석 대상 집단의 범위를 통제하는 효과를 거두고자 하였다.

또한 선행 연구 중에서 박치성(2012)은 단독 저술만 발표한 218명은 분석에서 제외하였고, 이수상(2013)의 연구도 공저 네트워크에 포함되지 않은 단독 저술만 발표한 저자를 제외하고 분석한 바 있다. 그러나 단독 저술만 발표한 연구자를 분석에서 제외하게 되면 공저 여부와 연구성과 사이의 상관관계를 제대로

분석하지 못하는 한계가 있게 된다. 따라서 이 연구에서는 단독 저술만 발표한 연구자와 공저 네트워크에서 분리된 독립적인 컴포넌트에 속한 연구자도 모두 포함하여 분석을 수행하였다.

3.2 측정 변수

이 연구에서 각 연구자의 공저 관련 변수로 사용한 것은 협력계수 CC(collaboration coefficient; Aijferuke, Burell, & Tague, 1988), 공저 네트워크에서의 연결정도중심성, 근접중심성, 매개중심성, SSR(Sum of Square Root; 이재운, 2014), 이웃중심성(이재운, 2015)과 삼각매개중심성(이재운, 2006a)이다.

협력계수 CC는 연구자가 단독저술만 발표하면 0이 되고 공동저술의 비중이 커질수록, 그리고 논문 당 공저자의 수가 많아질수록 1에 가깝게 높아지는 지표이다. 이를 이용해서 연구자의 공저 성향을 측정하였다. 선행 연구에서는 단독저술의 비율이나 공저한 저자 수와 같은 단순한 변수를 사용하였으나 공저 성향을 섬세하게 반영하지 못하는 한계가 뚜렷하므로 이 연구에서는 협력계수를 산출하여 변수로 사용하였다.

공저 네트워크에서 측정된 중심성 지수 6가지의 특성을 정리해보면 <표 5>와 같으며 측정 범위와 이진/가중 네트워크 고려 여부에 따라

<표 5> 연구성과와 비교된 중심성 지수

구분	연결정도 중심성	SSR	이웃중심성	매개중심성	근접중심성	삼각매개 중심성
측정 범위	지역	지역	지역	전역	전역	전역
이진/가중	이진	가중	가중	이진	이진	가중

서 고르게 구성되어 있음이 확인된다. 이 중에서 연결정도중심성, 매개중심성, 근접중심성은 기존 연구에서도 연구성과와의 상관성이 검토된 이진 네트워크 중심성 지수이다. SSR, 이웃중심성, 삼각매개중심성은 가중 네트워크를 위한 중심성 지수이다. 특히 SSR은 각 공동연구자와의 협력 건수의 제곱근을 합산한 값으로서 공저 네트워크에서 공저의 양과 공저의 폭을 균형있게 반영하는 지수라고 확인된 바 있다(이재윤, 2015).

각 연구자의 연구성과 변수로 사용한 것은 논문 수, h-지수(Hirsch, 2005), hs-지수(이재윤, 2006b), 인용빈도, 평균인용빈도의 5가지이다. 또한 이들 각각에 대해서 공저자 수에 따른 공저 기여도를 감안하여 보정한 보정 논문 수, 보정 h-지수, 보정 hs-지수, 보정 인용빈도, 보정 평균인용빈도를 추가로 산출하였다.

공저 기여도 산출은 논문 수와 같은 양적 성과나 인용빈도와 같은 질적 성과를 분수 계상(fractional counting)하기 위한 것이다. 이 연구에서는 제1저자와 다른 공저자의 기여도를 차별화하되, 제2저자 이후의 다른 공저자는 모두 동일한 기여도를 가지도록 하는 새로운 방식을 사용하였다. 이 방법은 Lee와 Chung(2014)의 복수저자 기여도 산정 방식 비교 연구에서 저자동시인용분석에 적용하였을 때 결과의 내적 타당도가 가장 높게 나타난 제1저자 가중 산식과 같은 발상이다. 교신저자의 기여도도 높게 설정하는 것을 고려해볼 수도 있으나 서지 데이터베이스에서 교신저자를 별도로 구분하고 있지 않기 때문에 이를 적용하지는 않았다. 이에 따라 저자 수가 n명일 때, 제1저자는 공저자의 수가 많을수록 감소하는 기여도를 가지게

하고 나머지 공저자는 제1저자의 기여도를 1에서 뺀 나머지를 균등하게 나눠가지도록 다음과 같이 공식을 고안하였다.

$$\text{제1저자의 기여도} = \frac{1}{\sum_{r=1}^n \frac{1}{r}}$$

$$\text{공동저자 한 명의 기여도} = \frac{1}{n-1} \times \left(1 - \frac{1}{\sum_{r=1}^n \frac{1}{r}} \right)$$

예를 들어 저자가 3명일 경우, 제1저자는 $1 / (1 + 1/2 + 1/3) = 0.545$ 의 기여도를 가지고, 제2저자와 제3저자는 $(1-0.545) / 2 = 0.227$ 의 기여도를 각각 가지게 된다. 이 방법에서는 한 논문의 저자들의 기여도 합계는 항상 1이 된다. 이와 유사한 Lee와 Chung(2014)의 복수저자 기여도 산정 방식에서는 제1저자에게 항상 기여도 1을 부여함으로써 한 논문의 저자들의 기여도 합계가 1에서 2 사이로 유동적이다. Lee와 Chung(2014)은 동시인용분석을 통해 저자들의 연구 주제를 파악하는 것이 목적이었지만, 이 연구에서는 연구성과 비교를 하는 것이 목적이므로 논문 한 편의 기여도 합계가 일정하게 되도록 고려하였다.

보정 논문 수는 각 논문마다 이렇게 산출한 기여도만큼을 합산하여 산출하였다. 인용에 기반하여 산출하는 나머지 연구성과 변수들은 각 논문의 실제 인용빈도에 기여도를 곱한 만큼만을 인용값으로 사용하여 이를 합산한 보정 인용빈도, 평균을 구한 보정 평균인용빈도를 산출하였다. 기여도를 곱한 인용값에 기반하여 보정 h-지수, 보정 hs-지수를 산출하는 과정은 이재윤(2016)에 따랐다.

이상의 변수를 사용하여 1,638명 연구자 각각의 공저 활동 내역과 연구성과를 측정된 후에 다음의 네 가지 분석을 실시하였다.

첫째, 논문별 저자 수에 따른 인용빈도의 차이를 분석하였다. 이를 통해 저자 수가 많은 논문일수록 인용이 많이 되는지를 확인해보았다.

둘째, 각 저자별로 발표한 논문 수와 인용지수 사이의 상관관계를 분석하였다. 특히 공저를 보정한 경우와 보정하지 않은 경우를 각각 분석하여 비교해봄으로써 공저 보정의 효과를 확인하였다.

셋째, 협력계수와 연구성과 사이의 상관관계를 분석하였다. 이를 통해 공저가 활발한 정도와 양적 및 질적 연구성과 사이에 어떤 상관성이 있는지 확인하였다.

넷째, 중심성 지수와 연구성과 사이의 상관관계를 분석하였다. 공저 네트워크에서 각 저자의 입지를 다양한 관점에서 측정된 6가지 중심성 지수와 양적 및 질적 연구성과 사이의 상관관계를 검증함으로써 선행 연구의 결과를 재확인하고 상반된 결과가 나타났던 경우에 대한 원인 분석을 시도해보았다.

이 연구에서 측정된 논문 수나 인용빈도와 같은 연구성과 변수는 정규분포가 아닌 것으로 알려져 있으므로 상관관계 검증은 모두 비모수 상관관계수인 스피어맨 상관계수를 사용하였다.

4. 분석 결과

4.1 논문별 저자 수에 따른 인용빈도 분석

저자 수가 1명인 논문은 2,290건, 2명인 논문은 976건, 3명 이상인 논문은 395건이었다. 이 세 집단의 인용빈도 평균은 각각 2,576, 2,436, 2,443으로 비슷하게 나타났으며 인용빈도 최댓값도 28, 22, 25로 큰 차이가 없었다. <표 6>에 나타난 인용빈도 분포를 요약한 수치도 거의 같게 나타났다.

저자 수가 각각 1명, 2명, 3명 이상인 세 논문 집단 사이의 인용빈도 차이를 비모수 검정 기법인 윌콕슨 순위합 검정으로 분석한 결과 p값이 0.8988로 저자 수별 논문 집단 간 차이가 유의하지 않게 나타났다. 이는 여러 학문분야를 대상으로 분석한 Franceschet와 Costantini(2010)의 연구와는 상이한 결과이다. Franceschet와 Costantini(2010)가 연구한 분야 중 비교 대상 논문 집합을 나누는 기준을 저자 1명, 2명, 그리고 3명 이상으로 적용한 분야인 수학/컴퓨터과학, 도시공학/건축, 경제학/통계학 등 세 영역에서는 모두 저자 수가 많은 논문 집합이 평균인용빈도도 더 높은 것으로 나타났다. 또한 Gazni와 Didegah(2011)가 Harvard 대학 연구자들의 10년 간 발표 논문 124,937 편을 분석한 결과에

<표 6> 저자 수별 논문 집합의 인용 빈도 요약 통계량

저자 수	해당 논문 수	인용 빈도 요약 수치					
		최솟값	1사분위수	중앙값	평균	3사분위수	최댓값
1명	2,290	0	0	1	2,576	4	28
2명	976	0	0	1	2,436	4	22
3명 이상	395	0	0	1	2,443	4	25

서도 분야를 막론하고 단독 논문보다 공저 논문의 인용빈도가 높게 나타났다. 그러나 국내 문헌정보학 분야에서는 논문의 저자 수에 따른 인용빈도의 차이가 없는 것으로 판단된다. 이런 차이의 원인에 대해서는 결론 부분에서 논의해보았다.

만약 저자의 수가 논문의 인용빈도를 좌우하는 것으로 나타났다면, 공동저술을 많이 하는 저자일수록 인용 지수가 높아질 가능성이 크다. 그러나 그렇지 않았기 때문에 이후의 분석에서 저자의 공동저술 활동에 따른 통계량(협력 계수, 공저 네트워크 중심성 등)이 인용빈도를 비롯한 인용 지수와 관련이 있는 것으로 나타난다면, 공동저술과 관련된 개별 저자의 활동 내용이 인용 지수에 영향을 끼치는 것으로 판단할 수 있을 것이다.

4.2 논문 수와 인용지수의 상관관계 분석

4.2.1 단순 논문 수와 인용지수의 상관관계 분석

이 절에서는 연구자가 발표한 논문 수가 인용지수와 어떤 상관관계가 있는지를 검증해보았다. 연구자별 인용지수로는 총 인용빈도, 평균인용빈도, h-지수, hs-지수의 네 가지 인용지수를 산출하였으며 이 네 가지 인용지수가 연

구자의 논문 수와 가지는 상관계수를 통계적으로 검증하였다. 검증 대상 집단은 전체 1,638명, 논문 5건 이상 276명, 논문 수 최상위 30명을 대상으로 하였다. 비모수 상관계수인 스피어맨 순위 상관계수를 산출하고 검증한 결과는 <표 7>과 같다.

인용빈도는 세 집단 모두에서 논문 수와 상관관계가 유의하게 나타났다. 총 인용빈도가 각 논문의 인용빈도를 합산한 것이므로 논문 수와 상관관계가 있는 것은 당연한 결과이다. 이뿐만 아니라 전체 1,638명과 논문 5건 이상 276명을 대상으로 분석하였을 때, 논문 수는 모든 인용지수와 상관관계가 유의한 것으로 나타났다. 그러나 논문 26건 이상인 상위 저자 30명을 대상으로 분석하였을 때에는 인용빈도만 논문 수와 상관관계가 유의하게 나타났고, 평균인용빈도, h-지수, hs-지수는 논문 수와 상관관계가 유의하지 않게 나타났다.

평균인용빈도는 총 인용빈도를 논문 수로 나눈 것이므로 이론적으로 평균인용빈도와 논문 수는 상관성이 없어야 한다. 그러나 논문 수 26건 이상인 핵심 저자 30명에 대한 분석에서만 상관관계가 유의하지 않게 나타났고, 저자들의 논문 수 편차가 큰 두 집합에서는 논문 수와 평균인용빈도 사이에 상관관계가 유의하게 나타났다. 이는 활발하게 논문을 발표하는 저자의

<표 7> 논문 수와 인용지수의 스피어맨 순위 상관계수

분석 대상	논문 수와 비교한 인용 지수			
	인용빈도	평균인용빈도	h-지수	hs-지수
전체 1,638명	.699**	.300**	.726**	.678**
논문 5건 이상 276명	.797**	.225**	.710**	.689**
논문 26건 이상 30명	.479**	-.022	.180	.206

** . 상관관계가 0.01 수준에서 유의함(양측).

논문이, 평균 이하로 소수의 논문을 발표하는 저자의 논문보다 인용빈도가 조금이라도 높을 가능성이 있음을 나타낸다. 이런 결과에는 저자의 자기인용 등도 일부 작용했을 가능성이 있다. 다만 논문 수가 26건 이상으로 활발하게 논문을 발표하는 저자들의 논문들은 인용될 기회의 편차가 거의 없으므로 평균인용빈도와 논문 수의 상관성은 나타나지 않는다.

평균인용빈도와 달리 h-지수와 hs-지수는 본질적으로 연구성과의 양적인 측면과 질적인 측면을 복합적으로 반영하는 지수이므로 논문 수와 통계적인 상관성이 있어야 한다. 그러나 핵심 저자 30명에 대한 분석에서는 상관성이 나타나지 않았다. 그 원인에 대해서는 공저를 고려한 논문 수와 인용지수 사이의 관계를 분석한 다음 단락에서 추가 논의해보기로 한다.

4.2.2 공저를 보정한 논문 수와 인용지수의 상관관계 분석

이번에는 단순 논문 수가 아니라 공저를 고려하여 보정한 논문 수와 보정한 인용지수의 상관관계를 검증해보았다. <표 8>에 제시된 검증 결과를 보면 보정한 논문 수 역시 전체 1,638명 저자 집합과 논문 5건 이상 276명을 대상으로 검증하였을 때, 모든 인용지수와 상관관계가 유의

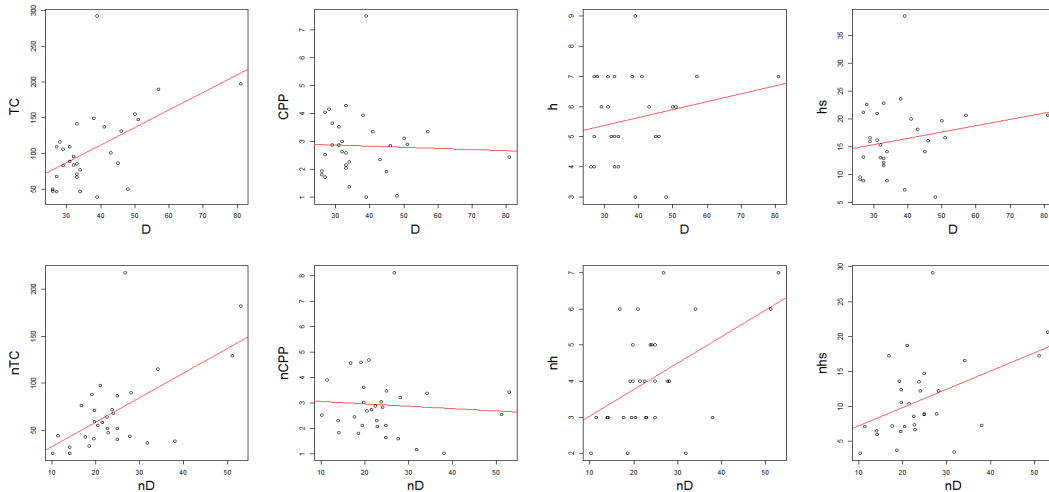
한 것으로 나타났다.

그러나 논문 26건 이상인 핵심 저자 30명을 대상으로 분석하였을 때에는 보정 평균인용빈도만 보정 논문 수와 상관관계가 유의하지 않았고, 보정 h-지수와 보정 hs-지수는 보정 논문 수와 상관관계가 유의한 것으로 나타났다. 상관관계를 산점도로 표현한 <그림 3>을 살펴보면, 보정하지 않은 논문 수와 h-지수 및 hs-지수 사이의 상관관계를 보여주는 윗 줄 오른쪽의 두 산점도에 비해서, 보정한 변수 사이의 상관관계를 보여주는 아랫줄 오른쪽의 두 산점도에서 각 점이 회귀선에 더 가깝게 분포하고 있음이 드러난다. 앞에서 언급했듯이, 평균인용빈도는 논문 수와 상관성이 없고 h-지수와 hs-지수는 상관성이 있다는 이론적인 예측과 일치하는 결과이다. 논문 발표가 활발한 저자 집단에 대해 분석한 결과에서 공저를 보정하지 않은 논문 수와 인용지수 사이의 상관관계가 부분적으로 이론에 부합되지 않은 반면에, 보정한 논문 수와 보정한 인용지수 사이의 상관관계는 이론과 부합하는 결과가 나타났다. 이는 공저를 고려하여 논문 수와 인용지수를 측정하는 것이 현상을 설명하는데 더 적합할 수 있음을 시사한다.

<표 8> 보정 논문 수와 보정 인용지수의 스피어맨 순위 상관계수

분석 대상	보정 논문 수와 비교한 인용 지수			
	보정 인용빈도	보정 평균인용빈도	보정 h-지수	보정 hs-지수
전체 1,638명	.677**	.215**	.722**	.714**
논문 5건 이상 276명	.815**	.143*	.755**	.734**
논문 26건 이상 30명	.464**	-.102	.430*	.464**

** . 상관관계가 0.01 수준에서 유의함(양측).
 * . 상관관계가 0.05 수준에서 유의함(양측).



〈그림 3〉 논문 수 상위 30명에 대해서 논문 수 D(가로 축)와 4가지 인용지수(인용빈도 TC, 평균인용빈도 CPP, h-지수, hs-지수) 사이의 산점도. 윗줄은 공저를 보정하지 않은 값의 산점도이고, 아랫줄은 공저를 보정한 값의 산점도이며, 실선은 회귀선임

4.3 공저 성향과 연구성과의 상관관계 분석

협력계수 CC는 단독저술만 하면 0이 되고 공저를 활발하게 할수록 1에 가깝게 높아지는 지수이다. 이를 이용해서 단독저술보다 공저 위주의 저자가 논문 수 및 인용지수도 높은지 여부를 검증해보았다. 〈표 9〉에는 공저를 보정하지 않은 연구성과 지수와 협력계수 사이의 상관관계 검증 결과를 제시하였고, 〈표 10〉에는 공저를 보정한 연구성과 지수와 협

력계수 사이의 상관관계 검증 결과를 제시하였다.

〈표 9〉를 보면 협력계수와 논문 수의 상관관계는 전체 저자 집합의 경우에만 통계적으로 유의하고 다른 두 경우는 통계적인 상관성이 나타나지 않았다. 이는 협력계수가 높은 연구자, 즉 단독 저술이 아닌 공저를 주로 하는 연구자가 논문을 더 많이 발표하는 것은 아님을 의미한다. 오히려 〈표 10〉을 보면 공저를 감안하여 보정한 보정 논문 수는 협력계수와 부정적인 상관관계를 보이고 있다.

〈표 9〉 협력계수와 성과 지수의 스피어맨 순위 상관계수

분석 대상	협력계수 CC와 비교한 성과 지수				
	논문 수	인용빈도	평균인용빈도	h-지수	hs-지수
전체 1,638명	-.207**	-.115**	-.007	-.135**	-.108**
논문 5건 이상 276명	-.100	.035	.112	.075	.080
논문 26건 이상 30명	.147	.200	.071	.002	.120

** : 상관관계가 0.01 수준에서 유의함(양측).

〈표 10〉 협력계수와 보정된 성과 지수의 스피어맨 순위 상관계수

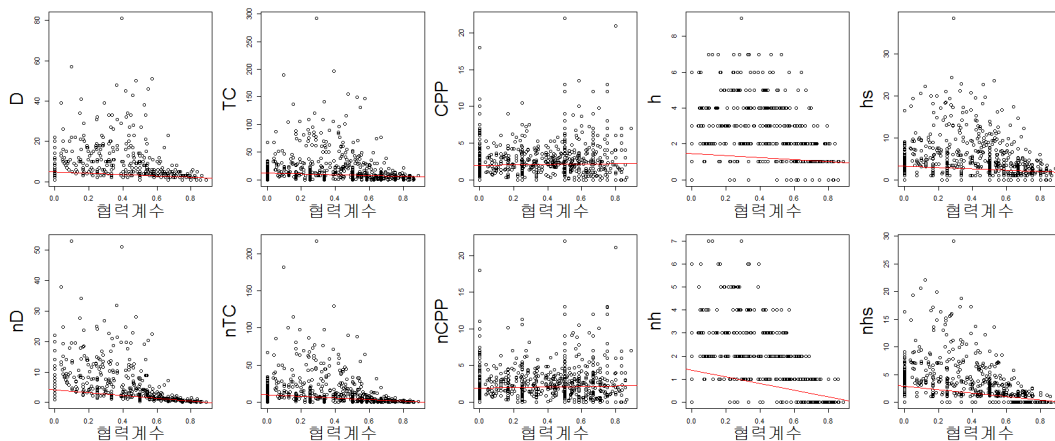
분석 대상	협력계수 CC와 비교한 보정 인용 지수				
	보정 논문 수	보정 인용빈도	보정 평균인용빈도	보정 h-지수	보정 hs-지수
전체 1,638명	-.686**	-.327**	-.026	-.445**	-.435**
논문 5건 이상 276명	-.476**	-.266**	.123*	-.353**	-.337**
논문 26건 이상 30명	-.597**	-.224	.071	-.450*	-.455*

** . 상관관계가 0.01 수준에서 유의함(양측).

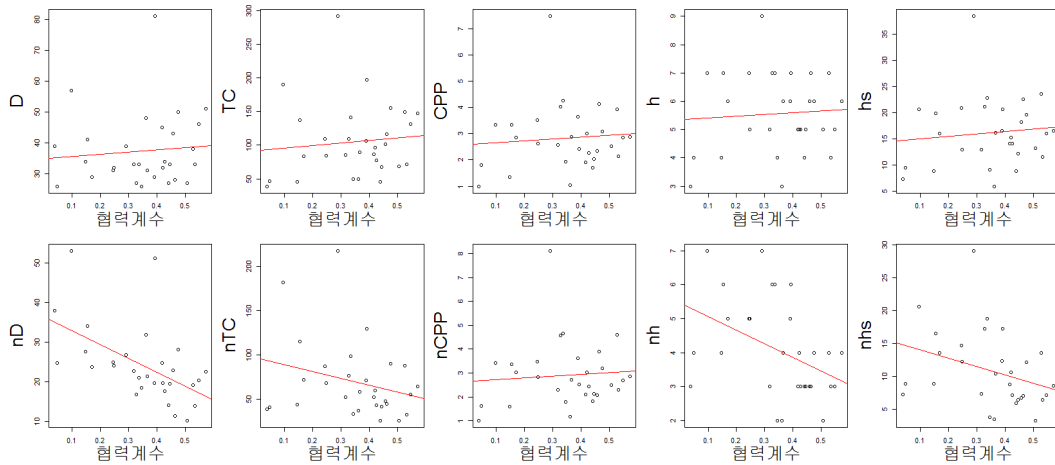
* . 상관관계가 0.05 수준에서 유의함(양측).

인용지수와 협력계수의 상관관계도 논문 5건 이상 276명 집단과 상위 30명 집단에서 모두 통계적으로 유의하지 않게 나타났다. 이는 정보시스템 분야에서 공저자 수와 인용빈도 사이에 상관관계가 통계적으로 유의하지 않았다고 보고한 Li, Liao, Yen(2013)의 연구와 일치한 결과이다. 오히려 공저를 보정한 〈표 10〉에서는 논문 수가 많을수록 유리한 보정 인용빈도와 보정 h-지수 및 보정 hs-지수는 모든 집단에서 협력계수와 통계적으로 부정적인 상관관계가 유의하게 나타났다. 그리고 총 인용빈도도 동

일한 경향을 보이는 것이 〈그림 4〉와 〈그림 5〉에서 확인된다. 특히 〈그림 5〉에서는 공저를 보정하지 않은 윗줄 그림에서 거의 평행했던 회귀선이, 공저를 보정한 아랫줄 그림에서는 세 번째의 평균인용빈도를 제외하면 양적인 성과가 반영되는 나머지 네 가지 지수에서 뚜렷하게 우하향하는 경향을 보이고 있다. 이는 공저를 고려하여 논문 수를 산출할 경우에는 오히려 단독 저술 위주의 저자가 공저 위주의 저자보다 양적인 연구성과가 더 높다는 것을 의미한다.



〈그림 4〉 전체 저자에 대해서 협력계수 CC(가로 축)와 5가지 성과지표(논문 수 D, 인용빈도 TC, 평균인용빈도 CPP, h-지수, hs-지수) 사이의 산점도. 윗줄은 공저를 보정하지 않은 값의 산점도이고, 아랫줄은 공저를 보정한 값의 산점도이며, 실선은 회귀선임



〈그림 5〉 논문 수 상위 30명에 대해서 협력계수 CC(가로 축)와 5가지 성과지표(논문 수 D, 인용빈도 TC, 평균인용빈도 CPP, h-지수, hs-지수) 사이의 산점도. 윗줄 그림들은 공저를 보정하지 않은 값의 산점도이고, 아랫줄 그림들은 공저를 보정한 값의 산점도이며, 실선은 회귀선임

단독 저술만 발표한 연구자를 분석에서 제외하였을 때 공저 여부와 연구성과 사이의 상관관계가 달라지는 지를 알아보기 위해서 협력계수가 0인 저자, 즉 단독 저술만 수행한 저자를 제외한 후에 협력계수와 연구성과 사이의 상관관계를 산출해보았다. 논문 26건 이상 저자 중에는 단독 저술만 발표한 저자가 없으므로 실제 차이는 논문 5편 이상인 저자 집단과 전체 저자 집단에서 발견할 수 있다. 〈표 11〉과 〈표 12〉를 각각 〈표 9〉, 〈표 10〉과 비교해보면 스피어만 순위 상관계수의 수치가 전체적으로 낮아졌을 뿐만 아니라 통계적인 유의성 여부가 달라진 경우도 나타났다.

이러한 경우, 〈표 9〉에서는 논문 5건 이상 발표한 저자들의 경우에 협력계수와 논문 수, 협력계수와 인용빈도 사이의 상관관계가 통계적으로 유의하지 않았으나, 〈표 11〉에서는 모두 부정적인 상관관계가 0.01 수준에서 유의한 것으로 바뀌었다. 즉, 공저를 전혀 안한 저자들을 제외하면 공저와 연구성과 사이의 상관성이 더

어떤 순위 상관계수의 수치가 전체적으로 낮아졌을 뿐만 아니라 통계적인 유의성 여부가 달라진 경우도 나타났다. 이를테면, 〈표 9〉에서는 논문 5건 이상 발표한 저자들의 경우에 협력계수와 논문 수, 협력계수와 인용빈도 사이의 상관관계가 통계적으로 유의하지 않았으나, 〈표 11〉에서는 모두 부정적인 상관관계가 0.01 수준에서 유의한 것으로 바뀌었다. 즉, 공저를 전혀 안한 저자들을 제외하면 공저와 연구성과 사이의 상관성이 더

〈표 11〉 단독 저술만 수행한 저자를 제외하고 산출한 협력계수와 성과 지수의 스피어만 순위 상관계수(음영 부분은 제외하기 전과 유의성 여부가 달라진 경우)

분석 대상	협력계수 CC와 비교한 성과 지수				
	논문 수	인용빈도	평균인용빈도	h-지수	hs-지수
전체 1,274명	-0.442**	-0.321**	-0.106**	-0.348**	-0.310**
논문 5건 이상 243명	-0.239**	-0.131*	-0.012	-0.084	-0.085
논문 26건 이상 30명	.147	.200	.071	.002	.120

** : 상관계수가 0.01 수준에서 유의함(양측).

<표 12> 단독 저술만 수행한 저자를 제외하고 산출한 협력계수와 보정된 성과 지수의 스피어맨 순위 상관계수(음영 부분은 제외하기 전과 유의성 여부가 달라진 경우)

분석 대상	협력계수 CC와 비교한 보정 인용 지수				
	보정 논문 수	보정 인용빈도	보정 평균인용빈도	보정 h-지수	보정 hs-지수
전체 1,274명	-.710**	-.465**	-.117**	-.533**	-.535**
논문 5건 이상 243명	-.552**	-.394**	.001	-.451**	-.449**
논문 26건 이상 30명	-.597**	-.224	.071	-.450*	-.455*

** . 상관관계가 0.01 수준에서 유의함(양측).

* . 상관관계가 0.05 수준에서 유의함(양측).

부정적인 방향으로 바뀌어 측정되는 것이다. 따라서 공저 여부와 연구성과 사이의 상관관계를 제대로 파악하기 위해서는 단독 저술만 발표한 연구자도 포함해야 한다.

4.4 중심성 지수와 연구성과의 상관관계 분석

공저 네트워크에서 각 저자의 입지를 측정하는 중심성 지수와 연구성과가 어떤 상관성을 나타내는지를 살펴보았다. 이하의 분석에서는 전체 1,638명을 대상으로 변수 간 상관관계를 분석한 결과와 논문 5건 이상인 236명을 대상으로 변수 간 상관관계를 분석한 결과의 통계적인 유의성 판정이 대부분 일치하였다. 예를 들어 논문 수와 6개 중심성 지수 사이의 상관관계는 전체 1,638명을 대상으로 분석한 결과와 논문 5건 이상인 236명을 대상으로 분석한 결과 모두에서 동일하게 0.01 수준에서 유의하게 나타났다. 또한 발표한 논문이 5건에도 미치지 못하는 연구자는 연구성과를 평가하기에는 근거가 부족한 편이다. 예를 들어 인용빈도 0인 저자는 모두 논문 수가 5편 미만이었으며, 논문 수가 5편 이상인 저자 중에서 인용빈도가 0인 저자

는 없었다. 따라서 전체 1,638명을 대상으로 분석한 결과는 분량 관계상 생략하고, 5건 이상 논문을 발표한 저자 236명을 대상으로 한 분석 결과와 논문수 상위 30명을 대상으로 한 분석 결과를 제시하였다. 우선 양적 성과인 논문 수와 중심성 지수의 상관관계를 분석한 이후에 질적 성과인 인용지수와 중심성 지수의 상관관계를 분석해보았다.

4.4.1 중심성 지수와 논문 수의 상관관계

논문 수 5건 이상 저자 276명과 논문 수 상위 30명에 대해서 각각 공저를 보정하지 않은 논문 수와 중심성 지수의 스피어맨 순위 상관계수는 <표 11>에, 산점도는 <그림 6>에 각각 제시하였다. 공저를 보정한 논문 수와 중심성 지수의 스피어맨 순위 상관계수는 <표 12>에, 산점도는 <그림 7>에 각각 제시하였다.

논문 5건 이상 발표한 276명을 대상으로 분석했을 때에는 근접중심성을 제외하면 모든 중심성 지수가 보정하지 않은 논문 수 및 보정한 논문 수와 상관관계가 유의하게 나타났다. 근접중심성은 공저 네트워크에서 최대 컴포넌트에 속하지 않고 분리되거나 고립된 작은 컴포넌트에 속한 저자가 지나치게 낮은 값을 가

지게 되므로(〈그림 6〉과 〈그림 7〉의 왼쪽에서 다섯 번째 근접중심성 그림에서 왼쪽에 치우친 저자들) 다소 일관되지 않은 결과가 나타났다. 결과적으로 논문 수 5명 이상 저자 집단에 대해서는 공저 보정에 상관없이 논문 수와 중심성 지수가 통계적으로 유의한 상관관계를 보였다.

논문 수 상위 30명에 대한 분석에서는 다른 결과가 나타났다. 공저를 보정하지 않은 〈표 13〉의 결과에서는 SSR만 논문 수와 상관관계가 유의하게 나타났으며 다른 중심성 지수는 논문 수와 상관성이 없었다. 공저를 보정한 〈표 14〉의 결과에서는 모든 중심성 지수가 공저 보정한 논문 수와 부정적 상관성을 보였으며, SSR을 제외한 5개 중심성 지수는 통계적으로 유의한 부정적 상관관계가 나타났다. SSR은 공저의 양이 직접 반영되는 지수이기 때문에 다른 중심성 지수와 다소 다른 결과가 나타난 것으로 추정된다.

전체적으로 선행 연구와 마찬가지로 중심성 지수가 높은 저자가 논문 수도 많은 것으로 나타났다. 이는 전체 저자를 대상으로 분석한 Abbasi, Altmann, Hossain(2011)의 연구와 임병학(2013)의 연구, 그리고 이수상(2013)의 연구에서와 일치하는 결과이다. 그러나 논문 수 상위 저자 집단의 경우에는 SSR을 제외한 모든 중심성 지수와 논문 수 사이의 상관관계가 통계적으로 유의하지 않았다. 공저를 보정한 경우에는 연결정도중심성, 이웃중심성, 매개중심성, 삼각매개중심성이 보정 논문 수와 오히려 부정적인 상관관계가 유의한 것으로 나타났다. 이로써 논문 수 상위 저자들의 경우에는 대부분의 중심성 지수가 논문 수와 상관성이 없으며 공저를 고려했을 때에는 오히려 보정 논문 수와 대부분의 중심성 지수가 부정적인 상관관계를 보인다고 결론지을 수 있다.

〈표 13〉 중심성 지수와 논문 수의 스피어맨 순위 상관계수

	논문 수와 비교한 중심성 지수					
	연결정도 중심성	SSR	이웃중심성	매개중심성	근접중심성	삼각매개 중심성
논문 5건 이상 276명	.469**	.473**	.457**	.469**	.302**	.512**
논문 26건 이상 30명	.246	.388*	.265	.099	.149	.277

** . 상관관계가 0.01 수준에서 유의함(양측).

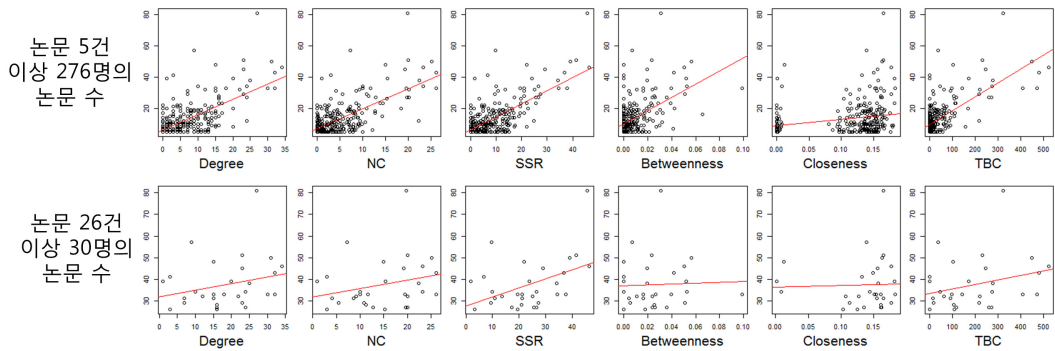
* . 상관관계가 0.05 수준에서 유의함(양측).

〈표 14〉 중심성 지수와 보정 논문 수의 스피어맨 순위 상관계수

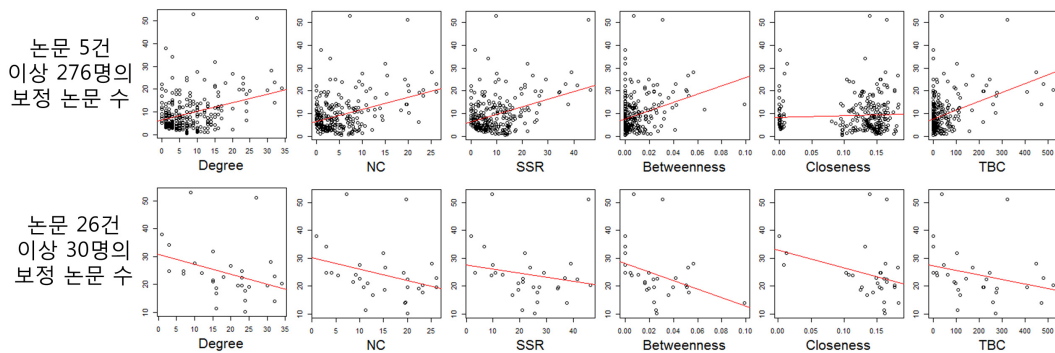
	보정 논문 수와 비교한 중심성 지수					
	연결정도 중심성	SSR	이웃중심성	매개중심성	근접중심성	삼각매개 중심성
논문 5건 이상 276명	.149*	.136*	.152*	.198**	.042	.209**
논문 26건 이상 30명	-.450*	-.304	-.429*	-.465**	-.306	-.432*

** . 상관관계가 0.01 수준에서 유의함(양측).

* . 상관관계가 0.05 수준에서 유의함(양측).



〈그림 6〉 논문 수(세로축)와 중심성 지수 6종(가로축) 사이의 산점도. 위는 논문 수가 5건 이상인 276명 저자의 분석 결과이고, 아래는 논문 수가 26건 이상인 30명 저자의 분석 결과임(실선은 회귀선)



〈그림 7〉 보정 논문 수(세로축)와 중심성 지수 6종(가로축) 사이의 산점도. 위는 논문 수가 5건 이상인 276명 저자의 분석 결과이고, 아래는 논문 수가 26건 이상인 30명 저자의 분석 결과임(실선은 회귀선)

4.4.2 중심성 지수와 인용지수의 상관관계

공저를 보정하지 않은 인용지수 4가지(인용빈도, 평균인용빈도, h-지수, hs-지수)와 함께 이들 각각에 대해서 공저를 보정한 인용지수 4가지를 추가하여 총 8가지 인용지수와 중심성 지수 6종 사이의 상관관계를 분석하였다. 〈표 15〉에는 논문 수 5건 이상 저자 276명에 대해서 산출한 스피어맨 상관계수를 제시하였고 〈표 16〉에는 논문 수 26건 이상 저자 30명에 대해서

산출한 스피어맨 상관계수를 제시하였다.

중심성 지수와 인용지수의 상관관계를 살펴보면, 논문수 5건 이상 저자 276명을 대상으로 한 분석의 결과와 논문 수 최상위 저자 30명에 대한 분석의 결과는 상당히 다르게 나타났다.

〈표 15〉의 논문수 5건 이상 저자 276명을 대상으로 한 분석 결과에서 인용빈도, 평균인용빈도, h-지수, hs-지수, 보정 인용빈도, 보정 평균인용빈도는 모두 6가지 중심성 지수와 통계

〈표 15〉 중심성 지수와 인용지수의 스피어맨 순위 상관계수(논문 수 5건 이상 저자 276명)

	인용빈도	평균 인용빈도	h-지수	hs-지수	보정 인용빈도	보정 평균 인용빈도	보정 h-지수	보정 hs-지수
연결정도중심성	.493**	.243**	.470**	.466**	.273**	.265**	.143*	.160**
SSR	.495**	.239**	.477**	.470**	.260**	.261**	.127*	.142*
이웃중심성	.465**	.214**	.454**	.441**	.260**	.240**	.129*	.146*
매개중심성	.517**	.304**	.505**	.502**	.333**	.317**	.222**	.236**
근접중심성	.384**	.280**	.369**	.388**	.183**	.265**	.102	.115
삼각매개중심성	.522**	.255**	.493**	.489**	.320**	.272**	.187**	.200**

** . 상관관계가 0.01 수준에서 유의함(양측).

* . 상관관계가 0.05 수준에서 유의함(양측).

〈표 16〉 중심성 지수와 인용지수의 스피어맨 순위 상관계수(논문 수 26건 이상 저자 30명)

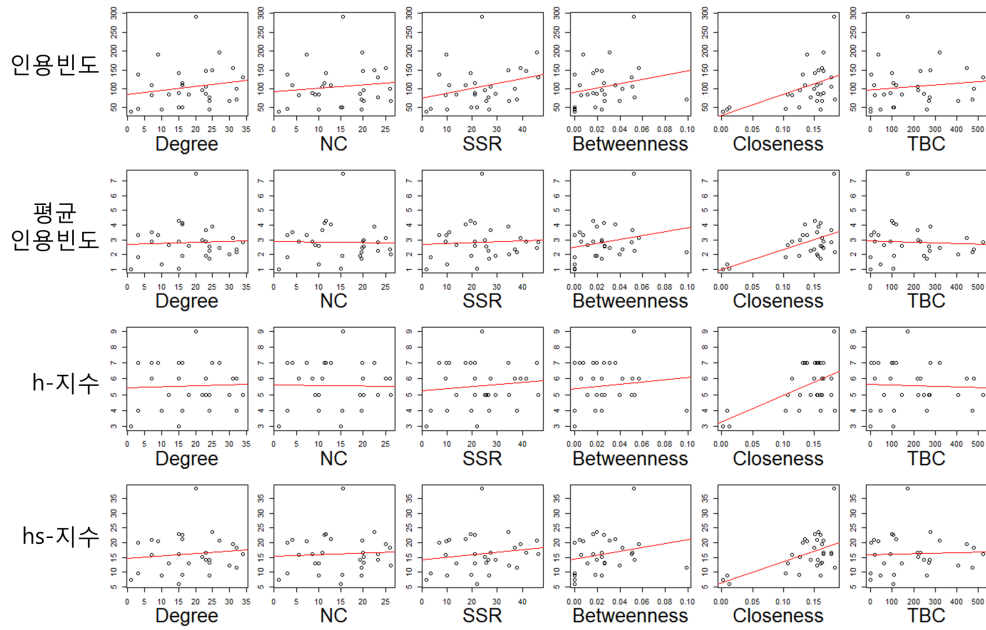
	인용빈도	평균 인용빈도	h-지수	hs-지수	보정 인용빈도	보정 평균 인용빈도	보정 h-지수	보정 hs-지수
연결정도중심성	.207	-.027	-.009	.089	-.142	-.015	-.352	-.347
SSR	.297	-.014	.024	.118	-.052	-.002	-.291	-.286
이웃중심성	.166	-.048	-.027	.074	-.154	-.037	-.384*	-.372*
매개중심성	.307	.221	.171	.237	.011	.220	-.139	-.150
근접중심성	.496**	.380*	.352	.397*	.176	.371*	.027	.021
삼각매개중심성	.206	-.040	-.015	.085	-.137	-.027	-.358	-.346

** . 상관관계가 0.01 수준에서 유의함(양측).

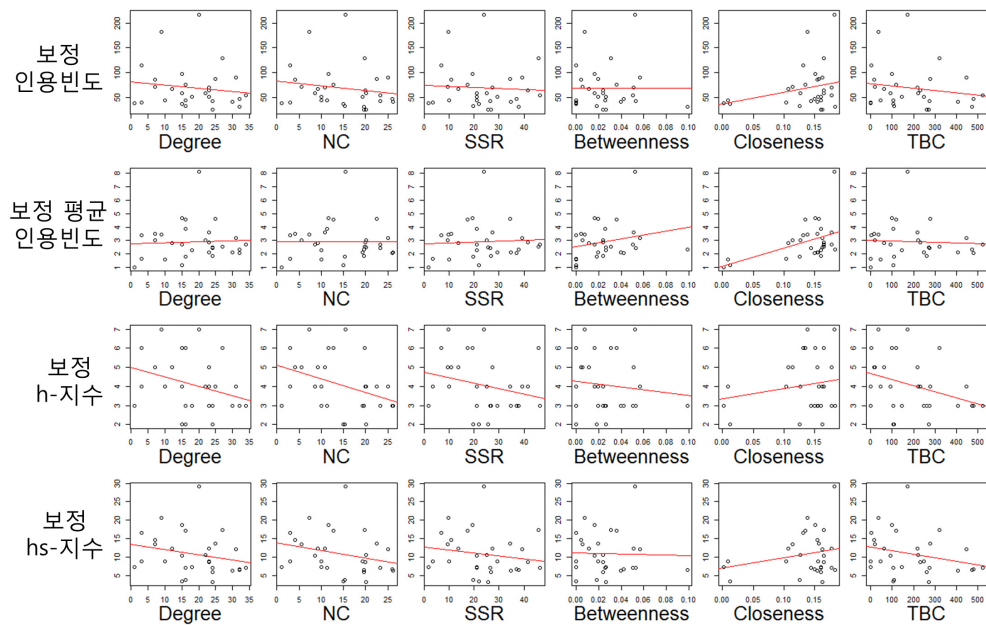
* . 상관관계가 0.05 수준에서 유의함(양측).

적으로 유의한 긍정적 상관관계가 나타났다. 앞서 논문 수와 중심성 지수 사이의 상관관계에서도 동일한 결과가 확인된 바 있다. 또한 보정한 인용지수 중에서도 보정 인용빈도와 보정 평균인용빈도는 두 집합에서 모두 6가지 중심성 지수와 통계적으로 긍정적인 상관관계를 보였다. 보정 h-지수와 보정 hs-지수도 근접중심성을 제외한 5가지 중심성 지수와 통계적으로는 상관관계가 유의하였지만 상관계수 수치는 다소 저하되었다. 전체적으로 근접중심성을 제외한 모든 중심성 지수가 공저 보정 여부에 상관없이 모든 인용지수와 통계적으로 유의한 상관관계를 보였다.

〈표 16〉의 논문 수 최상위 저자 30명에 대한 분석 결과에서는 이웃중심성과 근접중심성을 제외한 나머지 4종 중심성 지수가 인용지수 8종 모두와 통계적으로 유의한 상관관계를 보이지 않았다. 근접중심성의 경우, 인용빈도, 평균 인용빈도, hs-지수, 보정 평균인용빈도와 상관관계가 유의하게 나타났다. 그러나 〈그림 8〉과 〈그림 9〉의 산점도를 살펴보면 예외적으로 근접중심성이 낮은 3인이 상관관계를 크게 좌우하고 있는 것으로 보인다. 이들은 공저 네트워크의 최대컴포넌트에 속하지 않아서 대부분의 다른 저자들과 연결되지 않기 때문에 근접중심성이 현저하게 낮은 저자들이다. 이들 3인을 제



〈그림 8〉 논문 수 상위 저자 30명에 대한 인용지수 4종(세로축)과 중심성 지수 6종(가로축) 사이의 행렬 산점도(실선은 회귀선)



〈그림 9〉 논문 수 상위 저자 30명에 대한 보정 인용지수 4종(세로축)과 중심성 지수 6종(가로축) 사이의 행렬 산점도(실선은 회귀선)

외하고 스피어맨 상관계수를 산출해보면 근접중심성은 모든 인용지수와 통계적인 상관성이 유의하지 않은 것으로 나타난다. 이밖에 이웃중심성이 보정 h-지수 및 보정 h_s -지수와 통계적으로 유의한 부정적 상관관계를 보이는 것으로 나타났지만($\rho = -0.384, -0.372$), 상관계수의 절대값은 삼각매개중심성의 경우($\rho = -0.358, -0.346$)나 연결정도중심성의 경우($\rho = -0.352, -0.347$)와 큰 차이는 없었다.

전체적으로 논문수 5건 이상 저자 276명을 대상으로 한 분석에서는 근접중심성을 제외한 모든 중심성 지수와 모든 인용지수 사이에 통계적으로 유의한 상관관계가 나타난 반면에, 논문수 최상위 저자 30명에 대한 분석에서는 대부분의 인용지수와 중심성 지수 사이에 통계적으로 유의한 상관관계가 나타나지 않았다. 오히려 이웃중심성의 경우는 공저 보정 h-지수 및 보정 h_s -지수와 통계적으로 유의한 부정적 상관관계를 보였다. 이런 차이는 전체 저자를 대상으로 분석한 Abbasi, Altmann, Hossain(2011)의 연구에서 중심성과 인용지수 사이에 상관관계가 유의한 반면에, 상위 30명 저자를 대상으로 분석한 이수상(2013)의 연구에서는 연결정도중심성 및 매개중심성과 인용지수 사이에 상관관계가 유의하지 않게 나타났던 것을 설명해주는 결과이다. 근접중심성은 공저 네트워크의 최대 컴포넌트에서 분리될 경우에 극단적으로 낮은 값을 가지게 되므로 설명력이 떨어지는 점을 감안하여 논의에서 제외하는 것이 바람직하다. 결국 연결정도중심성과 매개중심성의 경우에 생산성이나 중심성이 낮은 하위 저자까지 포함하여 분석하면 인용지수와 통계적으로 유의한 상관성을 보이지만, 상위 저자만을 대상으로 분

석할 경우에는 인용지수와의 상관성이 통계적으로 유의하지 않다는 것이 확인되었다. 이런 경향은 가중 네트워크 중심성인 SSR과 삼각매개 중심성의 경우도 마찬가지로 나타났다.

5. 결론

2004년부터 10년 동안 문헌정보학 분야 KCI 등재지 9종 학술지에 발표된 논문 3,661편의 저자 1,638명을 대상으로 공저 관련 변수와 연구성과 변수를 측정하고 상관관계를 분석해보았다. 인용빈도 등의 연구성과 변수는 공저자 수를 고려하여 보정한 변수도 함께 측정하였다. 또한 모든 분석은 전체 1,638명 대상, 논문 5건 이상 발표한 236명 대상, 논문 수 상위 30명 대상의 세 가지 경우에 대해서 각각 실시하였다. 분석에서 얻어진 결과는 다음과 같다.

첫째, 논문별 저자 수에 따른 논문 인용빈도의 차이는 없는 것으로 나타났다. 이는 공저자가 많은 논문이 인용빈도도 높은 경향이 있다는 해외 선행 연구와 다른 결과이다.

둘째, 각 저자별로 발표한 논문 수와 인용지수 사이의 상관관계를 살펴본 결과, 공저 보정 여부와 상관없이 전체 1,638명과 논문 5건 이상 276명을 대상으로 분석하였을 때 논문 수와 각 인용지수 사이의 상관관계가 모두 통계적으로 유의한 것으로 나타났다. 그러나 논문 수 상위 저자 30명만을 대상으로 분석한 경우에는 공저를 보정하지 않을 경우 인용빈도만 논문 수와 상관관계가 유의하게 나타났고, 공저를 보정한 경우에는 h-지수와 h_s -지수도 논문 수와 상관관계가 유의하게 나타났으며, 평균인용빈도는

공저 보정 여부에 상관없이 일관되게 논문 수와 상관관계가 유의하지 않았다. 상위 저자 집단에 대해서 분석했을 때 공저 보정 여부에 따라서 h-지수와 hs-지수가 논문 수와 가지는 상관관계가 달라진 것이다. 본질적으로 연구성과의 양적인 측면과 질적인 측면을 복합적으로 반영하는 지수이므로 이론적으로는 논문 수와 통계적인 상관성이 있어야 함을 감안하면, 공저를 고려하여 논문 수와 인용지수를 측정하는 것이 현상을 더 잘 설명할 수 있는 방식이라고 할 수 있다.

셋째, 협력계수와 연구성과 사이의 상관관계를 살펴본 결과, 공저를 고려한 보정 논문 수와 보정 h-지수, 보정 hs-지수는 공저가 활발한 저자일수록 오히려 더 적은 것으로 나타났다. 즉, 공저를 고려할 경우 오히려 단독 저술 위주의 저자가 공저 위주의 저자보다 양적인 연구성과와 복합 연구성과가 더 높다는 것을 의미한다. 질적인 연구성과 중에서 공저를 보정한 인용빈도도 협력계수와 부정적인 상관관계를 보였으며 상위 저자집단에 대해서만 상관관계가 통계적으로 유의하지 않은 것으로 나타났다. 이로써 공저가 활발하다고 해서 연구성과가 더 좋은 것은 아니라고 결론내릴 수 있다.

넷째, 중심성 지수와 연구성과의 양적/질적인 측면 사이의 상관관계를 분석한 결과 전체적으로 선행 연구와 마찬가지로 중심성 지수가 높은 저자가 논문 수도 많고 인용지수도 높은 것으로 나타났다. 이는 전체 저자를 대상으로 분석한 Abbasi, Altmann, Hossain(2011)의 연구와 임병학(2013)의 연구에서와 일치하는 결과이다. 그러나 논문 수 상위 저자 집단의 경우에 공저를 보정한 논문 수는 중심성 지수가

높을수록 오히려 적은 경향이 나타났다. 또한 질적인 연구성과인 인용지수와 중심성 지수 사이의 관계는 전체적으로는 선행 연구(Abbasi, Altmann, & Hossain, 2011; Li, Liao, & Yen, 2013)와 마찬가지로 긍정적인 상관관계가 나타났으나, 상위 저자 30명만을 대상으로 분석한 결과에서는 상관관계가 유의하지 않은 것으로 나타났다. 이 경우는 중심성 상위 저자 집단을 대상으로 분석한 이수상(2013)의 연구와 일치하는 결과이다. 이로써 인용빈도와 중심성 지수와의 상관관계를 분석한 선행 연구들에서 나타난 결과가 불일치한 이유가 분석 대상 저자 집단의 범위가 달랐기 때문임이 입증되었다.

이처럼 저자 집단의 범위를 달리한 분석과 공저를 보정한 분석을 통해서 선행 연구의 결과를 재확인하고 상반된 것처럼 보였던 결과들을 설명할 수 있었다.

연구 결과 중에서 후속 연구가 특히 필요한 부분은 논문의 저자 수와 인용빈도 사이의 관계이다. 해외 타 분야와 달리 국내 문헌정보학계에서는 공저자가 많은 논문이 인용빈도도 높게 나타나지는 않았다. 해외의 타 분야에 대한 분석과 다른 원인으로서는 국내와 국외의 논문 소비 집단의 차이, 혹은 국내와 국외의 논문 생산 집단의 차이를 생각해볼 수 있다. 우선 논문을 소비하는 측의 차이에 주목하는 추측으로는, 국내 문헌정보학 분야의 연구자 집단 규모가 선행 연구에서 다루어진 분야의 해외 연구자 집단 규모보다 매우 작기 때문에 공저자 수에 따라 의미있는 차이를 낳을 정도의 인용이 충분히 이루어지지 못했을 수 있는 가능성이 있다. 이와 달리 논문을 생산하는 측의 차이가 원인이라고 추론해보면, 국내 문헌정보학 분야의 공저

연구가 단독 연구에 비해서 여러 명의 저자가 참여하여 기여한 만큼의 영향력을 실제로 가지지 못했을 가능성이 있다. 만약 국내 다른 학문분야에서도 단독 논문과 공저 논문의 인용 빈도에 차이가 없다면 첫째 추측처럼 논문 소비집단의

차이에 기인했을 가능성이 더 크다고 할 수 있다. 만약 그렇지 않다면 국내 문헌정보학 분야의 공동연구가 가지는 가치에 대해서 더 깊은 탐구와 고민이 있어야 할 것이다.

참 고 문 헌

- 김은주, 남태우 (2015). 정신의학 분야 국제공동연구의 지식구조 네트워크에 관한 연구. *정보관리학회지*, 32(3), 317-340. <http://dx.doi.org/10.3743/KOSIM.2015.32.3.317>
- 박치성 (2012). 행정학 학문공동체의 공동연구 네트워크 구조에 관한 연구: 1998-2009년간 24개 행정학 학술지의 논문공저자 네트워크의 분석을 중심으로. *한국사회와 행정연구*, 22(4), 129-153.
- 이수상 (2013). 연구자 네트워크의 중심성과 연구성과의 연관성 분석: 국내 기록관리학 분야 학술논문을 중심으로. *한국도서관·정보학회지*, 44(3), 405-428.
- 이재윤 (2006a). 계량서지적 네트워크 분석을 위한 중심성 척도에 관한 연구. *한국문헌정보학회지*, 40(3), 191-214. <http://dx.doi.org/10.4275/kslis.2006.40.3.191>
- 이재윤 (2006b). 연구성과 측정을 위한 h-지수의 개량에 관한 연구. *정보관리학회지*, 23(3), 167-186. <http://dx.doi.org/10.3743/KOSIM.2006.23.3.167>
- 이재윤 (2014). 공동연구 네트워크 분석을 위한 중심성 지수에 대한 비교 연구. *정보관리학회지*, 31(3), 153-179. <http://dx.doi.org/10.3743/KOSIM.2014.31.3.153>
- 이재윤 (2015). 가중 네트워크를 위한 일반화된 지역중심성 지수. *정보관리학회지*, 32(2), 7-23. <http://dx.doi.org/10.3743/KOSIM.2015.32.2.007>
- 이재윤 (2016). 공저자 수를 고려한 h-지수 산출. *정보관리학회지*, 33(3), 7-29. <http://dx.doi.org/10.3743/2016.33.3.007>
- 임병학 (2013). 논문 공동저자 네트워크가 연구 성과에 미치는 영향에 대한 연구: 로고스경영연구의 공동저자를 중심으로. *로고스경영연구*, 10(1), 1-20.
- Abbasi, A., Altmann, J., & Hossain, L. (2011). Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics*, 5(4), 594-607. <http://dx.doi.org/10.1016/j.joi.2011.05.007>
- Ajiferuke, I., Burell, Q., & Tague, J. (1988). Collaborative coefficient: A single measure of the

- degree of collaboration in research. *Scientometrics*, 14(5), 421-433.
<http://dx.doi.org/10.1007/BF02017100>
- Bordons, M., Aparicio, J., González-Albo, B., & Díaz-Faes, A. A. (2015). The relationship between the research performance of scientists and their position in co-authorship networks in three fields. *Journal of Informetrics*, 9(1), 135-144. <http://dx.doi.org/10.1016/j.joi.2014.12.001>
- Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics*, 69, 131-152.
<http://dx.doi.org/10.1007/s11192-006-0144-7>
- Franceschet, M., & Costantini, A. (2010). The effect of scholar collaboration on impact and quality of academic papers. *Journal of Informetrics*, 4(4), 540-553.
- Gazni, A., & Didegah, F. (2011). Investigating different types of research collaboration and citation impact: A case study of Harvard University's publications. *Scientometrics*, 87(2), 251-265. <http://dx.doi.org/10.1007/s11192-011-0343-8>
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), 16569-16572.
<http://dx.doi.org/10.1073/pnas.0507655102>
- Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26(1), 1-18.
- Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology*, 66(7), 1323-1332. <http://dx.doi.org/10.1002/asi.23266>
- Lee, Jae Yun, & Chung, EunKyong (2014). A comparative analysis on multiple authorship counting for author co-citation analysis. *Journal of the Korean Society for Information Management*, 31(2), 57-77. <http://dx.doi.org/10.3743/KOSIM.2014.31.2.057>
- Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*, 35(5), 673-702. <http://dx.doi.org/10.1177/0306312705052359>
- Li, E. Y., Liao, C. H., & Yen, H. R. (2013). Co-authorship networks and research impact: A social capital perspective. *Research Policy*, 42(9), 1515-1530.
<http://dx.doi.org/10.1016/j.respol.2013.06.012>

• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

- Kim, Eun-Ju, & Nam, Tae-Woo (2015). A study on the knowledge structure networks of international collaboration in psychiatry. *Journal of the Korean Society for Information*

- Management, 32(3), 317-340. <http://dx.doi.org/10.3743/KOSIM.2015.32.3.317>
- Lee, Jae Yun (2006a). Centrality measures for bibliometric network analysis. *Journal of the Korean Society for Library and Information*, 40(3), 191-214.
<http://dx.doi.org/10.4275/KSLIS.2006.40.3.191>
- Lee, Jae Yun (2006b). Some improvements on h-index: Measuring research outputs by citations. *Journal of the Korean Society for Information Management*, 23(3), 167-186.
<http://dx.doi.org/10.3743/KOSIM.2006.23.3.167>
- Lee, Jae Yun (2014). A comparative study on the centrality measures for analyzing research collaboration networks. *Journal of the Korean Society for Information Management*, 31(3), 153-179. <http://dx.doi.org/10.3743/KOSIM.2014.31.3.153>
- Lee, Jae Yun (2015). A generalized measure for local centralities in weighted networks. *Journal of the Korean Society for Information Management*, 32(2), 7-23.
<http://dx.doi.org/10.3743/KOSIM.2015.32.2.007>
- Lee, Jae Yun (2016). Calculating the h-index and its variants considering the number of authors in a paper. *Journal of the Korean Society for Information Management*, 33(3), 7-29.
<http://dx.doi.org/10.3743/2016.33.3.007>
- Lee, Soo-Sang (2013). Analytical study on the relationship between centralities of research networks and research performances. *Journal of Korean Library and Information Science Society*, 44(3), 405-428.
- Leem, Byung-Hak (2013). An effect of co-authorship network on research performance: Focusing on co-authoring of *Logos Management Review*. *Logos Management Review*, 10(1), 1-20.
- Park, Chisung (2012). A study on the network structure of the public administration academic community using the coauthor network from 1998 to 2009. *Korean Society and Public Administration*, 22(4), 129-153.

