

국가연구개발사업 계량평가 도입 효용성에 대한 실증적 분석

- 세계수준의 연구중심대학(WCU)육성사업을 중심으로 -

Usefulness Analysis of Developed Bibliometric Indicators for Evaluation of National R&D Programs

- Case Study of World Class University Program -

조우현(Woo-hyun Cho)*, 우정표(Jung-Pyo Woo)**, 송충한(Chung-Han Song)***

I. 서론

1. 연구개요

지난해 발표된 2009년 기준 The Times가 선정한 세계 우수대학 200위 이내에 국내 4개 대학¹⁾이 포함된 것으로 발표되었다. 이는 전년²⁾보다 향상된 결과이나 일본, 홍콩, 중국, 싱가포르 등 아시아 주요국과 비교하면 아직 미흡한 수준이다. 전 세계적으로 우수 대학들은 글로벌 경쟁력을 갖춘 교수진을 확보하고자 노력하고 있으며 이를 위해 정부차원의 다양한 프로그램이 활발히 운영되고 있다. 우리나라도 2008년 해외 우수학자의 대규모 유치를 통한 대학경쟁력 강화를 목표로 '세계수준의 연구중심대학(WCU) 육성사업'을 신규 도입하였다.

WCU사업은 우수한 해외 연구진의 유치라는 기존의 국내 연구사업과는 차별화된 새로운 요소가 주요사항으로 부각되게 됨은 물론, 정부차원의 예산지원이 5년간 약 8,250억원에 달하는 대형 국책사업으로 출범함에 따라, 그 어떤 연구사업보다 철저하고 객관적인 평가체계가 강하게 요구되었다. 이런 배경으로 WCU사업은 Peer Review 체계에 의존하던 기존의 국가연구개발사업과는 다른 새로운 과학적이고 객관적으로 분석할 수 있는 성과지표의 도입이 추진되었다.

연구자의 우수성을 나타내는 연구성과의 산출물은 다양한 형태로 나타날 수 있으나, 전 세계적으로 공통으로 적용할 수 있는 성과물은 논문과 특허, 편집위원참여실적 등으로 분류할 수 있다. 현재 국가R&D사업 평가에서는 통상 SCI 논문관련 지표(논문수, 피인용수(Citation), 영향력지수(Impact Factor) 등)를 제한적으로 적용하고 있다. 그러나 SCI논문 성과지표는 질적 성과보다는 양적 성과를 측정하는 한계를 가지고 있으며, 연구분야별 특성을 정확히 반영하지 못하는 단점도 극복하지 못하고 있다(Kostoff, 1995). 이후 과학계량학(Scientometrics)³⁾은 H-index(Hirsch, 2005), G-index(Egghe, 2006) 등의 다양한 형태로 발전하였으나, 실효성에 대한 논란은 여전히 지속되고 있다. 따라서 국가연구개발사업의 효과적이고 객관적인 평가시스템을 구축하기 위해 기존 지표의 한계를 극복하고 질적인 평가를 수행할 수 있는 과학적인 지표의 수립이 필요하고 실제 연구사업 평가에 적용함으로써 그 효용성을 검증할 필요가 있다.

* 조우현, 한국연구재단 연구원, whcho@nrf.go.kr, 042-869-6251

** 우정표, 한국연구재단 WCU육성팀장, jpwoo@nrf.go.kr, 042-869-6250

*** 송충한, 한국연구재단 R&D혁신센터장, chsong@nrf.go.kr, 042-869-6640

1) The Times는 미국 QS사와 매년 대학순위를 발표하고 있다. 2009년 대학순위 200위 내에 서울대, KAIST, POSTECH, 연세대 등 4개 대학이 포함되었다.

2) 2008년 기준 상위 200위권에는 서울대, KAIST, POSTECH 등 3개 대학만 포함되었다.

3) 과학(Science)와 계량(Metrics)이라는 뜻으로, 학자의 연구성과 계량적 측정 학문을 말한다.

이에 본 연구에서는 WCU사업 기획당시 철저한 검토를 통해 신규 채용된 계량평가지표로 상위 10% 저널 게재 논문수⁴⁾, 인당 피인용횟수, 편당 피인용횟수, 편집위원참여실적, 특허등록 실적(모든 지표는 각 연구분야내의 상대 우위를 기준으로 측정) 등을 적용한 실제 평가운영 결과를 분석하여 신규 설계된 연구실적 계량평가지표의 효용성을 검토하여, 국가연구개발사업에 계량지표의 적극적인 도입에 대한 가능성과 향후 발전방안 등을 제시하고자 한다.

2. 주요선행연구

우리나라뿐 아니라 전 세계 선진국의 국가 연구개발사업 평가를 비롯한 각종 평가에 동료평가(Peer Review)가 활용되고 있다. 동료평가는 연구나 연구제안서를 독립된 전문연구자들에 의해 철저히 검토되는 과정이다(Harding, 2002). 다시말하면, 평가대상이 되는 연구과제 또는 연구인력을 해당분야의 전문성을 갖춘 동료 연구자들이 연구의 우수성 및 연구방법의 체계성 등을 주관적으로 측정하는 방법을 의미한다.

동료평가는 오랜기간동안 가장 이상적인 연구개발사업 평가 체계(COSEPUP, 1999)로 인식되고 있으며, 실제로 전 세계적으로 가장 효과적인 방법으로 활용되고 있다. 동료평가가 객관적으로 정량화된 산출결과를 제공할 수는 없지만 연구의 질적인 측면을 평가하는 방법으로서 연구의 영향 평가에 가장 광범위 하게 사용되고 있으며 신뢰할 수 있는 수단으로 인정받고 있다(연경남 외, 2005). 전 세계 대학을 평가하여 순위를 매년 제공하는 The Times의 대학평가에서도 Peer Review의 활용도는 40%에 달한다.

그러나 동료평가는 전문가들에 의한 주관적인 판단에 의존하므로 가장 효과적인 연구과제의 질적평가 체계임에 틀림없으나 완벽한 평가 방법은 아니며, 개인의 이해관계나 주변 환경 등에 영향을 받을 수 있는 개연성이 높다. Harding(2002)은 동료평가의 평가시간(기간)이 짧기 때문에 정확한 평가가 어렵다는 점과 적정 평가위원의 섭외 승인률이 일반적으로 50% 미만이라는 문제 등을 동료평가의 비효율성으로 지적하였다. Kostoff(1995)는 동료평가의 문제점을 다음 6가지로 정리하였다. 첫째, 평가자가 해당 분야 전공과 일치하지 않을 경우 평가 결과가 왜곡될 수 있다. 둘째, 기존의 연구분야를 보호하기 위한 중견 학자들의 네트워크(Old boy network)가 새롭게 등장하는 학문분야의 연구비 수혜를 제한하는 결과로 평가를 왜곡할 수 있다. 셋째, 스타 과학자 등에 의한 후광효과(Halo Effect)가 작용할 수 있어 연구내용의 질적 수준과 무관하게 연구비를 지원받을 가능성이 높다. 넷째, 개별 평가자 마다 각기 다른 해석과 평가의 기준이 적용될 수 있다. 다섯째, 동료평가의 전제조건으로 질적으로 우수한 연구과제를 구성하는 요소가 무엇이고 이를 이끄는 과학자가 누구인지, 향후 연구 방향이 어느 쪽으로 진행될 것인지에 대해 평가자간 상당 수준의 동의와 의견일치를 요구한다는 점이다. 여섯째, 많은 비용이 소요되는 문제가 발생할 수 있다(연경남 외, 2005).

이러한 동료평가의 문제점을 개선하고자 객관적인 수치에 의한 평가제도의 도입에 대한 연구가 활발하게 이루어지고 있다. 연구실적에 대한 계량적인 평가는 1900년대 이전부터 연구가 시작되었으나(Cole and Eales, 1917) 1969년 “Bibliometric”의 개념이 개발되면서 계량지표 개발과 연구가 활성화 되었다(Prichard, 1969).

계량지표중 기본적으로 활용되고 있는 논문 피인용수를 활용하여 Hurt(1985)는 방법론적 피인용수의 차이와 개념적 피인용수의 차이를 분석하였으며, 피인용 빈도 순위를 우수 학자의 구분에 활용하였다. Velho와 Krige(1984)는 전통적인 접근 방법으로 연령별, 국적별 피인용수 분석을 추진하였다. 이렇듯 1980년대 이후 논문의 인용지수에 대한 다양한 연구가 진행되었으며, 이러한 과도기를 거쳐 2000년대에 들어서면서 SCI 성과지표에 대한 분석적인 연구가 활발히 진행되게 된다.

4) WCU사업은 기존의 연구사업과 다르게 SCI와 SCOPUS를 동시 적용하였다. 최근 The Times 선정 대학평가 등에서 SCOPUS가 활용되는 등 SCOPUS의 활용 비중이 전세계적으로 확대되는 것을 반영한 조치이다.

SCI의 논문 Index를 활용한 계량지표의 연구로 이한진(2004)은 Leeuwen, Wurff and Raan(2001), Aksnes and Taxt(2004) 등이 개발한 논문 피인용수를 활용한 논문수, 전체 피인용 횟수, 인용도가 없는 논문 비율, 논문 평균 피인용 횟수 등 다양한 형태의 지표를 소개하고 이를 분석하였다. 연경남 외(2005)는 Peer Review와 계량지표의 적용 비율 및 계량지표를 위한 SCI 논문의 적용범위, Impact Factor의 분야별 백분율 활용 등의 국가연구개발사업에의 계량지표 적용방안을 분석 및 제시하였다. 허정은 외(2006)는 NSC(Nature, Science, Cell) 저널에 게재된 논문현황 분석을 계량서지학적으로 추진하였고, Sen(1992)은 분야별 IF의 최대값을 지표로 제안했으며, Marshakova-Shaikovich(1996)는 분야별 상위 5%에 해당하는 IF값을 바탕으로한 새로운 지표를, Pudovkin(2004)은 분야내 순위만을 활용하는 순위보정 영향력 지수를 제안하였다. 허정은 외(2008)는 표준화된 순위보정 영향력지수(Modified Rank Normalized Impact Factor, mrnIF), 보완된 순위보정 영향력지수(Ordinal Rank Normalized Impact Factor, ornIF), 분야대비 영향력지수(Relative Field Impact Factor, rfiF), 기여율을 고려한 논문수 등을 제안하고 이에 대한 타당성을 분석하였다.

새로운 계량지표의 개발과 적용 타당성 분석 등이 활발히 전개되고 있으나 계량지표에 대한 비판역시 제기되고 있다. 정보과학 등 일부 분야에서는 세계 최고 수준의 학술대회 논문집이 SCI에 포함되지 않는다는 지적이 제기되었다(김원, 2005). 정진완 외(2005)는 인용횟수에 대한 문제점으로 7가지를 지적하였다. 첫째, 우수한 학술대회 논문의 배제, 둘째, 인용횟수는 해당분야 학자 수에 영향을 받는다는 문제, 셋째, 해당 분야 학자들의 인용 성향이 영향을 미친다는 문제, 넷째, 해당 논문이 Review논문인 경우 인용수가 상대적으로 커지는 문제, 다섯째, 해당저자나 공동연구자의 자기인용 문제, 여섯째, 비판을 위한 인용 문제⁵⁾(인용 동기 문제), 일곱째, 인용 횟수를 추출하는 연도의 문제⁶⁾가 그것이다.

이러한 계량평가의 문제점에도 불구하고 계량평가를 국가연구개발사업 평가에 적용하는 시도는 다양하게 진행되고 있다. 한국과학재단은 1995년부터 2년간 핵심전문연구지원사업에 연구자의 최근 5년간 SCI실적을 계량화하여 반영하였고, 2001년부터 4년간 선도과학자 육성지원사업 선정평가에 분야별 SCI의 영향력지수(IF) 표준을 제시하고 이를 상회하는 논문의 피인용 횟수를 점수화하여 50% 적용한바 있으며, 2005년 기초의과학연구사업 선정평가에 대표논문의 10편의 순위보정 IF값과 전체 논문의 IF를 각각 15% 적용하였다. 프론티어연구개발사업에는 2006년부터 2년간 순위보정 IF와 인용횟수/IF를 단계평가에 5% 적용한 사례가 있다. 한국학술진흥재단은 2005년부터 2년간 국가석학지원사업에 논문인용지표를 일정부분 충족하지 못하면 신청자격에 부여하지 않는 제도를 시행하였고, BK21사업의 경우 2007년 단계평가에서 논문실적, 취업률 등을 실적순위에 따른 계량평가를 수행하였으나, 2008년 정성적 방법으로 전환하였다.

여러 연구와 사례를 통해 살펴본 바와 같이 Peer Review와 계량지표 적용 평가 모두 장점과 문제점을 가지고 있다. 이 두 가지 평가방식의 효과적 복수 적용은 연구사업 평가의 효율성을 높이는데 기여할 수 있으나 그 실증적인 분석은 찾아보기 어렵다. 따라서 본 연구에서는 Peer Review와 신규 제안된 계량지표를 약 50%씩 적용한 최초의 국가연구개발사업인 세계수준의 연구중심대학(WCU) 육성사업의 평가점수 분석을 통해 Peer Review와 계량지표의 적용 사례 효용성을 분석하고 발전방안을 제시하고자 한다.

5) 논문의 오류를 비판하기 위해 인용된 논문도 인용수는 증가하게 된다.

6) 시간이 지나가야 논문의 인용횟수가 증가하게 된다.

II. 본문

1. 연구방법 및 설계

연구실적 계량평가는 그동안 국가연구개발사업의 평가, 대학의 교원평가 등에 제한적으로 활용되어왔다. 그러나 적용 방법은 SCI 논문수, 인용수, 특허등록 건수 등 단순 양적인 측정을 넘어서지 못하고 있다. 2장의 연구 사례에서도 계량지표에 대한 다양한 연구가 진행되었으나 SCI 논문지표의 범위를 벗어나지는 못하고 있다.

본 연구에서는 2008년 신규 국가 연구지원사업으로 추진된 세계수준의 연구중심대학(WCU) 육성사업에서 활용된 계량지표를 중심으로 그 효용성을 조사하였다. WCU사업은 미래 국가발전 핵심분야의 연구추진 및 학문후속세대 양성을 목표로 우수 해외학자 유치를 통해 국내 대학의 교육·연구풍토를 혁신하고 세계수준의 연구중심대학을 육성하는 목표로 추진되었다. 5년간 총 8,250억원이 투입되는 대규모 사업으로 3가지 유형으로 구분 운영된다. 1유형은 다수의 해외학자와 국내학자가 새로운 학과 또는 전공을 신설하는 과제로 과제당 평균 약 30억원이 투입되며, 2유형은 개별학과 초빙지원과제로 기존 학과에 해외학자를 초빙하여 공동연구 및 강의를 진행하며, 과제당 평균 약 7억원이 투입된다. 3유형은 세계수준의 해외석학을 단기간 초빙하여 특강, 세미나 등 교육활동 및 공동연구를 수행하는 형태로 추진된다.

WCU사업에서 도입된 계량지표는 기존에 활용되던 양적 실적지표보다 크게 개선되어 질적으로 점수가 차등 적용되는 새로운 계량평가 지표를 개발하여 적용하였다. 질 중심의 처음 도입된 계량평가 지표 구성은 아래와 같다.

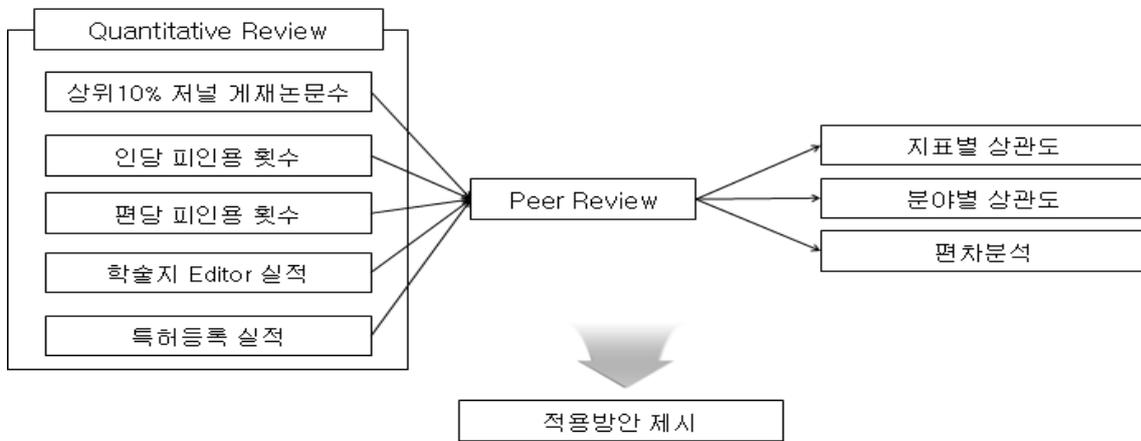
<표 1> 세계수준의 연구중심대학(WCU) 육성사업 계량평가 지표

WCU사업 계량지표
1. 학문분야별 상위 10% 저널 게재 논문수 (SCI급, SCOPUS급 게재 논문수)
2. 1인당 피인용 횟수
3. 논문 1건당 피인용 횟수
4. SCI급 및 SCOPUS급 학술지의 Editor 경력
5. 국내외 특허등록 실적 (공학중심 과제에 한함)
6. 연구비 수주실적

각 지표의 내용은 다음과 같다. 첫째, 상위 10% 저널 게재 논문수는 SCI와 SCOPUS DB에서 세부분야별 영향력 지수(IF) 상위 10%에 해당하는 저널 리스트를 추출하여 각 사업팀의 해당 저널 게재논문 수를 논문저자 구분에 따라 차등점수화 하였다. 이는 기존의 단순 논문수를 평가하는 방식에서 획기적으로 개선한 방법으로 영향력 지수라는 계량 수치를 이용한 우수저널에 게재된 논문에 점수를 부여하는 질적 평가지표의 특징을 갖는다. 둘째, 1인당 피인용 횟수는 논문저자 구분에 따라 차등 점수화된 국내외 참여교수의 피인용 총횟수의 총합을 참여교수 수로 나눈 수치로 피인용 횟수를 평균적으로 적용한 질적 평가지표의 특징을 갖는다. 셋째, 논문 1건당 피인용횟수는 역시 논문저자 구분에 따라 참여교수별 논문 1건당 피인용 횟수를 산출하고 이를 참여교수의 수로 나눈 지표로 피인용 횟수가 높은 우수 논문을 게재한 학자에게 높은 배점이 돌아가는 질 중심의 지표 특징을 갖는다. 넷째, 학술지 Editor 경력은 SCI, SCOPUS급 저널의 편집위원 참여실적

을 개월수로 카운팅한 실적으로, 각 학자의 학계의 지명도를 점수화 하는 지표로, 논문 이외의 새로운 형태의 연구자 수준평가 지표로 도입되었다. 다섯째, 국내외 특허등록 실적은 공학단위 과제에만 해당되는 지표로 등록된 특허에 한해 발명인수 기준에 따라 점수화하는 지표로 도입되었다. 여섯째, 연구비수주실적은 참여학자들의 타 연구 수행 실적 및 역량을 판단할 수 있는 지표로 도입되었다. 그러나 연구비의 규모나 국가별 판단 기준 등이 다소 모호한 단점으로 인해 객관성이 미흡할 개연성이 있으므로 동 분석에서 제외하였다.

WCU 육성사업은 국가연구개발사업에서는 최초로 선정평가에 정성평가와 계량평가 비중을 각각 50%로 적용한 사례이다. 계량평가에 대한 비판이 꾸준히 제기되는 가운데 대안으로 개발된 계량지표가 그동안 국가연구개발사업 평가에 있어서 세계적으로 적용되고 있는 Peer Review와의 상관도 분석을 통해 적용의 가능성을 정확히 분석할 수 있는 의의를 갖는다. 따라서 본 연구에서는 상기 사업을 중심으로 다양한 형태의 계량지표라는 투입변수가 Peer Review 결과와 어떠한 연관성을 갖는지 고찰해 보고, 평가를 실제 수감한 전문가들로부터의 의견을 수렴하는 과정을 통해 계량평가 적용의 효용성 및 적용방안을 도출하고자 한다.

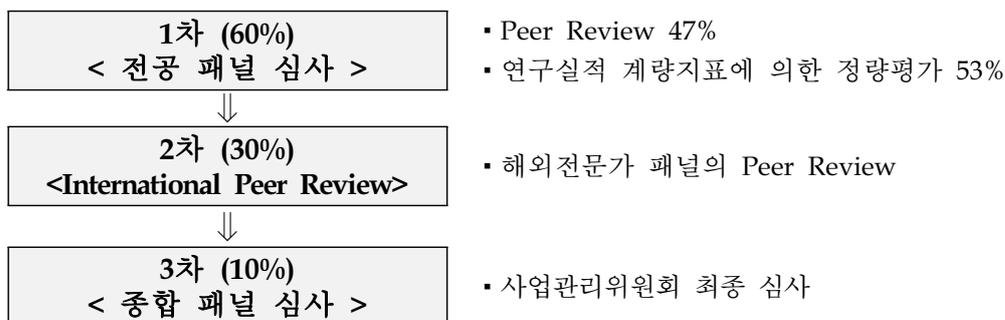


(그림 1) 연구설계도

2. 계량평가 실증분석

1) WCU사업 평가체계

WCU사업(유형1, 2)은 기존의 국가연구개발사업과는 다소 차별화된 평가제도를 도입하여 운영하였다. WCU사업의 3단계 평가체계는 아래 그림과 같다.



(그림 2) WCU사업 평가체계도

WCU사업 평가체계의 가장 큰 특징은 1차 전공패널심사의 계량지표에 의한 평가가 53%를 차지한다는 점이다. 또한 WCU사업 평가에서 활용된 계량지표는 기존의 양 중심의 지표를 개선한 질 중심의 평가지표가 대거 도입된 사례로 볼 수 있다. 유형1사업 1차심사 평가지표별 세부 배점 기준은 아래 표와 같다.

<표 2> WCU사업 유형1 선정평가 1차심사 지표구성

평가지표	배점			
· 유치·활용 계획서 평가 : 45% - 신설 전공·학과 분야의 타당성 - 전공·학과 구성 인력의 적합성 및 우수성(해외유치인력 포함) - 교육과정 구성 및 운영 계획의 적합성 - 참여교수별 매학기당 담당 교과목 배분 및 학생지도계획의 적절성 - 연구계획서(Research Proposal)의 우수성 - 연구 및 교육의 학문적·경제사회발전 기여도				
	자연과학	생명과학	공학	인문사회
· 연구팀 1인당 연구실적 평가 : 45% ※ 연구실적 산정 기간 : 최근 5년간 - 학문분야별 상위 10%저널에 게재된 논문 수 (인문사회분야의 경우 SSCI, A&HCI 및 SCOPUS급 저널, 해외 유명 저널에 게재된 논문수, 해외 유명 저널은 전공패널심사에서 심의·인정) - 연구팀 피인용 총횟수 - 논문 1건당 피인용횟수 - SCI(E),SSCI,A&HCI, SCOPUS급 학술지의 Editor 경력(인문사회분야의 경우 해외 유명 저널 포함, 해외 유명 저널은 전공패널심사에서 심의·인정) - 연구비 수주 실적 - 국내 및 국외 특허	10%	10%	10%	25%
	10%	10%	5%	5%
	10%	10%	5%	5%
	10%	10%	10%	10%
	5%	5%	10%	-
	-	-	5%	-
· 대학의 여건 및 지원계획 평가 : 10% - 대학의 해당분야 지원 계획의 충실성 - 교육·연구 인프라 여건 ① 대학원생수 대비 최근 5년간 박사학위 배출 실적 ② 교원 1인당 학생수(학부생수+대학원생수) ③ 학부생수 대비 대학원생수(특수대학원생 제외)				
			2%	
			8%	
			(2%)	
			(3%)	
			(3%)	

<표 2>에서 볼 수 있듯이, 전체의 47%(유치·활용계획서평가 45%와 대학의 여건평가 중 2%)가 Peer Review에 의한 평가이고, 53%(연구실적평가 45%와 대학 여건평가 중 8%)가 계량지표에 의한 평가이다.

계량평가의 점수산출은 해당 패널 최고 실적치를 지표내 만점으로 처리하고 비례적으로 점수를 산출하는 방식을 적용하였다.

2) 분석대상 자료현황

최초 게시된 WCU사업에 총 475개 과제가 접수되었다. 본 연구를 위한 자료로 신청된 전체 과제 중 유형3을 제외한 나머지 과제의 평가결과가 활용되었다. WCU사업은 지원방법 등에 따라 3

개 유형으로 구분되어 있으며, 그 중 유형3의 경우 계량평가가 미적용 됨에 따라 해당 과제 161개는 분석대상에서 제외하였다. 유형1과 유형2과제는 Peer Review에 의한 정성평가와 계량지표에 의한 정량평가가 각각 47%와 53% 적용됨에 따라 동 평가결과를 100%활용하여 상관분석 및 정책적 시사점을 모색할 수 있었다.

3) Peer Review와 계량지표에 의한 평가 현황 분석

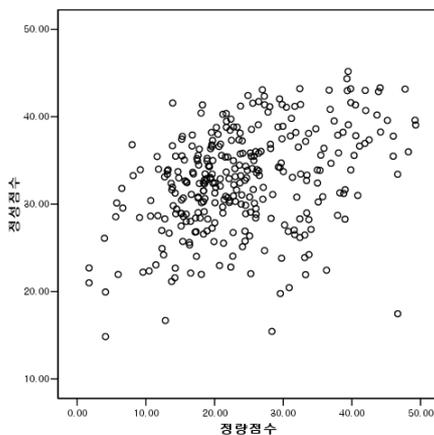
WCU사업 선정평가 Peer Review 결과 점수분포를 분석하였다. 유형1, 2 접수과제 총 314과제는 자연과학 63개 과제, 공학 130개 과제, 생명과학 75개 과제, 인문사회 46개 과제로 각 분야별 평균 및 편차는 아래 표와 같다. 인문사회 분야의 평균 점수가 대체로 타 분야에 비해 낮은 것을 볼 수 있으나 나머지 분야에서는 평균과 편차가 분야간 유사한 수준을 보인다. 평가결과의 평균이 Peer Review가 계량지표에 의한 평가보다 월등히 높은 것과 편차가 계량지표 평가에서 크게 높은 것이 특징으로 나타났다.

<표 3>WCU사업 분야별 점수분포

분야	과제수	Peer Review				신규계량지표 적용				편차비 (B/A)
		최저점수	최고점수	평균	편차 (A)	최저점수	최고점수	평균	편차 (B)	
자연과학	63	46.73	91.84	70.65	10.73	12.54	90.10	46.87	19.38	1.81
생명과학	75	47.27	91.94	73.25	11.12	10.87	92.81	47.94	18.05	1.62
공학	130	35.52	96.13	71.20	10.39	23.25	92.99	44.74	15.41	1.48
인문사회	46	31.60	86.62	57.49	13.42	3.26	88.06	41.75	22.77	1.70

4) Peer Review와 평가지표별 상관도 분석결과

유형1, 2 전체 과제에 대한 Peer Review와 신규 계량지표 적용 평가의 상관도를 분석하여 아래와 같은 결과를 얻었다.



상관계수

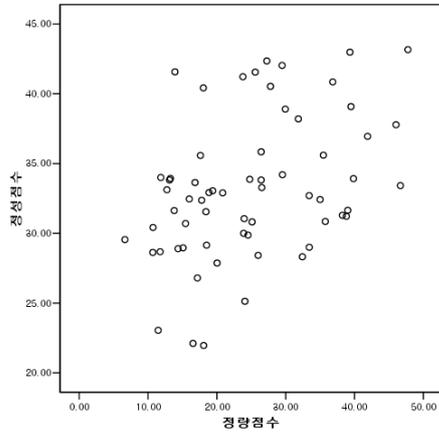
		정성점수	정량점수
정성점수	Pearson 상관계수	1	.388(**)
	유의확률 (한쪽)		.000
	N	314	314
정량점수	Pearson 상관계수	.388(**)	1
	유의확률 (한쪽)	.000	
	N	314	314

** 상관계수는 0.01 수준(한쪽)에서 유의합니다.

(그림 3) WCU사업 Peer Review와 계량지표평가 상관도

산점도를 보면 Peer Review와 신규 계량지표 적용 평가 사이에 선형 양의 상관관계를 볼 수 있으며, 1% 유의수준(Pearson계수 0.388)에서 상관도가 나타나는 것을 알 수 있다.

서두에 밝힌바와 같이 WCU 사업 평가는 자연과학, 생명과학, 공학, 인문사회의 4개 분야분류에 따라 평가를 진행하였다. 각 분야별 상관도는 아래와 같다.

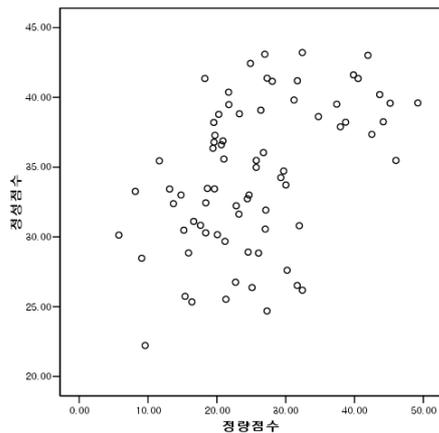


상관계수

		정성점수	정량점수
정성점수	Pearson 상관계수	1	.407(**)
	유의확률 (한쪽)		.000
	N	63	63
정량점수	Pearson 상관계수	.407(**)	1
	유의확률 (한쪽)	.000	
	N	63	63

** 상관계수는 0.01 수준(한쪽)에서 유의합니다.

(그림 4) WCU2차 사업 Peer Review와 계량지표평가 상관도(자연과학)

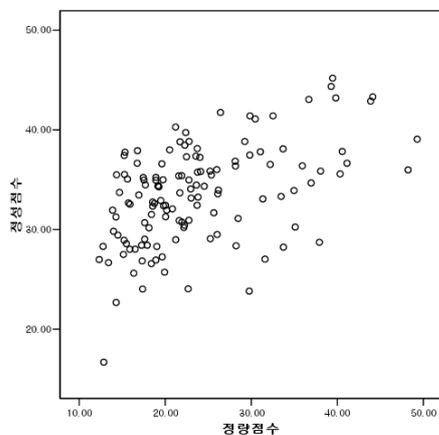


상관계수

		정성점수	정량점수
정성점수	Pearson 상관계수	1	.458(**)
	유의확률 (한쪽)		.000
	N	75	75
정량점수	Pearson 상관계수	.458(**)	1
	유의확률 (한쪽)	.000	
	N	75	75

** 상관계수는 0.01 수준(한쪽)에서 유의합니다.

(그림 5) WCU사업 Peer Review와 계량지표평가 상관도(생명과학)

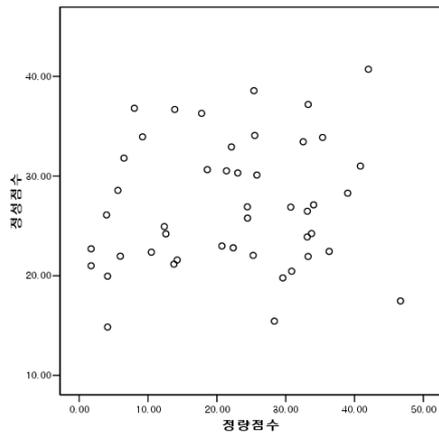


상관계수

		정성점수	정량점수
정성점수	Pearson 상관계수	1	.497(**)
	유의확률 (한쪽)		.000
	N	130	130
정량점수	Pearson 상관계수	.497(**)	1
	유의확률 (한쪽)	.000	
	N	130	130

** 상관계수는 0.01 수준(한쪽)에서 유의합니다.

(그림 6) WCU사업 Peer Review와 계량지표평가 상관도(공학)



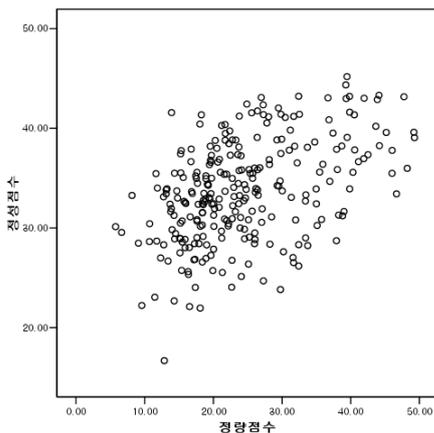
상관계수

		정성점수	정량점수
정성점수	Pearson 상관계수	1	.135
	유의확률 (한쪽)		.185
	N	46	46
정량점수	Pearson 상관계수	.135	1
	유의확률 (한쪽)	.185	
	N	46	46

(그림 7) WCU사업 Peer Review와 계량지표평가 상관도(인문사회)

분야별 분석결과는 매우 의미있는 결과를 보여준다. 자연과학, 생명과학, 공학 전분야에서 Pearson 상관계수는 각각 0.407, 0.459, 0.497로 분석되었으며 모두 1% 유의수준에서 양의 상관관계를 갖는 것으로 나타났다. 그러나 인문사회분야의 경우 0.135로 분석되어 유의한 상관도를 볼 수 없었다. 이는 다시말하면 신규 계량지표 적용 평가는 이공학 전 분야에서 Peer Review 결과와 유사한 결과를 갖는 반면, 인문사회의 경우는 논문, 에디터 등의 신규 연구실적 계량지표 적용 평가가 Peer Review 결과와 연관성이 낮다는 사실을 보여준다. 따라서 신규 개발된 계량지표에 의한 평가는 이공학 분야에 적용하는 경우 그 실효성을 얻을 수 있으며, 인문사회 분야의 경우 분야 특성에 적합한 연구실적 평가지표의 개발이 향후 연구되어야 할 필요성이 있다는 결론을 얻을 수 있다.

Peer Review와 신규 계량지표 적용 평가의 상관성이 없는 인문사회 분야를 제외한 전체 이공학 분야의 상관도는 아래와 같이 분석되었다. 인문사회 분야 모집단 46개를 제외한 268개 과제의 점수 상관도는 1% 유의수준(Pearson계수 0.461)에서 양의 상관관계를 나타낸다.



상관계수

		정성점수	정량점수
정성점수	Pearson 상관계수	1	.461(**)
	유의확률 (한쪽)		.000
	N	268	268
정량점수	Pearson 상관계수	.461(**)	1
	유의확률 (한쪽)	.000	
	N	268	268

** 상관계수는 0.01 수준(한쪽)에서 유의합니다.

(그림 8) WCU사업 Peer Review와 계량지표평가 상관도(이공학과제)

WCU사업에서 도입한 공통된 신규 계량지표는 상위 10%저널 게재 논문 지표, 인당피인용횟수 지표, 편당피인용횟수 지표, 에디터실적 지표, 특허실적 지표 등 5개이다. 이들 지표별 상관관계를 통해 각 지표가 갖는 의미를 분석할 수 있다. 따라서 신규 계량지표별 분석은 Peer Review와 계

량지표 평가에서 상관성이 입증된 이공학 분야 전체 과제를 대상으로 분석하였다.

<표 4> WCU사업 신규 계량지표별 Peer Review와의 상관관계

구분		Peer Review	Top 10% Journal	Citation per person	Citation per paper	Editor	Patent
Peer Review	Pearson 상관계수	1	.372(**)	.334(**)	.277(**)	.277(**)	.264(**)
	유의확률 (한쪽)		.000	.000	.000	.000	.001
	N	268	268	268	268	268	130
Top 10% Journal	Pearson 상관계수	.372(**)	1	.732(**)	.425(**)	.445(**)	.371(**)
	유의확률 (한쪽)	.000		.000	.000	.000	.000
	N	268	268	268	268	268	130
Citation per person	Pearson 상관계수	.334(**)	.732(**)	1	.765(**)	.424(**)	.269(**)
	유의확률 (한쪽)	.000	.000		.000	.000	.001
	N	268	268	268	268	268	130
Citation per paper	Pearson 상관계수	.277(**)	.425(**)	.765(**)	1	.237(**)	.182(*)
	유의확률 (한쪽)	.000	.000	.000		.000	.019
	N	268	268	268	268	268	130
Editor	Pearson 상관계수	.277(**)	.445(**)	.424(**)	.237(**)	1	.091
	유의확률 (한쪽)	.000	.000	.000	.000		.151
	N	268	268	268	268	268	130
Patent	Pearson 상관계수	.264(**)	.371(**)	.269(**)	.182(*)	.091	1
	유의확률 (한쪽)	.001	.000	.001	.019	.151	
	N	130	130	130	130	130	130

** 상관계수는 0.01 수준(한쪽)에서 유의합니다.

* 상관계수는 0.05 수준(한쪽)에서 유의합니다.

신규 계량 지표별 상관도 분석결과 유의수준이 약간의 차이를 보이기는 하지만 이공학 전체 분야에 적용된 신규 계량지표 모두 1% 유의수준에서 Peer Review결과와 양의 상관관계를 갖는 것을 볼 수 있다.

5) 상관계수를 이용한 계량지표 배점기준 설정

동 결과를 통해서 신규 적용된 4개 지표의 타당한 배점 비율을 도출할 수 있다. Peer Review와의 상관도를 지표별로 추출하면 아래 표와 같다.

<표 5> WCU사업 신규 계량지표별 Peer Review와의 상관계수

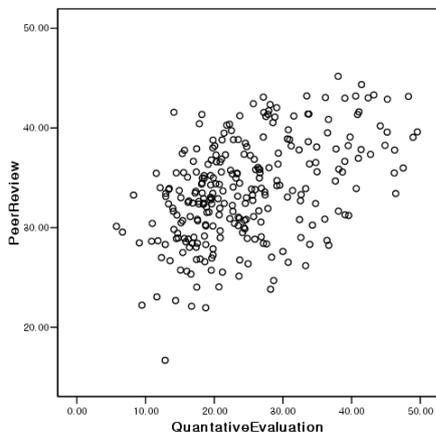
구분	Top 10% Journal	Citation per person	Citation per paper	Editor	Patent
Pearson 상관계수	0.372	0.334	0.277	0.277	0.264

Peer Review와의 상관계수를 지표별 배점기준으로 도입하면, 아래와 같은 결과를 얻을 수 있다.

<표 6> WCU사업 계량지표별 상관계수 반영 배점표

구분	기존		상관도 비율 변환	
	자연/생명	공학	자연/생명	공학
Top10	10	10	12	9
인당피인용	10	5	10	8
편당피인용	10	5	9	6
Editor	10	10	9	6
특허	-	5		6
배점 합	40	35	40	35

위 표의 새로운 배점표에 의한 과제 점수 재산정 결과를 Peer Review와 상관분석을 수행하여 이공학 과제 268개에 대한 분석결과보다 상관계수가 다소 높은 결과를 도출하였다.



상관계수

		정성점수	정량점수
정성점수	Pearson 상관계수	1	.473(**)
	유의확률 (한쪽)		.000
	N	268	268
정량점수	Pearson 상관계수	.473(**)	1
	유의확률 (한쪽)	.000	
	N	268	268

** 상관계수는 0.01 수준(한쪽)에서 유의합니다.

(그림 9) WCU사업 Peer Review와 계량지표평가 상관도(이공학과제)

3. 결론

본 연구에서는 그동안 많은 연구를 통해 논란이 되어온 계량지표의 연구사업 평가 적용을 실증적으로 검증하는 분석을 진행하였다. 일부 연구사업 평가에서 논문게재 수 등을 소폭 적용한 사례는 많이 있으나 표준화된 계량지표의 전면적인 도입사례는 찾아볼 수 없었으며, 거의 모든 국가연구개발사업에 대한 평가는 Peer Review시스템에 의존하고 있는 현실이다. 2008년 신규 기획되고 추진된 세계수준의 연구중심대학(WCU) 육성사업은 수회의 전문가자문 등을 거쳐 선정 1단계 평가에서 신규 계량 지표를 설정하고 약 50% 수준의 비율로 적용시키는 최초의 사례가 되었다.

신규 도입된 계량지표는 다음의 몇가지 특징을 갖는다. 첫째, 논문관련 지표는 SCI와 SCOPUS DB를 동시에 적용하였다. 그동안 계량평가에 사용된 논문지표는 거의 SCI에 의존하고 있다. 그러나 최근 세계 대학평가(The Times)에서는 논문지표로 SCOPUS를 활용하는 등 DB의 다양성이 확대되고 있다. 따라서 WCU 사업에서는 SCI와 SCOPUS 두 개 DB를 모두 활용하여 지표화 하였다. 둘째, 두 개 서지학 DB의 세부 분야분류에 의한 상위 10% 저널 리스트 추출 및 해당 저널 게재논문에 대한 평가를 수행하였다. 이는 단순히 양적으로 SCI와 SCOPUS DB에 등재된 저널 게재 논문 수를 평가하는 것을 넘어선 분야별 IF가 상대적으로 높은 저널에 게재된 논문을 평가함으로써 질적으로 인증된 논문의 수를 측정하는 특징을 갖는다. 셋째, 모든 계량지표는 각 실적의 참여 방법에 따라 점수가 차등 부여된다. 논문지표의 경우 단독저자인 경우 100%, 주저자 또는 교신저자의 경우 80%, 공동저자인 경우 1/n%로 수치화되며, 특허의 경우도 단독발명, 공동발명에

따라 100%, 1/n%로 수치화 된다. 넷째, 논문, 특허 등 기존의 계량실적 외에 편집위원 참여실적이 도입되었다. SCI, SCOPUS급 저널의 Editor는 학계의 저명도를 나타내는 지표가 될 수 있음을 고려하여 신규 도입된 지표이다.

앞서 분석한 내용과 같이 WCU 평가 결과를 통해 계량지표의 적용 효용성을 분석하고 다음과 같은 의미있는 결과를 도출하였다.

첫째, 계량지표에 의한 평가결과는 Peer Review 시스템보다 변별력이 크다는 점을 알 수 있다. Peer Review 결과는 분야에 따라 일부 차이가 있으나 평균을 중심으로 소폭의 편차를 보이는 반면, 계량지표에 의한 평가는 표준편차가 상대적으로 높아 평가결과에 Peer Review보다 큰 영향을 주는 것으로 분석된다. <표 3>의 Peer Review결과와 계량지표에 의한 평가결과의 편차비를 분석하면, 인문사회를 제외한 이공학분야 편차비의 평균이 1.64로 나타났다. 이는 계량지표에 의한 평가가 변별력에 1.64배 높은 비중으로 영향을 미쳤다는 것으로 해석할 수 있고, 이공학분야 평가에 있어서 Peer Review의 배점을 계량지표 적용 배점보다 다소 높일 필요가 있다는 결론을 도출할 수 있다. 이에 대한 검토는 향후 연구를 통해 진행될 필요성이 있다.

둘째, 연구실적 계량지표 점수와 Peer Review와의 상관도 분석을 통해 계량지표의 적용이 가능한 분야를 검증하였다. 이공학 전분야에서 계량지표 적용결과와 Peer Review가 1%유의확률에서 양의 상관관계를 보였다. 이를 통해 이공학 전분야의 평가에 있어서 신규 계량지표 적용의 효용성이 실증적으로 검증된 결과이다. 인문사회 분야의 경우 계량적으로 측정된 결과가 Peer Review와 일관된 상관관계를 보이지 못하였다. 따라서 인문사회분야의 연구실적 계량평가는 분야 특성에 적합한 새로운 지표의 개발이 필요하다는 결과를 얻을 수 있었다.

셋째, WCU사업은 철저한 질 중심의 계량지표를 개발하고 추진하였다. 지표역시 단순한 논문수를 벗어나 IF 순위 분야별 상위 10%저널게재 논문수, 참여학자당 피인용수 등 새로운 지표를 도입 활용하였다. 이공학 분야의 각 지표별 상관도 분석 결과 설정된 모든 계량지표가 Peer Review 결과와 1% 유의수준에서 양의 상관도를 보이면서 지표설정의 타당성을 입증하게 되었다.

넷째, 신규 개발된 5개 지표에 대한 상관계수 반응을 통해 상위 10%저널 게재 논문수, 참여학자당 피인용수, 논문당 피인용수, 편집위원참여실적, 특허실적의 배점비율을 <표 6>과 같이 조정하는 것을 제안할 수 있다.

서론에서 언급한 바와 같이 국가연구개발사업의 효과적인 지원이 국가경쟁력 발전에 기여할 수 있다. 공정성과 전문성이 확보된 객관적인 평가체계를 구축하기 위해서 계량지표의 도입은 매우 효과적인 방법이 될 수 있다. 본 연구를 통해 검증된 계량지표 외에 우수한 연구성과를 측정할 수 있는 새로운 지표의 개발 및 인문사회 분야에 적합한 지표의 탐색 등 후속 연구가 필요하다.

참고문헌

- 김원 (1995), “SCI에 의존하는 한국의 정보통신 분야 연구실적평가의 문제점 및 해결책”, 『정보과학회지』, 17(5): 61-64
- 여운동 · 이우형 · 이상필 (2007), “과학계량학을 이용한 연구활동 평가 : OLED 사례분석”, 『기술혁신학회지』, 10(2): 230-254
- 연경남 · 이성중 · 이종현 · 송충한 (2005), “연구계획서 평가시 정량지표 도입의 타당성에 관한 분석”, 『기술혁신학회지』, 8(1): 261-276
- 이한진 외 (2004), 『과학재단 연구과제의 질적 수준 분석』, 한국과학재단
- 정진완 · 문수복 · 민상렬 · 이재용 · 최진영 (2004), “SCI:과연 정보과학 분야에서도 유효한 연구업적 평가척도인가?”, 『정보과학회지』, 22(8): 70-76
- 허정은 · 김해도 · 최태진 · 김성백 · 조영돈 · 김인호 (2006), “Nature, Science, Cell지에 게재된 한국

- 인 과학자의 논문현황 분석”, 『기술혁신학회지』, 9(3):558-577
- 허정은 · 김해도 · 조영돈 · 조석민 · 조순로 (2008), “국가연구개발사업의 과학적 성과분석을 위한 새로운 계량지표 개발에 관한 연구”, 『기술혁신학회지』, 11(3): 376-399
- COSEPUP(Committee on Science, Engineering, and Public Policy, USA), (1999), Evaluating Federal Research Programs: Research and the Government Performance and the Results Act, Washington D.C.
- Cole. F. J and Eales. N. B (1917), “The history of comparative anatomy. Part 1: A statistical analysis of the literature. Sci. Progr.”(London), 44
- E. Harding (2002), “Peer Review”, Parliamentary Office of Science and Technology Postnote, 182: 1-4
- Hurt, C. D (1985), “Identification of important authors in science : a coparison of two methods of identification”, Information Processing & Management, 21:177-186
- J. E. Hirsch (2005), “An index to quantify an individual’s scientific research output”, PNAS, 102(46):16569-16572
- Jeong Dong Youl (1990), “Values and Future Research Issues in Bibliometrics ”, 『도서관학』, 19: 243-261
- Kostoff, Ronald N.(1995), “Federal Research Impact Assessment: Axioms, Approaches, Applications”, Scientometrics, 34:163-205
- L. Miles Raisig (1962), “Statistical Bibliography in the Health Sciences”, Bull Med Libr Assoc, 50(3): 450-461
- Leo Egghe (2006), “Theory and practise of the g-index”, Scientometrics, 69:131-152
- Marshakova-Shaikevich, I. (1996), “The Standard Impact Factor as an Evaluatiohn Tool of Science and Scientific Journals”, Scientometric, 25: 283-290
- Prichard. A. (1969), “Statistical bibliography or bibliometrics”, Journal of documentation, 358-359
- Pudovkin A. I. and Garfield, E. (2004), “Rank-Normalized Impact factor: A Way to Compare Journal Performance Across Subject Categories”, Proceedings of the 67th ASIS&T Annual Meeting, 41: 507-515.
- Sen, B. K (1992), “Normalized Impact Factor”, Journal of Documentation, 48:318-329
- Velho, L and Krige, J. (1984), “Publication and citation practiecs of Brazilian agricultural scientists”, Social Studies of Science, 14:45-62