

## 연구계획서 평가시 정량지표 도입의 타당성에 관한 분석

### A Study on the use of quantitative indicators for the research proposal assessment

연경남\* 이성종\* 이종현\* 송충한\*

#### 〈목 차〉

- |                  |                            |
|------------------|----------------------------|
| I. 서론            | III. 정량지표사용 사례 분석 및 대안의 제시 |
| II. 기존 연구에 대한 개관 | IV. 요약 및 결론                |

### Abstract

Peer review is a well-known, pervasive and most reliable method for assessment of research proposals. But with the strength of peer review, it also has some various limitations and shortcomings. This study suggests the use of bibliometric indicators as a counterbalance of shortcomings and mistakes in peer review. Bibliometric indicators suggested in this study are somewhat different from the conventional ones. Bibliometric indicators with limited number of papers, use of impact factor and discrimination first and corresponding authors from co-authors may produce a more reasonable indicators. In this way bibliometric indicators can be considered as complementary to a peer review.

**Keyword** : peer review, bibliometric indicators, research assessment,

---

\* 한국과학재단, chsong@kosef.re.kr

## I. 서론

평가는 사회 제반문제를 다루는데 있어 과학적인 방법을 이용하여 해결하고자 하는 노력에서 출발하였다. 근대적인 평가의 기원은 대략 17세기 정도까지 거슬러 올라가나 체계적이고 자료에 근거한 평가는 비교적 최근에 활발하게 이루어지고 있다.(김인목 외, 2002) 이러한 평가에 대해 Rossi and Freeman(1989)은 평가란 공공정책의 수행에 따른 종합적인 분석을 위해 여러 학문분야의 복합적인 기법과 개념을 도입한 과학적 연구방법론을 적용하는 것이라고 정의하고 있다. 따라서 평가의 개념은 정부 정책과 프로그램의 개념화 설계, 기획 및 실행의 모든 단계에서 적용될 수 있을 것이다. 그러나 여기서는 이러한 광범위한 평가의 개념을 한정하여 연구개발활동에 대한 평가로 국한하고자 한다. 특히, 연구개발활동 중에서도 연구과제의 선정단계에서 이루어지는 평가로 범위를 좁히고자 한다. 왜냐하면 연구개발단계에 적용될 수 있는 광범위한 개념의 평가(Assessment)는 선정단계에서의 평가(Appraisal), 중간평가(Monitoring), 중간점검(Review), 결과평가(Evaluation)로 구분될 수 있는데(김인목 외, 2002), 각 단계마다 평가의 목적과 사용하는 방법론에서 차이가 있기 때문에 여기서는 분석대상을 선정단계에서의 평가로 국한하고자 한다.

본 연구의 주요 목적은 연구과제의 선정평가 단계에서 주로 적용되는 동료평가(peer review)방법에 대한 대안(代案) 또는 보완 수단으로서 계량서지학(bibliometrics)을 이용한 계량서지학적 지표(bibliometric indicators)와 특허 등을 모두 포함하는 정량지표의 도입 방안에 대해 분석하고자 하는 것이다<sup>1)</sup>. 이를 위해 본 연구에서는 동료평가의 개념과 동료평가의 장단점, 정량지표의 개념과 장단점, 그리고 동료평가와 정량지표지표의 관계 등에 대해 살펴본 후, 실제로 적용되고 있는 사례분석을 통해 정량지표의 도입이 과연 타당한지 그리고 정량지표를 도입하는 경우 어떻게 도입하는 것이 바람직한지에 대해 살펴보려고 한다.

1) 여기서 지표의 용어를 '정량지표'로 사용한 것은 지표를 구성하는 주요 내용이 계량서지학(bibliometrics)을 이용한 계량서지학적 지표(bibliometric indicators)뿐만 아니라 특허 등 기타의 지표도 포함하기 때문이다.

## II. 기존 연구에 대한 개관

동료평가(peer review) 방식은 동료 과학자의 연구활동에 대해 해당분야의 전문적인 지식을 가진 동료 전문가가 설정된 기준의 의거하여 연구의 질적 우수성을 체계적으로 그리고 주관적으로 측정하는 평가방법이며 이러한 동료간 조정기능을 통해 해당 분야 연구활동의 방향이 정립되고, 연구결과의 가능성이 설정되며 또한 한정된 과학기술 자원이 효과적으로 그리고 효율적으로 배분된다고 정의하고 있다.(Chubin & Hackett, 1990; OECD, 1987)

GAO(1999)는 동표평가를 연구의 성격과 목적의 차이, 연구기간의 차이, 연구성과의 광범위한 측면, 자금지원 방식의 변화(연구장려금, 연구협약 등)등 연구개발활동의 평가시 각각의 경우에서 발생하는 변동사항을 모두 반영할 수 있는 적절하고 필수적인 방법으로 인식하고 있다. COSEPUP(1999)도 동료평가(peer review)방법이 정부의 연구개발 프로그램을 평가하는 가장 효과적인 방법이라고 결론짓고 있다. 이와 같이, 동료평가는 객관적으로 수량화되지는 않지만 연구의 질적인 측면을 평가하는 방법으로서 연구의 영향 평가에 가장 광범위하게 사용되고 있으며 신뢰할 수 있는 수단으로 인정받고 있다.<sup>2)</sup>

그러나 동료평가는 인간의 지적활동에 중심을 두고 있기 때문에 인간의 심리적인 측면에서 이해의 충돌로부터 발생하는 다양한 주관적인 편파성이 평가에 부정적인 영향을 미치게 된다. 따라서 동료평가는 이러한 인간의 주관적 판단이 주변의 여건에 의해 영향을 받아 보편타당한 객관성과 공정성을 일정 부분 상실하고 있다는 점에서 결정적으로 비판을 받고 있다(김인목 외, 2002). 동료평가가 갖는 이와 같은 문제점에 대해 Kostoff(1999)는 다음과 같이 정리하고 있다. 첫째, 평가자가 해당 분야의 전공과 정확하게 일치하지 않을 경우 평가결과가 왜곡될 수 있다. 이러한 문제점은 동료평가의 결과가 50% 이상 운에 따라 결정될 수도 있다는 비판(Nederhof, 1988)을 상기시키기도 한다.<sup>3)</sup> 둘째, 기존의 분야를 보호하기 위한 중진들의 네트워크('Old Boy' network)이 평가를 왜곡시킬 수 있다. 즉, 새롭게 나타나는 첨단 분야보다 기존의 전통적 연구 분야에서 연구비 수혜비율이 훨씬 높게 나타나고 있으며, 이러한 점에서 동료평가는 과학기술혁신과정에서 그다지 효과적인 평가방법이 아니라는 비판이

2) 동료평가는 그 방식에 따라 우편평가(mail-type peer review)와 위원회평가(panel-type peer review) 그리고 혼합평가(combination of mail and panel)로 구분될 수 있다.

3) Aksnes and Taxt(2004)에서 재인용.

제기되고 있다. 셋째, 유명한 연구자, 소속기관 등에 대해 후광효과('Halo' effect)가 작용할 수 있어 연구내용의 질적 수준과 무관하게 연구비를 지원 받을 가능성이 상대적으로 높다. 넷째, 개별 평가자마다 각기 다른 해석과 평가의 기준이 적용될 수 있다. 즉, 동료평가는 평가자 각자의 주관적인 판단에 의해 이루어지기 때문에 동일한 판단기준을 적용하기가 매우 어렵다.<sup>4)</sup> 다섯째, 동료평가의 전제조건으로 질적으로 우수한 연구과제를 구성하는 요소가 무엇이고 이를 이끄는 과학자가 누구인지, 향후 이러한 혁신적인 방향이 어느 쪽으로 진행될 것인지에 대해 평가자간 상당 수준의 동의와 의견일치를 요구한다는 점이다. 여섯째, 많은 비용이 소요되는 문제가 발생할 수 있다.<sup>5)</sup>

따라서 동료평가가 올바른 평가방법이 되기 위해서는 몇 가지 특성을 갖추어야 한다. 이에 대해 Chubin(1994)은 올바른 동료평가 방법이 갖추어야 하는 특성으로서 효과적인 자원 배분기능, 과학의 책무(science accountability)를 증진시키는 기능, 합리적이고 공정한 과정, 과학적 성과에 대한 유효하고 신뢰할 수 있는 측정기능 등을 들고 있다<sup>6)</sup>. 또한 Kostoff(1997)는 훌륭한 동료평가가 갖추어야 할 조건으로서 첫째, 평가의 수단, 조직, 기준이 특정한 평가 상황에 맞게 조정되어야 하며 둘째, 평가단계에서 프로그램 혹은 프로젝트의 목표가 중요하게 고려되어야 하고 셋째, 평가의 목적이 명확하게 체계화되어야 하며 넷째, 평가의 신뢰성이 확립되도록 하여야 하고, 끝으로 평가의 디자인 단계에서 평가결과의 효과적인 활용을 고려하여야 한다는 것 등을 제시하고 있다.

연구결과를 평가함에 있어 정량지표를 이용한 접근방법은 1세기정도 전부터 시도되어 왔지만 실질적인 의미의 정량적 분석방법은 1960년대부터 본격적으로 시작되었다. 1969년에 Pritchard는 당시에 활발해지기 시작한 정량적 분석 시도와 일련의 연구를 하나의 포괄적인 학문분야로 표현해보기 위해 당시에 사용했던 'statistical bibliography'라는 용어를 대신하여

4) 동료평가가 가지고 있는 이러한 문제점을 해결하기 위해 '연구결과의 절반이상이 세계적 우수성을 가지고 있으며, 나머지도 영국에서 우수한 수준에 포함되는 경우', '모든 연구결과가 영국의 우수한 수준에 해당되며 몇 개의 연구는 세계적으로 우수한 경우' 등 주관적인 평가의 척도를 제공하는 경우도 있음.(HEFCs, 1998)

5) 동료평가에 소요되는 비용은 직접비용과 간접비용으로 구분될 수 있는데, 직접비용은 평가에 직접 소요되는 비용이며, 간접비용은 대부분 평가인력이 소비하는 시간으로 구성된다. 평가인력이 평가에 참여함으로써 소비하는 시간은 비용으로 계산할 때 결코 무시할 수 없는데, 특히 평가에 참여하는 인력이 수준 높은 연구자일 경우 그 비용은 더욱 커지게 된다.

6) Chubin(1994), Kostoff(1997)에서 재인용

'bibliometrics'라는 새로운 개념을 창출해 냈으며 이것이 정량적 지표평가 및 분석의 기원이 되고 있다(OECD, 1997; Pritchard, 1969). 연구결과를 평가하는데 주로 사용된 계량서지학이 연구계획서의 평가에 도입되는 것은, 대부분의 연구계획서 평가단계에서 연구책임자의 연구 수행 능력을 평가에 반영하고 있고, 연구책임자의 연구수행 능력의 평가는 연구책임자가 발표한 논문의 양적 및 질적 수준을 통해 이루어지기 때문이다.

Leeuwen, Wurff and Raan(2001), Aksnes and Taxt(2004), CWTS(2000) 등은 계량서지학을 이용한 정량지표의 종류를 다음과 같이 정의하고 있다. 첫째, 개인(그룹)이 발표한 전체 논문 건수(P) 둘째, 자체인용(self-citation)을 포함한 전체 피인용 횟수(C) 셋째, 논문편당 평균 피인용 횟수(CPP=C/P) 넷째, 자체인용을 제외한 논문편당 평균 피인용 횟수(CPPex) 다섯째, 분석 대상 기간동안 한번도 인용되지 않은 논문의 비율(%Pnc) 여섯째, 각 논문의 평균 피인용 횟수(CPP)와 논문이 게재된 저널들의 평균 피인용 횟수(JCSm)간의 비율(CPP/JCSm; 이 비율이 1보다 큰 경우 개인(그룹)의 논문이 해당 저널에 게재되어있는 논문들보다 영향력이 크다는 것을 의미한다). 일곱째, 각 논문의 평균 피인용 횟수(CPP)와 해당 논문이 제출된 분야에서의 저널들의 평균 피인용 횟수(FCSm)간의 비율(CPP/FCSm; 이 비율이 1보다 큰 경우 개인(그룹)의 논문이 해당 분야의 저널들의 영향력보다 크다는 것을 의미한다) 이다.

이외에도 정량지표들은 다양한 방법으로 변형되어 사용되기도 한다. 예를 들어 Lewison, Cottrell and Dixon(1999)은 인용횟수 등을 절대값으로 사용하지 않고 각 논문의 피인용 횟수가 해당 분야의 논문들이 갖는 피인용 횟수 중에서 어느 정도의 위치(rank)에 있는지를 파악하여 이 논문이 인용횟수 상위 1/4에 위치하면 4의 가중치를 부여하고 그다음 1/4에 위치하면 3의 가중치를 부여하는 등 인용횟수와 관련된 지표를 상대값으로 변환하여 사용하고 있다. 이러한 방식은 논문 게재, 인용 등 분야가 갖는 특성에 따른 인용값의 변화를 비교가능한 지표로 변환시켜 사용한다는 장점을 가지고 있다.

그러나, 일반적으로 정량지표들은 동료평가 과정에 수학적 엄격성과 정확성을 부여하는 도구임에도 그 자체로는 질적 측면을 고려하는데 한계를 가지고 있다 (김인목 외, 2002). Kostoff(1995) 및 van Raan(1996) 등은 정량지표가 갖는 문제점들을 다음과 같이 지적하고 있다. 첫째, 인용횟수는 연구의 질이 아닌 양적인 실적만을 측정할 뿐이다. 둘째, 학술지에 게재되는 논문의 유형은 분야별로 다양할 수 밖에 없는데 단순한 건수 위주의 측정은 분야별 다양성과 특성을 전혀 반영하지 못한다. 셋째, 자기인용 및 조각논문발표(7)에 영향을 많이

받는다. 넷째, 복수저자 및 연구자그룹에 의한 논문에 있어 개인별 업적이나 실질적 기여도 등을 구별하지 못한다. 다섯째, 과학기술 분야에서 지식의 의사소통 방식에 있어 논문이 아닌 방식에 대하여는 이를 무시하게 된다. 여섯째, 영어권 국가에서 발간하는 학술지가 다수를 차지하고 있어 영어 중심의 언어 편기(bias)가 존재한다. 일곱째, 긍정적 인용과 부정적 인용의 특성을 구별하지 못한다. 여덟째, 연구활동이 종료된 시점과 이의 연구결과에 의한 논문발표 시점 및 인용된 시점과의 상당한 시간적 차이가 있다.

위에서 살펴본 바와 같이, 동료평가와 정량지표를 이용한 평가는 모두 나름대로의 장점과 단점을 가지고 있다. 각각의 방식이 가지고 있는 장단점을 보완하기 위해 많은 연구자들이 동료평가와 정량지표의 보완적 성격을 강조하고 있다. Luukkonen(1991)은 인용횟수가 동료평가 순위와 대략적인 상관관계를 가지는 경향이 있다는 것을 분석하였으며, Rinia et al(1998)은 응집물질물리 분야의 연구프로그램을 평가한 결과 여러 가지 인용지수가 동료평가 순위와 밀접한 상관관계를 나타낸다는 결과를 얻었고, Oppenheim(1997)는 영국의 유전학, 해부학, 고고학 연구에 대한 1992년도 연구 수행결과를 평가한 동료평가 순위와 인용지수 사이의 강한 긍정적 상관관계가 존재한다는 것을 발견하였다. 아울러, Aksnes and Taxt(2004)는 계량서지학적 지표가 동료평가의 단점과 실수를 막기 위한 평형추로서의 역할을 할 수 있다고 주장하고 있다.<sup>8)</sup> 따라서, 계량서지학적 지표는 동료평가에 대한 보조장치로서 활용되는 것이 바람직하다고 제안하고 있다.

### III. 정량지표사용 사례 분석 및 대안의 제시

계량지표를 실제적으로 평가에 사용한 사례는 그리 흔하지 않다. 이는 앞서 언급한 바와 같이, 계량지표가 갖는 근본적인 제약도 있지만 현실적으로 계량지표에 접근하는 것이 용이하지 않다는 점과 함께 계량지표를 유효한 지표로 활용하기 위해서는 많은 시간과 금전적 비용이 소요되기 때문이다. 여기서는 영국의 Wellcome Trust에서 적용한 사례와, 한국과학재단에서 적용한 사례 그리고 최근 대학의 교수업적평가에서 사용되고 있는 사례를 살펴보고

7) 하나의 연구결과로부터 비슷한 내용의 논문을 다수로 만들어 논문을 발표하는 것으로 논문실적 강조 추세에 대하여 자신의 논문 발표 건수를 증가시켜 보려는 의도에서 나온 부정적인 경향이다.

8) Aksnes and Taxt(2004)에서 재인용.

자 한다. 대학의 교수 업적평가 사례를 살펴보고자 하는 것은 연구자의 연구수행능력에 대한 평가가 결국 연구자가 이룩한 성과를 토대로 이루어질 수밖에 없기 때문이다. 따라서 이는 교수 개개인의 연구성과 평가와 맥을 같이하게 된다.

## 1. 영국의 사례

매년 2000만 파운드의 연구비 배분에 대해 조언하는 영국 Wellcome Trust의 Neuroscience Panel은 1995년 주요 연구비 지원에 대한 peer review결정에 계량지표(bibliometric indicator)에 관한 정보를 제공하기로 결정하였다. 이러한 서지계량적 통계의 사용은 다른 곳에서도 이루어졌지만(Rinia et al, 1998; van Raan, 1998), 서지계량적 방법이 대학사회에서 흔히 사용되는 방법이 아니기 때문에 투명하고 엄격한 과정이 만들어지는 것이 바람직한 것으로 판단되었다.(Lewison, Cottrell and Dixon, 1999)

사무국은 우선 신청자들로 하여금 신청자와 견줄 수 있는 해당분야의 세계적 연구자들을 추천하도록 하였고 아울러 신청자의 지난 10년간 SCI 혹은 SSCI에 게재된 article, notes, review에 한정된 논문실적을 제출하도록 요구하였다. 사무국에서는 신청자와 비교대상 과학자들이 최근 5년간 발표한 저널들을 조사하여 impact factor 가 높은 것에서 낮은 순으로 정리한 후, 상위 1/4이 저널에는 가중치 4를 부여하고, 그 다음 1/4의 저널에는 가중치 3을 그리고 나머지에 각각 2와 1의 가중치를 부여한다.

이와 함께, 사무국은 인용횟수와 관련된 정량지표도 작성한다. 사무국은 인용횟수와 관련된 정량지표의 작성을 위해 우선 논문 기준집단(norm group of papers)을 도출한다. 논문 기준집단은 키워드에 근거한 검색을 통해 통상적으로 수백편의 논문으로 구성되는데, 논문 기준집단은 적어도 신청자의 논문내용의 2/3가 나타날 수 있도록 조합되어야 한다. 그 후, 신청자 논문의 인용횟수가 기준그룹의 인용횟수의 분포 중에서 어디에 해당되는지에 따라 각각 4, 3, 2, 1점의 가중치를 부여한다. 아울러, 논문인용횟수가 상위 5-10%에 속하는 경우에는 해당 논문에 별표의 표시(\*표시)를 하여 해당 논문이 갖는 영향력을 패널 멤버들이 인지하도록 한다.

패널 멤버들은 상기의 방식에 의해 작성된 다양한 정량지표를 이용하여 신청자들의 연구실적을 평가하는데 활용한다. 한 가지 특이한 점은 서지계량적 분석을 이용하여 신청자들을

상대적인 관점에서 평가하는 것이 아니라 신청자들의 기 출판된 논문들이 만족할 만한 수준에 도달하였는지 즉, 신청자들이 연구를 수행할 자격이 있는지의 유무를 판단하는데 이를 사용한다는 것이다.

## 2. 과학재단의 사례

한국과학재단은 다양한 연구비지원 프로그램 중에서 선도과학자육성지원사업(이하 ‘선도과학자사업’)의 경우 신규과제 선정 평가시 정량지표를 사용하고 있다<sup>9)</sup>. 선도과학자사업은 연구능력이 검증된 기초과학분야의 우수 연구자를 선별 지원 하여 세계적으로 높은 평가를 받을 수 있는 연구성과창출과 창의적 인력을 양성하는 것을 목적으로 하고 있기 때문에 연구능력의 검증을 위해 정량지표를 사용하고 있는 것이다. 또한, 동 사업의 지원범위가 수학, 물리학, 지구과학, 화학, 생명과학분야 등 5개의 자연과학 분야로 한정되어 있어, 정량지표의 도입에 큰 문제가 없기 때문이다.

선도과학자사업은 두 번의 단계에서 정량지표를 사용한다. 하나는 신청단계이고 다른 하나는 평가단계이다. 신청단계에서는 미 ISI사가 출판한 JCR(Journal Citation Report)의 인용지수(Impact Factor)가 분야별로 아래 기준이상인 학술지에 책임저자(Corresponding Author)로서 논문을 게재한 연구자로 신청자격을 제한하고 있다.

<표 1> 선도과학자사업 신청자격(분야별 impact factor)

분야	수학	지구과학	물리학	화학	생명과학
인용지수	0.8	1.5	2.0	2.5	5.0

선도과학자사업의 경우 연구자가 해당 impact factor에는 못 미치는 저널에 논문을 게재하였지만, 우수한 논문으로 인정되어 많은 인용횟수를 기록하는 경우 이를 토대로 우수논문으로 인정하여 신청자격을 부여하고 있다. 다만, 인용횟수를 적용하는 경우에는 impact factor의 2배에 해당하는 점수를 요구하고 있다.

9) 선도과학자육성지원사업은 2004년까지 과학재단이 수행하였고 2005년부터는 교육인적자원부로 이관되어 학술진흥재단에서 사업을 수행하고 있음.

&lt;표 2&gt; 분야별 피인용 횟수에 대한 인정점수

구 분	수 학	지구과학	물리학	화 학	생명과학
피인용 횟수*	2회 이상	3회 이상	5회 이상	5회 이상	5회 이상
인정점수 (인용지수환산점수)	0.8	1.5	2.0	2.5	5.0

선도과학자사업에서는 신청과제 평가시 평가점수에 정량지표 50%, 동료평가 50%를 반영하고 있다. 정량지표의 계산에서는 <표2>의 기준에 따라 인용이 일정 수준 이상인 경우 해당 분야의 인정점수를 부여하는 등 연구자의 논문실적 전체의 impact factor 누계를 계산하여 이를 점수로 환산하였다. 환산방법은 해당 분야의 신청자중 인용지수 누계가 가장 높은 점수를 50점, 신청자 전원의 인용지수 누계점수의 평균치를 35점으로 환산한 이를 중심으로 신청자의 각 점수를 비율 배분하였다. 여기서 평균치를 35점으로 정한 것은 분야에 따라 인용지수 누계의 최고치와 최저치가 많은 편차를 보일 경우 이에 따라 발생하게 되는 점수분포의 불규칙성을 최소화하기 위한 것이다.

### 3. 대학의 교수 업적평가 사례

대학에도 점차 교수 연봉제가 도입됨에 따라 연구실적 평가 등에 정량적인 기준이 도입되는 추세이다. 다만 구체적인 산출기준은 각 분야의 특성에 따라 여러 가지 다른 양상을 보인다. 일반적으로 사용되는 정량지표는 주로 게재 저널의 Impact Factor, 논문편수, 피인용 횟수 등이고, 그 외에 특허등록, 해외초청강연, 국제학술회의 발표, 상업화, 저서, 편저 또는 번역 등도 활용되고 있다. 여기서는 국내 A 대학의 사례를 검토하고자 한다.

대학에서의 업적평가는 각 분야에 따라 많은 차이가 있다. 이는 분야에 따라 연구활동의 특성이 다르기 때문인 것으로 파악된다. 여기서는 자연과학분야, 생명과학분야, 공학분야의 업적평가를 각각의 사례로 살펴보고자 한다. 우선 자연과학분야에 속하는 분야의 경우 논문 실적의 평가에 있어서 모든 논문을 동일하게 간주하는 논문수로 평가하지 않고 각 논문의 질을 나타내는 값을 사용하고 있다. 즉, 해당 분야에서 특정 논문이 가지는 질적 지표를 해당분야의 전체 논문 수 중에서 해당 논문이 위치한 백분율을 적용하고 있다. 특 해당 분야의 저널들에 게재된 논문이 총 100편이고 해당 논문이 게재된 저널에 실린 논문이 10편이며

그 분야에서 해당 저널보다 낮은 저널들에 게재된 논문이 60편이라고 한다면, 이 논문의 위치는  $(60+10/2)/100$ 으로서 65점의 점수를 가지게 된다. 또한 이 분야의 경우 모든 공동저자의 기여도는 동일한 것을 원칙으로 하고 있다.

다음으로 생명과학분야를 살펴보면, 생명과학분야는 학술실적, 연구실적 및 국제활동 실적을 모두 업적평가에 포함하고 있다. 학술실적은 SCI등재 논문과 비등재 논문을 구분하여 점수를 부여하고 있으며, 저자의 역할을 구분하여 실적을 인정하고 있다. 저자의 역할은 주저자(first or corresponding author)와 공저자(co-author)를 구분하여 업적을 차별적으로 인정하는데, 공저자의 경우 해당 논문이 갖는 점수를 1/2만큼 인정하되 공저자로 합산된 점수의 누계가 무한정 인정되는 것이 아니라 자신이 주저자(first or corresponding author)로 역할을 한 논문을 통하여 얻은 점수의 50%까지만 인정하고 있다. 이처럼 공저자의 역할을 해당 논문이 갖는 점수의 1/2까지 인정하는 것은 분야 및 연구주체의 특성에 따라 공저자가 상당한 기여를 하는 경우가 발생할 수 있기 때문에 이를 보완하기 위한 것이며, 공저자로 얻은 점수가 주저자로 얻은 점수의 50%이상이 되지 못하도록 한 것은 여기저기 이름만 올려서 논문실적을 쌓는 악습을 방지하기 위한 방안으로 제시된 것이다. 연구실적 및 국제활동 실적과 관련해서는 국외특허등록을 국내특허등록의 4배의 점수로 인정하고 있으며 출원의 경우는 인정하지 않고 있다. 학술활동실적은 국제학술회의에서 경비를 지원받는 경우에 한하여 국내특허등록과 동일한 비중으로 업적을 인정하고 있으며 실용화 또는 상업화 사례가 있는 경우 구체적인 내용에 대해 평가를 실시하여 인정 여부를 결정하고 있다.

공학 분야의 경우에는 상기의 사례 이외에도 논문, 저서, 학술회의 등에 대해 각각의 점수를 부여하여 각 교수별로 총점수를 계산한 다음 교수들 간의 평균과 표준편차를 이용하여 교수들 간의 등급을 부여하는 경우도 있다.

위에서 살펴본 바와 같이 논문을 기준으로 하는 교수의 업적평가 방식은 논문의 impact factor를 고려하는지 그리고 저자의 역할을 고려하는지에 따라 각기 다른 기준이 적용되고 있다. 우선 impact factor를 고려하는 경우에는 각 분야별 저널이 가지는 impact factor의 특성에 따라 분야별로 상당한 차이를 보이고 있다. 우선, 수학이나 지구과학 등 저널별 impact factor의 차이가 미미한 분야의 경우, impact factor의 절대값을 그대로 사용하는 경우 변별력을 확보하기 어려운 문제가 발생한다. 이 경우, 해당 저널이 속한 연구분야 내의 각 저널별 impact factor 순위에서 상위 몇%에 위치하는 지에 따라 가중치를 부여하는 방법을 사용하기

도 한다<sup>10)</sup>. 반대로 생명과학이나 공학분야와 같이 저널별 impact factor의 차이가 많이 나는 경우에는 저널 또는 논문의 순위(rank)를 사용하지 않고 해당 저널의 impact factor를 그대로 사용하기도 한다. impact factor를 고려하지 않는 경우에는 해당 논문이 SCI 등재 논문이지의 여부, 국제 또는 국내 논문인지의 여부 등에 따라 단순하게 가중치를 부여하여 적용하는 것이 통상적인 방법이다. 아울러, 대부분의 경우 논문의 인용횟수(citation number)는 거의 사용을 하지 않고 있다. 이는 각 논문의 인용횟수를 평가에 사용하는 경우 논문DB에 접근하는데 상당한 비용이 소요될 뿐 아니라, 이를 조사하는 단계에서도 많은 시간과 노력이 소요되기 때문이다.

#### 4. 대안적 정량지표의 제시

우리는 위에서 동료평가와 정량지표를 활용한 평가방법에 대해 살펴보았다. 한 가지 우리가 더 고려해야 할 것은 평가방식을 디자인할 때 동 평가방식이 연구자들에게 긍정적인 동기부여를 할 수 있어야 한다.(송충한, 1999) 예를 들어, 평가과정에서 논문의 편수를 중요시하는 경우, 많은 연구자들은 심층적인 연구를 통해 영향력이 매우 큰 논문을 발표하려고 하기 보다는 조그만 결과들을 쪼개어 보다 많은 수의 논문을 만들어 내려고 할 것이다. 반면에, 논문의 편수 보다는 논문의 영향력을 보다 중요시하는 경우 연구자들은 보다 심도 있는 연구를 통해 훌륭한 논문을 발표하려고 노력할 것이다. 이처럼 평가제도는 단순히 우수한 연구과제선 선정하여 연구비를 지원하는 것에 그치지 않고 이에 관련된 많은 효과를 수반하게 된다.

따라서 정량지표를 도입하고자 하는 경우 우선적으로 다음과 같은 몇 가지 원칙이 정립될 필요가 있다. 첫째, 우수한 논문과 원천특허를 모두 반영하도록 하여야 한다. 최근 원천특허로 등록된 창의적 연구 개념이 기업의 기술경쟁력 나아가 국가의 경쟁력에 큰 영향을 미치므로 우수한 논문뿐만 아니라 우수한 특허도 평가에 반영함으로써 연구자들이 경제사회적으로 기여할 수 있도록 유도하는 것이 바람직하기 때문이다. 둘째, 경제사회적으로 큰 영향력(impact)을 미칠 수 있는 연구가 수행되도록 유도해야 한다. 연구자들은 대체로 연구지원

10) 수학 분야의 경우에는 긴 논문일수록 일반적으로 가치가 높고 게재가 어렵기 때문에 논문의 페이지수가 평균보다 많을 경우 상향 평가하기도 한다.

기관의 평가기준에 매우 민감하게 반응하게 된다. 따라서 평가기준이 영향력이 적은 논문은 논문편수가 많아도 정량지표에 큰 영향을 미치지 못하고 반대로 영향력이 큰 논문은 논문편수가 적어도 정량지표에 큰 영향을 미치도록 하는 경우 연구자들은 비록 편수가 적어도 우수한 논문을 생산하고자 할 것이기 때문이다. 셋째, 학문분야에 따른 특성이 평가에 혼란을 초래하지 않도록 해야 한다. 연구자들의 연구활동은 학문 분야별로 많은 차이점이 있기 때문에 이러한 차이점을 반영해야 한다. 예를 들어, 논문이 상대적으로 많은 생명과학분야와 논문이 상대적으로 적은 이론수학분야에 대해 일률적으로 impact factor를 기준으로 하면 생명과학분야는 상대적으로 우위를 점하게 되는 것 등이다. 넷째, 논문에서의 역할을 구분해야 한다. 논문에서의 역할은 주저자와 공동저자로 구분할 수 있는데, 공동저자로서 많은 논문실적을 올리는 경우에는 그 실적을 제한적으로 인정할 필요가 있다. 이는 아직도 고참연구자가 논문에 대한 기여 없이 이름을 올려 무임승차하는 경우가 간혹 있기 때문이다.

이러한 정량지표 도입의 기본방향에 부합되면서 동료평가를 보완할 수 있고, 또한 연구자에게 긍정적인 동기부여를 할 수 있는 정량지표의 대안은 다음과 같다.

첫째, 동료평가와 정량지표의 점수 배분비율로 각각 70%와 30%를 제안한다. 정량지표는 동료평가에 대한 보조장치로 활용되는 것이 바람직하기 때문이다.(Aksnes and Taxt, 2004) 그러나 보조장치로 활용되는 경우에도 정량지표가 평가에 어느 정도 영향을 미쳐야 하므로 30%수준이 적절한 것으로 판단된다.

둘째, 과거 일정기간 동안 발표된 논문 중 일정 숫자만을 대상으로 하는 것이 바람직하다. 예를 들어, 과거 5년간 가장 우수한 논문 5편 또는 과거 3년간 가장 우수한 논문 3편 등으로 하는 것이다. 이는 과거 일정기간 동안의 모든 논문을 모두 다 정량지표 대상으로 하는 경우 아주 우수한 논문 한편을 쓰기보다는 중간 등급의 논문 여러 편을 쓰는 것이 평가에 유리하게 작용할 수 있기 때문이다. 심사대상 논문 편수의 제한은 연구자들로 하여금 우수한 논문을 쓰도록 유도할 것이며, 이러한 연구 성향은 궁극적으로 미래의 원천기술을 확보하는데 결정적으로 기여하게 될 것이다.<sup>11)</sup>

셋째, 각 논문이 게재된 저널의 impact factor를 그대로 사용하지 말고, 해당 분야 저널들의 impact factor순위를 백분율로 적용하여 이에 따른 가중치를 적용하는 방식이 바람직하다. 해

11) 하준경(2004)은 실증분석 결과 우리나라의 R&D 생산성이 미국 등에 비해 낮은 것으로 나타나고 있는데, 그 원인이 R&D의 질적인 측면에 기인하는 것으로 파악하면서, R&D의 질적 수준 제고를 권고하고 있다.

당 분야 내에서 저널의 impact factor를 기준으로 순위를 부여한 후 이 순위를 기준으로 논문에 대한 점수를 부여하면 분야별 impact factor의 차이에 따른 분야간의 문제를 해소할 수 있기 때문이다. 분야 내에서 저널의 순위에 따른 점수의 배분(안)은 다음과 같다.

해당분야내의저널순위	상위10%	상위20%	상위40%	상위70%	나머지
점수부여	5점	4점	3점	2점	1점

넷째, 논문의 공동저자(co-author)인 경우의 점수 합계가 논문의 주저자(first or corresponding author)인 경우의 점수합계의 50%를 넘을 수 없도록 제한하는 것이 바람직하다. 예를 들어 공동저자로 등록된 경우 논문의 1/2을 실적으로 인정한다고 할 때 논문이 40편인 경우 20편의 실적이 인정되지만, 주저자로 등록된 논문이 4편이라면 4편의 50%인 2편의 범위 내에서 공공연구 논문실적을 인정하는 방식이다. 이 방식은 공동저자인 경우에도 논문실적을 인정해 주는 대신, 자신이 주저자로 저술된 논문 실적에 비례하여 공동저자 논문 실적을 인정해 주는 바람직한 방법이다.

이와 같이 새로운 정량지표를 도입하는 경우, 연구자가 심층적 연구를 하도록 촉진함과 동시에 주관적 가치 판단에 의존하는 동료평가의 편기를 보정하는 보완적 효과도 기대할 수 있을 것이다. 가장 바람직한 평가는 어느 정도 통일된 판단기준을 갖는 평가자들이 사심 없이 평가를 수행하는 것이지만, 평가자 풀(pool)이 상대적으로 적은 우리나라의 경우에는 평가자와 피평가자 간에 학연, 지연 등 주관적 판단을 왜곡시킬 수 있는 요소를 완전히 배제하는 것은 거의 불가능한 것으로 판단된다. 따라서 주어진 현실에서 보다 객관적이고 공정한 평가를 수행하기 위해서는 동료평가를 기본으로 하되, 정량지표를 통해 동료평가가 초래할 수 있는 문제를 보완하는 것이 가장 바람직한 방법이 될 것이다.

#### IV. 요약 및 결론

연구개발활동에 관한 평가를 수행할 때, 가장 많이 사용되어 온 방법이 동료평가(peer review)방법이다. 동료평가 방법은 질적인 평가를 수행할 수 있는 매우 이상적인 평가방법임에도 불구하고 주관적 판단에 전적으로 기인해야 하는 문제 때문에 일부 문제점들을 가지고 있기도 하다. 여기서는 이러한 동료평가의 문제점을 보완하는 방안으로서 정량지표

(bibliometric indicators)를 도입하는 방안에 대해 살펴보았다. 정량지표를 이용한 평가방식은 동료평가와 달리 객관적이고 수량적으로 평가를 수행하는 기법이다. 물론, 정량지표를 이용하는 경우에도 질적인 측면이 반영될 수 있도록 인용횟수, 저널의 impact factor 등을 활용하지만, 연구의 질적인 측면을 완전히 반영하기에는 한계가 있을 수밖에 없다. 따라서, 동료평가를 기본으로 하되 정량적 평가지표를 도입하여 동료평가를 보완하는 방법이 적극적으로 검토되고 도입될 필요가 있다. 아울러, 정량지표를 도입하는 경우에도 연구자들에게 긍정적인 동기를 부여할 수 있도록 평가지표를 설계해야 할 것이다.

여기서는 이러한 정량지표로서 첫째, 동료평가와 정량지표의 점수 배분비율로 각각 70%와 30%로 하고 둘째, 과거 5년간 가장 우수한 논문 10편만을 대상으로 정량지표를 작성하며 셋째, 각 논문이 게재된 저널이 속한 분야 저널들의 impact factor순위를 백분율로 적용하여 이에 따른 가중치를 적용하는 방식이 바람직하며 넷째, 논문에서의 저자 역할에 따라 점수를 차등화하고 공동저자의 논문 실적이 주저자의 논문실적에 의해 결정되도록 하는 방식을 제안하고 있다. 이와 같은 정량적 평가지표의 도입은 그동안 정량적 평가지표가 갖는 문제점들을 상당 부분 극복할 수 있을 뿐만 아니라, 국가차원의 원천기술 확보에도 크게 기여할 것으로 생각된다.

아울러, 대안으로 제시한 정량지표에는 포함되어있지 않지만 인용횟수는 논문의 질적 수준을 반영할 수 있는 중요한 지표중의 하나이므로 향후 인용횟수를 활용하는 정량지표에 대해서도 적극적인 방안을 모색하는 것이 필요하다고 생각된다.

## References

- 김인목 외, 「기초과학연구지원사업별 평가전략 정립을 위한 정책조사연구」, 한국과학재단 2002.
- 송충한, “기초연구의 성과평가 기법”, 한국기술혁신학회 춘계학술발표대회. 1999.
- 하준경, 「R&D와 경제성장: 논쟁과 우리나라에 대한 시사점」, 금융경제연구원, 2004.
- Aksnes, W. Dag, and Randi Elisabeth Taxt, “Peer review and bibliometric indicators: a comparative study at Norwegian university”, *Research Evaluation*, vol. 13, no. 1, April 2004.
- Chubin, DE, “Grants Peer-Review in Theory and Practice”, *Evaluation Review*, Vol 18. 1994
- Committee on Science, Engineering, and Public Policy(the joint committee of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine), *Evaluating Federal Research Programs: Research and the Government Performance and the Results Act*, Washington D.C., Feb. 1999.
- CWTS, *Bibliometric Profiles of Academic Electrical and Electronic Engineering Research in the Netherlands(1989-1998)*, 2000.
- General Accounting Office(GAO), *Measuring Performance: Strengths and Limitations of Research Indicators*, Washington D.C., May, 1997.
- General Accounting Office(GAO), *Peer Review Practices at Federal Science Agencies Vary*, Mar. 1999.
- HEFCs, *A Guide to the 2001 Research Assessment Exercise*, 1998
- Kostoff, Ronal N., “Federal research impact assessment: Axioms, approaches, applications”, *Scientometrics* 34(2) 1995, pp163-205
- Kostoff, Ronald N., *The Handbook of Research Impact Assessment*, Office of Naval Research, 1997
- Leeuwen, T. N. van, L. J. van der Wurff and A. F. J. van Raan, “The use of combined bibliometric methods in research funding policy”, *Research Evaluation*, 10(3), 2001. pp195-201.
- Lewis, Grant, Robert Cottrell and Diane Dixon, “Bibliometric indicators to assist the peer review process in grant decisions”, *Research Evaluation*, 8(1), 1999. pp 47-52.
- Luukkonen, T., “Citation indicators and peer review: their time-scales, criteria of evaluation, and

biases", *Research Evaluation*, 1(1) 1991.

Ndedrhof, A. J., "The validity and reliability of evaluation of scholarly performance", in A. F. J. van Raan(editor), *Handbook of Quantitative Studies of Science and Technology*, NorthHolland, Amsterdam, 1988.

OECD, "Bibliometric indicators and analysis of research systems; methods and examples", STI working papers 1997/1, 1997.

Oppenheim, C., "The correlation between citation counts and the 1992 research assessment exercise ratings for British research in genetics, anatomy and anachaeology", *Journal of Documentation*, 53(5), 1997 pp477-487

Prichard, A. "Statistical bibliography or bibliometrics", *Journal of documentation*", 1969, p358-359

Rinia, E. J., T. N. van Leeuwen, H. G. van Vuren and A. F. J van Raan, "Comparative analysis of a set of bibliometric indicators and central peer review criteria: evaluation of condensed mater physics in the Nethelands", *Research Policy*, 27, 1998, pp 95-107