

Title	研究成果の「卓越性」指標の多様性
Author(s)	林, 隆之; 土屋, 俊
Citation	年次学術大会講演要旨集, 29: 798-801
Issue Date	2014-10-18
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/12565
Rights	本著作物は研究・技術計画学会の許可のもとに掲載するものです。This material is posted here with permission of the Japan Society for Science Policy and Research Management.
Description	一般講演要旨

研究成果の「卓越性」指標の多様性

○林 隆之、土屋 俊（大学評価・学位授与機構）

1. はじめに

大学の研究評価において、研究業績評価に係る指標の設定やデータ活用に変化が生じている。大学の研究評価は、英国、仏国、オランダ、豪州など様々な国で行われ(Geuna & Martin, 2003; Hicks, 2012)、その多くはピアレビューを主たる方法としつつも、「informed peer review」と称されるように、評価対象の基礎的なデータや指標が評価者へ提供される。中でも近年、個別研究業績についてビブリオメトリクスデータをよりシステムティックに評価者や大学側へ提供することが英国や豪州などで行われるようになった。同時にビブリオメトリクスの利便性が向上しているがゆえに、偏重利用されることや専門知識なしに不適切に利用されることに対して懸念も増している(e.g. DORA 2012, Canadian Academies Council 2012, Cronin & Sugimoto 2014)。さらに並行して、学術面以外の社会・経済・文化面でのインパクトも評価で重視されるようになり(Lutz Bornmann, 2013; Martin, 2011)、研究業績を評価する多面的な指標群を模索する動きがみられるようになってきている(e.g. HEFCE 2014)。

日本に目を移せば、国立大学法人評価においても研究成果の評価は行われている。学部・研究科を代表する卓越した研究業績を大学が提出し、学術面だけでなく社会・経済・文化面の効果をピアレビューアが段階判定する。日本の特徴は、学部・研究科自身が各研究業績が卓越していると考えられる理由を指標・データを根拠に引きながら説明する必要があることである。前回の評価においても、引用数などのビブリオメトリクスデータを用いる事例は多く見られた。前回の評価実施以降、日本ではURAの専門職化などによりビブリオメトリクスデータの利用は急速に一般化している。そのような流れの中で、日本の大学評価においてもよりシステムティックな形でビブリオメトリクスデータを利用する可能性を模索すると同時に、ビブリオメトリクス以外の考えうる指標群を明示していくことが必要となっている。それによりビブリオメトリクス指標に適合しない研究業績が正当に評価されず、研究者の研究活動や成果発表方法に悪影響をもたらす恐れを回避しなければならない。

以上を背景に本稿では、次の問いを扱う。日本でも指標利用が進む可能性を踏まえて、1) 指標の一つであるビブリオメトリクスデータが英語圏でない日本において現実の評価場面でどれほど使える可能性があるか、2) それらのデータがピアレビュー結果と一致し、レビューを支援できる可能性があるか、3) ビブリオメトリクスデータ以外の学術面ならびに社会・経済・文化面の卓越性の指標としてどのようなものがあるかを考えることができるか。これらの問いを通じて、研究業績の卓越性の代理指標(proxy indicator)の多様化の必要性を実証的に確認する。

2. 分析対象データ：大学の学部・研究科を代表する研究業績群

本稿で分析対象とするのは、日本の第一期国立大学法人評価（2008年度実施）において全国立大学の学部・研究科から提出された研究業績説明書である。研究業績説明書は、学部・研究科を代表する業績（教員数の半数を提出上限）と、大学全体の重点的研究領域における研究業績の2種類が提出され、あわせて19,626件であった。また、日本の結果を解釈するための対照として、英国における研究評価であるRAEの前回実施分（RAE2008）において大学から提出された研究業績リスト214,287件を用いる。英国では研究実施教員(Active Research Staff)一人あたり4件を上限に提出するため、日本の10倍以上の業績が提出されている。

3. ビブリオメトリクス指標の適用可能範囲

まず、大学自身が評価に提出するような優れた研究業績について、そもそもジャーナル論文の割合はどの程度であるのか、その中でどの程度が論文DBの一つであるトムソン社のWeb of Science (WoS) に収録されていたかを分析する。分析は次の方法で行った。日本での研究業績の提出書類においては、業績種類の多様性を考慮して書誌情報などの記載様式を画一的には定めなかったため、著者・論文名・ジ

ジャーナル名などの記載は多様であり、標記揺れも多い。そのため、記載された文章から単純に WoS を検索することはできない。そのため、次の 2 つの方法による結果の和集合を検索結果とした。一つは、標記方式が不定でも対応可能である Crossref の DOI 検索を用いて DOI を検索し、その DOI により WoS を検索した。もう一つは、WoS 中の全論文と研究業績リストの文字列を対照して最も類似した論文を選ぶ方法である。各研究業績ごとに、WoS 中の論文で出版年・巻・ページ数の数字が研究業績文字列に含まれ、かつ、論文タイトルと研究業績の部分文字列との編集距離(Levenshtein Distance)指標が最も高い(最も類似した) WoS 論文を選び、目視で適合性を確認した。

英国 RAE2008 の提出論文リストは書誌情報が画一的な様式で記入され、一部の業績には DOI も記載されている。英国でも次の二つの結果の和集合を検索結果とした。一つは DOI による検索、もう一つは出版年・巻・ページ数が一致し、かつ、ジャーナル名の一致あるいはタイトルの編集距離に近い論文を選定した結果である。

以上の方法で検索した結果、全分野を集計すれば、表 1 のような結果となった。まず、日英ともに提出業績のうちで、論文データベース収録の主たる対象となりうるジャーナル論文に相当する業績は 7~8 割であった。さらに、WoS

に収録されていた割合は、全提出業績の 65%程度にすぎず、論文中では 80%前後であった。また、図 1、2 では分野別の結果を示している。提出された研究業績に占める各分野の割合を円グラフで示し、各分野の中で WoS に収録されていた割合を面積比で黒く塗っている(より詳細な分科レベルのデータは紙面制約上、割愛する)。

表 1 日英の全提出業績中の論文割合・WoS 収録割合

	提出業績数 (A)	うち、論文 「論文」(日本)あるいは「Journal article」(英国)カテゴリーの業績数 (B)		WoS 収録数 (C)	研究業績全数での WoS 収録割合 C/A	論文中の WoS 収録割合 C'/B
					C/A	C'/B
日本	19,626	16,158	82.3%	12,770	65.1%	78.8%
英国	214,287	161,261	75.3%	131,485	61.4%	81.5%

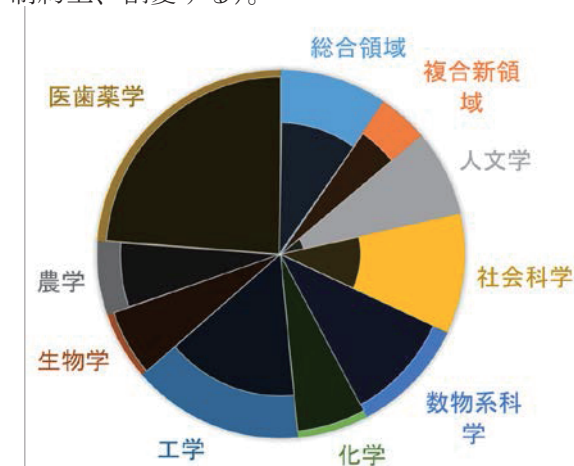


図 1 日本の国立大学法人評価での提出研究業績の分野割合と WoS 収録状況

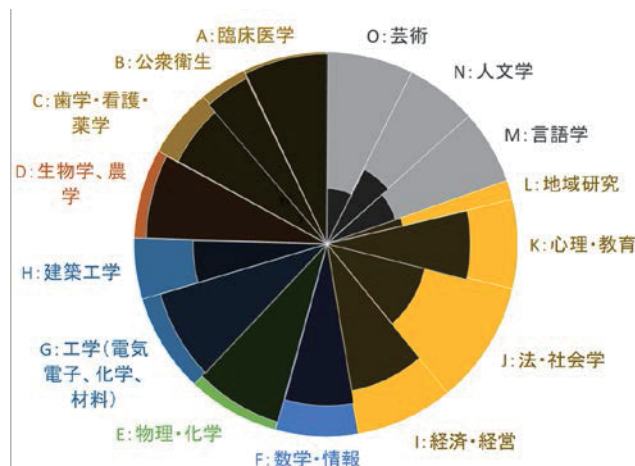


図 2 英国 RAE2008 での提出研究業績の分野割合と WoS 収録状況

ビブリオメトリクス利用が工学、社会科学、人文学では適切でないことは一般的にしばしば指摘されることであるが、本分析結果からは、大学を代表するような優れた業績に関しても限界があることが実証的に確認された。実際には詳細な分科レベルでの違いも大きく、業績評価においては細目単位での利用の必要性を検討することが望まれる。同時に、大学単位での分析をする場合には、分析対象となっているのは大学の研究活動のうちで、円グラフの左側に極端に偏った部分でしかないことを改めて認識する必要があると示唆された。さらに日・英間で比較すれば、そもそも提出業績の分野構成に大きな差異がみられる。英国では人文学・社会科学に相当する研究業績の提出がおおよそ半数を占める。人文学・社会科学の業績の WoS 収録率は英国のほうが高いが、全分野を集計すれば、研究業績のうちの「論文」の割合ならびに WoS 収録割合は日本のほうが高くなるという構造がある。このことは日本国内における人文・社会分野の量的な弱さを示唆している。

3. 3 ピアレビューとの関係

では WoS に収録されている研究業績について、ビブリオメトリクス指標はピアレビューによる評価結果と相関するのか。両者の関係の分析は既にいくつか存在し(林 2003 やそこでの引用文献を参照)、ある程度の

整合性を認めている。しかし、本稿の分析対象とするのは、大学を代表する優れた業績の中でもさらにそれを「卓越」「優秀」「それ以外」と3区分したものである(表2)。このような上位層の中での判断における整合性は明らかでない。

本分析では、WoS の 250 程度の分野分類(subject category)、ならびに ESI の分野分類を用いて、各論文が引用数でみて、該当する分野分類の同一出版年の論文(Article, Letter, Review)の中で上位何%に位置するかというパーセンタイル指標を用いた。複数の分野分類がふられている場合には、もっともパーセンタイルの値が良い(値が小さい)

分野を用いた。また、評価実施年は 2008 年であったために、評価での利用可能性の検証のために、2008 年までの引用数と、その後 3 年が経過した 2011 年までの引用の 2 つを用いて差異を確認した。

結果では、全ての論文を合計すれば、SS と判断された業績の中央値は WoS 分類を用いれば上位 11% 程度、ESI 分類を用いれば 8% であった。引用期間によって中央値に大きな差異はない。

分散も考慮し、WoS への収録割合が高い分野のみを取り出して 3 群での差のノンパラメトリック検定(クラスカル・ウォリス検定)を行った。

結果、SS と S とでは差異が見られるが、S と S 未満では統計的に有意な差異が見られない分科がみられた。すなわち、ピアレビューで極めて卓越しているとされる業績については引用数でも差が付きやすいが、それ以下での区分では整合性が十分ではない分科もある。

また、個別業績レベルでみれば、図 3 に示すように、たとえ引用数が上位 10%に入る場合でも、SS, S のどちらに判断されたかは確定的でない。全体的に半数程度が判断の中間である S となり、引用数の高低によって、SS や S 未満に判定される確率変動している。これまでも、ビブリオメトリクスデータは機関全体などの集計値には有効であるが、個別の論文や研究者を対象とする場合には不十分であることは指摘されてきた。本分析からも、個別論文単位ではビブリオメトリクスの判断がピアレビューを単純に代替できる可能性は低いことが示唆される。他方で、レビューアーの判断も完全ではないことを前提とすれば、複数の指標群のもとで判断

表 2 段階判定の定義

	学術面	社会、経済、文化面
SS (卓越)	「当該分野において、卓越した水準にある」研究業績の独創性、新規性、発展性、有用性、他分野への貢献などの点において、客観的指標等から判断して、当該分野で学術的に最も優れた研究の一つであると認められ、当該分野ないし関連する分野において極めて重要な影響をもたらしている水準	「社会、経済、文化への貢献が卓越している」以下の領域(略)において、客観的指標等から判断して、極めて重要な影響や幅広い影響をもたらしている水準
S (優秀)	「当該分野において、優秀な水準にある」SS にまでは至らないが、当該分野で学術的に優れた研究の一つであると認められ、当該分野ないし関連する分野において重要な影響をもたらしている水準	「社会、経済、文化への貢献が優秀である」SS にまでは至らないが、重要な影響や幅広い影響をもたらしている水準
S 未満(A: 良好, B: 相応, C: それ未満)		

表 3 ピアレビューによる 3 段階判定ごとの引用数パーセンタイルの中央値

	SS 中央値	S 中央値	S 未満 中央値
WoS 分野によるパーセンタイル (2008 年まで)	11.6%	26.3%	38.5%
WoS 分野によるパーセンタイル (2011 年まで)	11.0%	24.4%	35.4%
ESI 分野によるパーセンタイル (2008 年まで)	8.7%	25.6%	40.6%
ESI 分野によるパーセンタイル (2011 年まで)	8.3%	23.6%	36.5%
掲載雑誌の IF によるパーセンタイル (2008 年)	4.8%	10.5%	14.9%

表 4 収録割合が高い分野でのピアレビュー結果との関係

No	分科	提出数	うち「学術」への提出	「学術」のうちの WoS 収録率	引用数のパーセンタイル			クラスカル・ウォリス検定		
					SS 中央値	S 中央値	S 未満中央値	判定結果による差	多重比較: SS-S 間に差があるか	多重比較: S-S 未満に差があるか
11	神経科学	348	334	98%	13.6%	20.8%	33.6%	**	**	**
23	ゲノム科学	70	66	92%	1.5%	17.2%	14.5%	**	**	-
24	生物分子科学	90	82	96%	9.5%	24.6%	33.9%	**	**	-
46	基礎化学	387	382	96%	17.3%	22.7%	39.1%	**	**	**
47	複合化学	567	534	94%	9.8%	16.7%	30.4%	**	**	**
58	生物科学	691	680	97%	9.3%	24.3%	41.5%	**	**	**
68	薬学	304	280	98%	8.0%	23.5%	30.1%	**	**	-
69	基礎医学	1401	1328	96%	5.7%	19.8%	32.9%	**	**	**
72	内科系臨床医学	1453	1355	94%	11.0%	21.0%	32.6%	**	**	**
73	外科系臨床医学	609	564	97%	9.2%	25.4%	34.7%	**	**	**
74	歯学	519	482	95%	10.9%	30.8%	32.2%	**	**	-

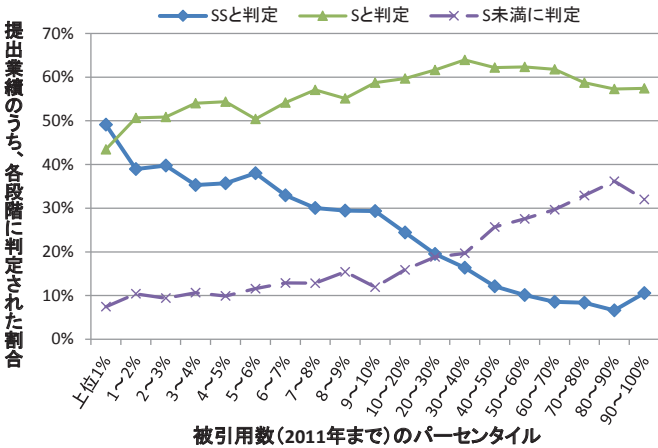


図 3 引用数による各判定の確率

ができる環境構築が望まれる。

4. 卓越性を示す多様な代理指標

上記まででビブリオメトリクスについて検討をすすめた。実際の

国立大学法人評価では、大学側（教員）が、各研究業績がなぜ卓越していると考えられるかを、根拠データ等をもとに示し、レビューアーは最上位の SS の判断を下す場合には、判断材料に用いた根拠データをマーキングすることが求められた。このマーキングされた根拠データが各分野で教員・評価者が重要と考える代理指標の種類を示すと考えられる。

表 5 ではその中で数が多かったものを示している。引用数は数物系、IF は医歯薬や生物学で多く見られる。他方、人文学では一般紙等での書評（件数ではなく、書評内容の文面）、社会科学では査読のある雑誌への掲載自体、工学では招待講演、生物学では Faculty of 1000 や解説記事などの発表後の研究者間での評価を踏まえた指標が特徴的に現れている。ただし、これは集計値が多かった事項にすぎず、たとえば表 6 に示す工学の例では様々な根拠データが示されている。これらを集積して、各分野に適切な根拠データを研究者間や評価者間で合意し、多様な指標を示しやすくしていく環境形成が望まれる。

5. おわりに

ビブリオメトリクスの限界はしばしば指摘されるが、実際に利用する場面では限界は無視されることが多い。本分析では現実の評価での利用可能性と、ビブリオメトリクス偏重を抑える方策を検討したものである。現在、分野ごとの検討会を設けて、研究業績の指標のみならず、研究活動さらには教育活動・成果についての分野特有の視点についての検討を開始している。各分野の研究・教育活動や成果の価値を研究者があらためて省察する機会として、評価が活用されることが今後望まれる。

主な参考文献

- 林隆之(2003)「ビブリオメトリクスによるピアレビューの支援可能性の検討—理学系研究評価の事例分析から」『大学評価』Vol.3, pp.167-187.
- Bornmann, L. (2013). What is societal impact of research and how can it be assessed? a literature survey. *Journal of the American Society for Information Science and Technology*, 64(2), 217–233.
- Martin, B. R. (2011). The Research Excellence Framework and the “impact agenda”: are we creating a Frankenstein monster? *Research Evaluation*, 20(3), 247–254.
- Cronin, B. and Sugimoto, C. R. eds. (2014), *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*, The MIT Press

表 5 「卓越性」の主な代理指標の分野間の違い

	引用			掲載雑誌 のIF			受賞			新聞・一 般紙で の書評・紹 介			学術誌・ 専門書で の書評・ 紹介・書 評			掲載雑誌 名			査読のある 雑誌への 掲載			招待講 演・基調 講演			Faculty of 1000			他研究者 による解説 記事: nature誌 News and Views, etc.			特許化			製品化・ 実用化			研究を 行った研究 費																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
		うち引用 数																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										

表 6 工学分野において提出された根拠データの例

学術面	- 研究成果に基づく、学術面での受賞 - 新聞、一般雑誌、業界誌、テレビでの研究成果の紹介・批評 - 学術誌や専門書での研究成果の紹介・批評 - 著名な学術雑誌への掲載、学術雑誌のインパクトファクター - 被引用数。高被引用論文への選出 - 著名な論文や講演、レビュー論文、教科書・辞典等における研究成果の引用・紹介とその扱われ方 - 論文のアクセス数やダウンロード数。それらの値が高い論文への選出 - 掲載論文における注目論文や優秀論文としての選出 - 著名な学術雑誌における研究動向解説論文・記事などによる解説 - 招待講演、基調講演 - 著名な学会や採択が厳しい学会における発表の選定。競争性の高い選定（ポストデッドライン論文など） - 再録雑誌への採択 - 研究活動のための競争的研究費。研究成果に基づいて新たに獲得した競争的研究費 - 研究費による事後評価の結果
社会・経済・文化面	- 社会・経済・文化面を重視した受賞（地方自治体、産業界などからの受賞） - 新聞、一般雑誌、業界誌、テレビでの紹介・批評 - 研究成果物の展示会やその来場者数 - 国内および国際特許化。ライセンス契約やその収入 - ソフトウェア、データ、装置・研究試料の開発・公開、利用状況や利用者の成果 - 研究成果に基づく起業 - 国際標準への選定、政府・産業団体等でのロードマップにおける選定 - 製品化・実用化、それによる企業の売上高や期待される市場規模 - 書籍の出版と出版部数 - 研究成果の教材としての利用状況 - 企業等との共同研究の状況や、その後の共同研究の申し出状況 - 社会・経済・文化面への貢献を重視した研究費の獲得 - 政策や規制・ガイドライン等への貢献 - 政府や地方自治体などにおける委員の就任とそこでの研究成果の反映 - 公共サービスでの研究成果の活用 - 医用工学などにおける臨床応用への展開や利用状況 - 研究成果やそれに基づく製品の利用者における、環境・エネルギー面の効果