

特許・論文間引用に関するジャーナル評価指標の検討

— Journal Commercial Impact Factor の性質 —

Journal evaluation on patent-paper citations

国立研究開発法人科学技術振興機構

山下 泰弘

独立行政法人産業技術総合研究所技術と社会研究センター、山形大学評価分析室等を経て現職。計量書誌学分析に従事。

✉ yasuhiko.yamashita@jst.go.jp

国立研究開発法人科学技術振興機構

吉田 秀紀

株式会社東芝研究開発センターを経て、現在 JST でファンディング業務や分析業務に従事。

1 はじめに

研究課題は、基礎的なものから、実用化を目途とする応用的なものまで、連続するフェーズの中に位置づけられると考えられる。将来の技術応用を見込んだファンディングを行う上では、研究課題(あるいはトピック)が、技術応用までに、どのくらいの距離が残されているのかを把握することが重要である。

筆者の一人(山下)は、技術応用に近い研究課題を識別する目的で、論文のリファレンスに注目した分析を行った(Yamashita, 2020)。具体的には、論文に含まれる「特許から引用されたリファレンス」、「企業著者が含まれるリファレンス」を集計し、それらが多い論文ほど将来特許から引用される傾向が高いことを見出した。これらの方法は、特許・論文間の引用リンクデータがあれば任意の最新論文に対して適用可能であるが、マクロな状況を把握するケースにおいてはデータ量や計算量が嵩むことは否めない。ファンディング現場の要求に応える、適時適切なエビデンスを提供するためには、多少ラフではあっても、直観的にわかりやすく、より安定した指標が望ましい。また、比較的若い論文に適用されるケースが多いため、分析対象論文(群)自体が特許か

ら得た引用を前提としない指標である必要がある。ジャーナル単位の指標は、個別論文の特徴量とは独立に扱うことができるため、この面において好ましい。

Bikard と Marx (2020) は、クラリベイト社の Journal Citation Reports で提供されている Journal Impact Factor (JIF) を模して、Journal Commercial Impact Factor (JCIF) を提案した。JCIF は、JIF の計算における分子を論文間引用から特許・論文間引用に置き換えたもので、広く認知されている指標をベースとしているため、直観的に受け入れやすいと考えられる。一般的に、特許からの論文引用は、論文間引用や JIF と高い相関を示すことが知られているが、Bikard と Marx は、JIF と JCIF を独立変数の一部とする回帰モデルにおいて、産業界の特許からの引用と、前者は正、後者は負の相関関係となることを示している。しかしながら、JCIF は提案されて間もないため、実際に分析に使用した場合の挙動については不明な点が多い。本稿では、JCIF あるいはそれをベースとした指標の利活用に向けた、JST 経営企画部エビデンス分析室で進めている検討の内容について報告する。

2 Journal Commercial Impact Factor (JCIF)

前述の通り、JCIFはJIFの特許版ともいえる指標である。ジャーナル*j*の*y*年のJIFは、(1)式のように定義される。ここで、 $N_{j,y-1}$ は*y*-1年にジャーナル*j*に掲載された論文の数、 $C_{j,y-1,y}$ は*y*-1年にジャーナル*j*に掲載された論文が*y*年に論文から得た引用数である。

$$JIF_{j,y} = \frac{C_{j,y-1,y} + C_{j,y-2,y}}{N_{j,y-1} + N_{j,y-2}} \quad (1)$$

JCIFは、JIFの分子を企業の特許からの引用数に置換したものである。すなわち、JCIFは、測定年の前年と前々年の2年間にあるジャーナルで発行された論文が、測定年に企業の特許から得た引用数の論文1報あたり平均値である。おそらく産業応用可能性に焦点を当てたため、大学や政府機関の特許からの引用については、計算から除外されている。ジャーナル*j*の*y*年のJCIFは(2)式で表される。

$$JCIF_{j,y} = \frac{C'_{j,y-1,y} + C'_{j,y-2,y}}{N_{j,y-1} + N_{j,y-2}} \quad (2)$$

$C'_{j,y-1,y}$ はジャーナル*j*の*y*-1年の論文が*y*年に出願された企業の特許から得た引用数である。JCIFは、JIFを模して作られているが、特許の出願年ベースで計測されている点において、JIFと大きく異なる。JIFの分子は、分母に含まれる論文が、計測年に出版された論文から得た引用数であり、計測年より後であれば、いつ同じ年のJIFを計測しても同じ値となる。それに対して、JCIFの分子の値は、分子を構成する特許の登録が進むにつれて増加することになる。

3 Reliance on Science (RoS) データ

Marxらは、Microsoft Academic Graph (MAG)と米国特許データより独自に構築した非特許文献引用データ（以下ではMarxらのウェブサイト名称と論文タイトル“Reliance on Science”より“RoSデータ”と呼ぶ）を用いて、分析を行うとともに、同データを広く研究者に提供している（Marx & Fuegi, 2020）。特許文献において、非特許文献の記載はフリーフォーマットであり、機械的に対応する文献を一意に同定できる

ケースから、人間が見ても対応する文献を同定できないケースまである。Marxらは、識別したリンクの確からしさに応じて10段階の評点（confscore）を付している。JCIFの計算にはconfscoreが5以上の引用リンクを含めている。

RoSデータのように構築プロセスや精度評価結果が明瞭な非特許文献引用データは非常に貴重であり、本稿でもRoSデータを活用することとする。Marxらは、JCIFについてもRoSデータのver.15以降で提供している。非特許文献引用リンクデータについては、適宜更新されており、構築方法も常にブラッシュアップされているが、JCIFについては、ver.15での発表以降の更新はなく、ver.32においても同一データが提供されている。したがって、より新しい年の値や、ブラッシュアップされた引用データに基づく値が必要な場合は、Marxらが公開しているプログラムを使用するなどして計算を行う形となる。

4 JCIFの再計算

前述のようにJCIFの値はMarxらが公開しているが、JCIFの測定条件による値の変動を検証するために、JCIFについては独自に計算を行った。米国では2000年に特許の公開制度が導入されたため、その前後で引用傾向が変化している可能性が高い。そこで、2000年以前の出願を含めない範囲で長期間の観測を行うため、2002年以降のJCIFを分析に使用することとした¹。2021年8月現在RoSデータの最新版はver.32であるが、Marxらと計算条件を合わせるため、本稿では彼らの使用したver.15を使用した。ただし、米国特許データについてはRoSデータに含まれていないため、2021年7月～8月にかけてPatentsViewのウェブサイト（<https://patentsview.org/>）からダウンロードしたより新しいデータを使用している。また、MAGでは収集するジャーナルに特に制約が設けられていないが、本稿では、安定して論文が発行されるジャーナルのみに限定する。すなわち、2000～2018年の期間に毎年1件以上の論文がRoSデータに含まれるジャーナル（12,525誌）が、以下本稿での分析対象となる。

1 2001年出願特許には、若干ではあるが、法改正以前の種類の公報（“A”）が含まれていたため、分析対象から外した。

特許からの論文引用は比較的まれな事象であり、単年で計測すると、偶発的な引用による年ごとの値の変動が大きい。例えば、2002年のJCIFにおけるトップジャーナルは Diagnostic Molecular Pathology 誌であるが、各年のJCIFを見ると、2002年に7.375と突出した値となっている以外、ほとんどの年で0.1以下に止まる(図1)。より安定した値を得るため、本稿では3年間(2002~2004)のJCIFの平均値を分析に使用する。また、特に観測時期について記載しない限り、JCIFの値は、使用するRoSデータに含まれる全期間の引用データに基づいた集計値(すなわち2018年末に観測される値)を指す。

JCIFについては、その特性について未知の部分が多いため、広く認知されているJIFと比較する形を取る。JIFについては、JCIFと異なり、観測時期による値の変動は発生しないことから、独自の値の算出は行わず、RoSデータで提供されている値を使用する²。

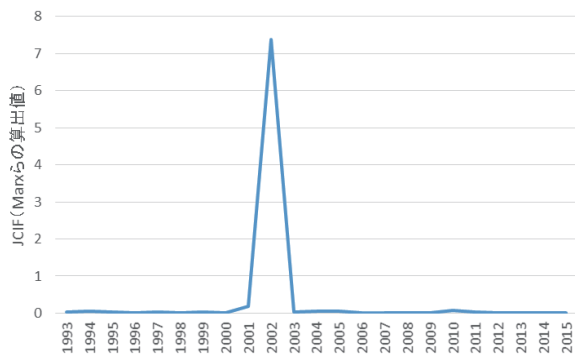


図1 Diagnostic Molecular Pathology 誌の各年のJCIF (Marx らによる算出値)

5 JIF と JCIF の比較

JCIFの値の分布は図2のようになる。比較のために2002年のJIF (Marx らによる算出値)と並置した。3年平均を取っているにも関わらず、JCIFの値が0というジャーナルが多く、全体の68.5%を占める。一方、2002年のJIFが0のジャーナルは10.8%に過ぎず、最頻のクラスは0.1超1以下である。

2 JIFの算出に必要なRoSデータ(ver.15)の論文間引用リンクテーブルについては、数度に渡りダウンロードを試みたが、その都度サーバエラーを生じたため、筆者らは入手できていない。7GB超の容量があるので、サーバに過大な負荷がかかったためと思われる。

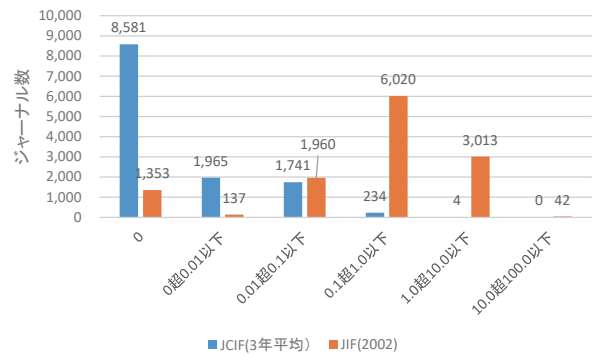


図2 JCIF (3年平均) と JIF (2002) の分布

JCIFが高いジャーナルは、企業の特許出願において相対的に重要な役割を果たすと考えられるが、内訳はどのようなものであろうか。JIFが学術面、JCIFが技術面でのインパクトを表すと考えると、両者を直交する2軸により、4種のジャーナルに分類することができる(図3)。図2に示したようにJIFとJCIFの分布が大幅に異なり、JCIFは指標値0が極端に多いこと、正であっても値が小さいものは、ジャーナルの性質というより偶発的な引用によると考えられること、などから、全体を性質が異なる4領域に区分することは難しい。そこで、中庸なジャーナルについては捨象し、それぞれの軸について突出して大きい/小さいジャーナルを例示することとする。ここでは、ジャーナル全体の上位1%(125誌)を指標値の高いグループ、指標値0を低いグループとみなした³。

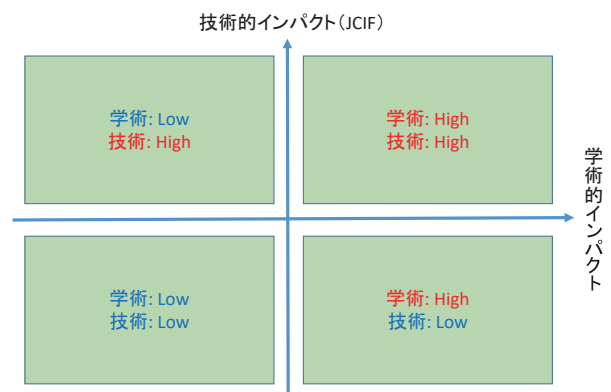


図3 学術・技術的インパクトによるジャーナル類型

各類型のジャーナルを表1に示す。学術・技術両面において高インパクトなジャーナルは41誌あり、その中

3 各象限で数十程度のサンプルが見込めるようトップ1%で区分したが、125位のJCIFは0.167、特許からの被引用数も3年間で2回に過ぎない。閾値については引き続き検討が必要と思われる。

には、Cell、PNAS、Scienceなどの著名誌が含まれる。学術・技術それぞれの観点から高インパクトなジャーナルの約1/3が相互に重なっていることから、少なくとも

も高インパクトなジャーナルに限って言えば2つの軸にある程度の相関関係があるものと考えられる。JCIFが高く、JIFが0であるジャーナルは1誌(Journal

表1 JCIF及びJIFによるジャーナルの類型と内訳

		JIF (学術)	
		高 (Top1%)	低 (JIF=0)
JCIF (技術)	高 (Top1%)	Endocrine reviews Annual review of immunology Cell Journal of clinical investigation Surface science reports IEEE signal processing magazine Proceedings of the national academy of sciences of the United States of America Annual review of neuroscience Nucleic acids research Nature genetics Accounts of chemical research Quarterly reviews of biophysics Pharmacological reviews Annual review of pharmacology and toxicology Annual review of medicine Journal of experimental medicine Current opinion in biotechnology Cancer research Genes & development Current opinion in immunology Microbiology and molecular biology reviews Materials science & engineering r reports ACM transactions on computer systems IEEE transactions on pattern analysis and machine intelligence Physiological reviews Nature medicine Science Annual review of genomics and human genetics Nature immunology Chemical reviews Genome research The plant cell Annual review of biochemistry Molecular aspects of medicine IEEE ACM transactions on networking Plant journal Progress in materials science ACM transactions on computer human interaction Annual review of genetics Annual review of cell and developmental biology ACM transactions on internet technology	Journal of optoelectronics and advanced materials
	低 (JCIF=0)	Living reviews in relativity Annual review of ecology evolution and systematics Journal of economic literature Reviews of geophysics Psychological science in the public interest Review of financial studies Journal of economic growth Quarterly journal of economics Framed field experiments Annual review of astronomy and astrophysics Psychological review GSA today Psychological bulletin Biological reviews Annual review of psychology Economic policy	African historical review Agricultural information research Chinese studies in history International medical journal Journal of Korean powder metallurgy institute Pakistan veterinary journal Pattern recognition and image analysis Histoire urbaine Magyar allatorvosok lapja The Japanese journal of personality 等 1,333 誌

of optoelectronics and advanced materials) のみであった。しかしながら、Web of Science (WoS) で調査したところ、2002年の論文から、このジャーナルの2000～2001年の論文が引用されていることが確認された。したがって、WoSから算出すれば、JIFは0にはならない。この原因についてはまだ検証しきれていないが、RoSデータのソースとなっているMAGについて一部の引用リンクの欠落が報告されており (Visser, van Eck & Waltman, 2021)、元データの正確さについても調査する必要がある。

学術的インパクトが高く、技術的インパクトが低いジャーナルは16誌あり、経済学・心理学・地理学等の分野が多く含まれる。

4つ目の区分の学術・技術の両側面からインパクトが低いジャーナルは1,333誌が含まれる。全サンプルの10%超を占める膨大な数であり、すべてのジャーナルを精査することは困難であるため、表1には任意の10ジャーナルのみ挙げた。ジャーナル名のみから性質を判断することは難しいが、非英語圏の地域的なジャーナルが多く含まれている。JIFが0であるジャーナル1,353誌のうち20誌を除く98.5%がこの区分に含まれる。すなわち論文から引用されない論文は特許からも引用されていない。ただし、いくつかの論文についてGoogle Scholarで目視確認したところ、2002年には引用されていないが、長期的には引用されているものもあった。このようなジャーナルの評価方法については、別途検討を行う必要がある。

6 JCIFの観測時期による変動

JIFは観測年を以て完結する指標である。引用の取得範囲等を見直さない限り、ある年のあるジャーナルのJIFは観測年以降変動することはない。一方、JCIFは、分子の特許からの引用が漸増するので、観測時期によって値や順位が変動することになる。したがって、指標として使用する上で、どの程度の観測期間を確保すべきかを把握しておく必要がある。

前節のプロセスで独自に計算した3年平均JCIF上位30誌のJCIFが、計測時期(ジャーナルを引用する特許の登録年)によってどのように変化するかを示したものが、図4である。初期段階において値や順位の変動が激しいことがわかる。登録特許公報から引用情報を取得

しているため、引用取得にかなりの年数がかかっているケースもあり、30位までの内訳が安定したのは2015年で、2018年まで30誌内での順位の入替えが発生している。相対的に小さい被引用数に基づく指標であるため、特許権の存続期間中は散発的に顕現する引用により、微妙な順位の変動が発生することが予想される。

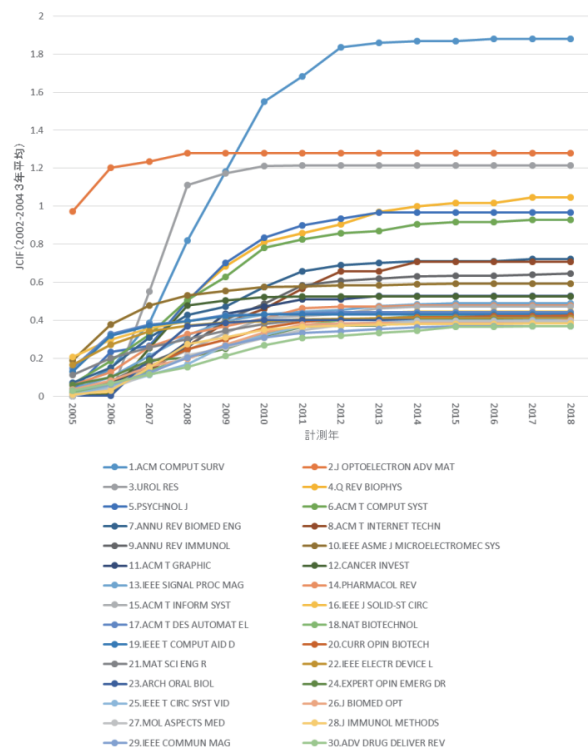


図4 上位30誌の観測時期によるJCIF(3年平均)の変化

一方、3年平均JCIFが0であるジャーナル数についてみると、2010年頃以降傾斜が緩やかになっている(図5)。そのことから、少なくとも、JCIFの観測時期として、6～7年程度は確保すべきと思われる。

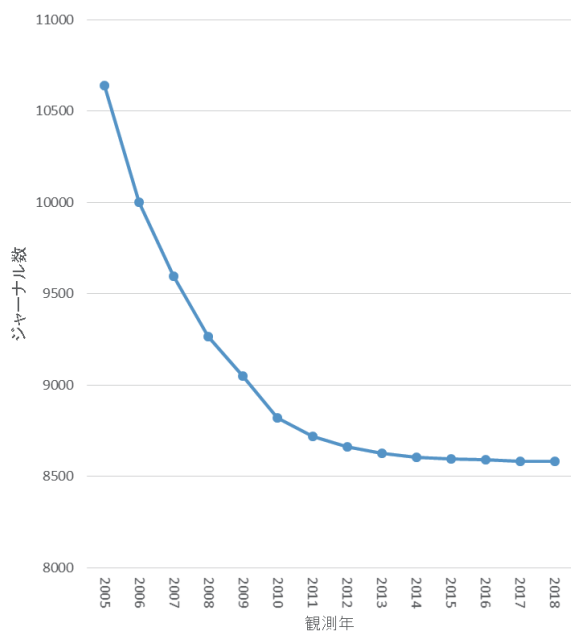


図5 観測時期による3年平均JCI = 0のジャーナル数の変化

7 まとめ

本稿では、BikardとMarxが提案したジャーナル指標JCIFの特性について、指標値の分布、JIFとの組み合わせでの類型化と内訳の精査、観測時期による挙動、の3つの点から検討を行った。本稿での検討内容はごくプリミティブなものであり、同指標の利活用可能性を明らかにするために、以下の点を引き続き検討する必要がある。

まず、第1の課題は、観測時期と収録論文の特許からの引用との相関関係の検証である。本稿では、JCIFが特許からの論文引用と相関するというBikardとMarxの報告を所与として、値や順位の安定性の側面から検討を行い、実際の相関性については踏み込まなかった。しかしながら、JCIFの値が安定するまで数年を要することを考慮すると、観測時期ごとに相関の強さも変化する可能性が高い。逆に、値が小さくても相関関係が不変であるならば、指標として有用性があるといえよう。本稿の段階での結論は、JCIFは直近年のものではなく、最低でも6年以上前のものを使用すべきということになるが、それだと最近勃興した分野の新しいジャーナルは評価対象となりえない。それは、比較的新しいプログラム/プロジェクトの成果分析や、最新の研究動向の把握といった、ファンディングの現場ニーズに応える分析

に使いにくいことを意味している。

第2の課題は、より多くの引用をJCIFの分子に含めることである。具体的には、引用ウインドウの長期化と取得する引用の種類の変換である。JCIFの引用ウインドウは、JIFに合わせて論文出版直後の1年間に出版された特許に限定されているが、統計的に意味のある数値を得るには、より長期間の出願を対象とする必要があると思われる。引用特許の出願年の分布についてさらなる検討を行い、最適な引用ウインドウの設定を行う必要がある。また、JCIFの算出に使用されているRoSデータver.15では、特許のフロントページからのみ引用が取得されているが、最近のバージョンでは特許本文に記載された引用までも含まれている。フロントページと本文での引用の意味の違いを考慮する必要があるが、JCIFの計算に含めた場合の挙動について検討の余地はあろう。

第3の課題は、比較対象のJIFの集計条件の見直しである。本稿ではJIFについては特に検討をせず、単年(2002年)の値を使用したが、JCIFに合わせて3年平均を取るなどにより、より安定した値を得て、さらなる分析を進める必要がある。

本稿では、学術・技術の2軸からジャーナルの類型化を試みた。このように区分することにより、ファンディングの現場では、例えば、ファンディングしたプロジェクトの成果の学術指向性と、産業技術との関連性を直観的にとらえることが可能になる。学術と技術の2軸での研究成果把握については、論文の被引用の観点から本誌で報告しているが(山下&治部, 2017)、即時的な現状把握が困難な点がプロジェクトのマネジメントへの適用上の障害であった。2軸モデルを適切なジャーナル評価指標と組み合わせることができれば、プロジェクトのモニタリングにも活用し得よう。

HuとRousseau(2018)は、Tijssen(2010)が論文著者の所属機関のセクタに基づいて定めた6つのジャーナルアプリケーションドメインを援用して、WoSの分野を発見志向分野と応用志向分野に分類し、引用の連鎖により、発見志向分野の研究で生産された知識が、応用志向分野を経て、技術分野(すなわち特許)に引用されるまでのプロセスを分析した。ジャーナルの分類を精緻化できるのであれば、本稿で用いた区分も、知識のフェーズ変化の分析に応用できる可能性はある



う。例えば、プロジェクト開始直後と終了時点で、プロジェクトメンバーの投稿先ジャーナルの比較をするなどの使い方も考えられる。

参考文献

Bikard, M & Marx, M. (2020). Bridging academia and industry: How geographic hubs connect university science and corporate technology. *Management Science*. 66: 3425-3443. <https://doi.org/10.1287/mnsc.2019.3385>

Hu, X. J. & Rousseau, R. (2018). A new approach to explore the knowledge transition path in the evolution of science & technology: From the biology of restriction enzymes to their application in biotechnology. *Journal of Informetrics*, 12: 842-857. <https://doi.org/10.1016/j.joi.2018.07.004>.

Marx, M. & Fuegi, A. (2020). Reliance on science: Worldwide front-page patent citations to scientific articles. *Strategic Management Journal*. 41: 1572-1594. <https://doi.org/10.1002/smj.3145>

Tijssen, R. J. W. (2010). Discarding the 'basic science/applied science' dichotomy: A knowledge utilization triangle classification system of research journals. *Journal of the American Society for Information Science and Technology*. 61: 1842-1852. <https://doi.org/10.1002/asi.21366>

Visser, M., van Eck, N. J. & Waltman, L. (2021). Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. *Quantitative Science Studies*. 2: 20-41. https://doi.org/10.1162/qss_a_00112

山下泰弘, 治部真里 (2017). 「特許からの論文引用に関する定量的評価指標の検討—日本の研究機関についての試行的分析—」. *Japio YEAR BOOK 2017*. 144-155.

Yamashita, Y. (2020). An attempt to identify technologically relevant papers based on their references. *Scientometrics*. 125: 1783-1800. <https://doi.org/10.1007/s11192-020-03673-5>



2

特許情報の高度な活用