HOKKAIDO UNIVERSITY

Title	論文評価指標の要点整理と動向紹介（特集 医学図書館員が知っておきたい最新情報）
Author(s)	千葉, 浩之
Description	91ページの「最新版のJIFを閲覧できる」の記述について、本稿執筆時点では最新の2018年版（2019年6月発表）が閲覧できていたが、2019年版（2020年6月発表）の発表後は2018年版は閲覧できるものの、最新版（2019年版）は閲覧できていない。（2020年10月19日 著者）
Citation	医学図書館 = Journal of the Japan Medical Libraries Association, 67(2), 89-93
Issue Date	2020-06
Doc URL	https://hdl.handle.net/2115/79556
Type	journal article
File Information	JJMLA_67_2_89-93.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

論文評価指標の要点整理と動向紹介

千葉　浩之*
北海道大学北キャンパス図書室

Ⅰ．はじめに

研究活動を支える図書館員にとって論文に関わる業務は多い。研究者が必要とする論文の入手のためにジャーナル契約や文献の取り寄せを行い，片や機関リポジトリの運営を通して発表論文のセルフアーカイブを支援する。こうした業務を背景として論文評価指標についても研究者から問い合わせを受けることがあるだろう。

今や論文評価指標は，個々の研究者の業績評価や購読ジャーナルの選定にとどまらず，国レベルの研究政策といった，より大きな問題にも用いられる。また，かねてより指標の誤用が懸念されており，2015年には「研究計量に関するライデン声明」がNature誌上で発表された[1)]。さらに情報技術や計算技術の発展により今後も新たな指標の登場が予想される。

こうした状況に鑑み，論文評価指標に対しては最新の知識はもとより，先々にも活かせる深い理解が必要と考える。そこで本稿では，先行研究[2)]を踏まえたうえで，論文評価指標の前提を掘り起こして整理しつつ，最近の動きを紹介したい。

Ⅱ．研究評価のひとつとしての論文評価

1．研究評価指標の需要の高まり

研究資金獲得競争の激化に加え，エビデンスに基づいた資金配分や説明責任が求められるなか，客観的な研究評価指標の需要が高まっている。たとえば令和元年度の国立大学法人運営費交付金の配分において一部の国立大学に対する配分指標として「運営費交付金等コスト当たりTOP10％論文数」が試行的に用いられた[3)]。また，第5期科学技術基本計画では，目標値のひとつとして「総論文数」の増加とともに「総論文数に占める被引用回数トップ10%論文数の割合」が設定されている[4)]。

*Hiroyuki CHIBA：〒001-0021 北海道札幌市北区北21条西10丁目.
chiba@lib.hokudai.ac.jp
（2020年4月15日　受理）

有限な資源の配分にあたり，何らかの研究評価指標を根拠に用いることは，もう少し身近な出来事，たとえば研究者の採用（研究ポストの配分）や読むべき論文の選択（研究時間の配分）においても見受けられる。

2．研究評価対象としての論文

上述の例のように，論文を研究評価の対象と見なすことは多い。確かに論文は主要な研究成果であるが，2012年に米国細胞生物学会で採択された「研究評価に関するサンフランシスコ宣言」によれば「科学研究の成果というものは数多くかつ多様であり，新たな知見を報じる研究論文，データ，試薬，ソフトウェアや，知的財産権，あるいは熟練した若手研究員等もこれに含まれるといえる」[5)]ので，論文だけで研究を評価することの不十分さも認識しておきたい。医学分野との関わりは薄いかもしれないが，書籍や会議発表が主要な評価対象となる研究分野の存在も押さえておくのが健全だろう。

また，一口に論文と言っても，掲載媒体や論文の種類は様々であることにも留意したい。前者は紀要，国内査読誌，国際査読誌等があり，後者は原著論文，レビュー論文，速報等がある。

3．論文の量的評価と質的評価

それでは論文の何を評価するか。機関や研究者を評価する場合，単純に論文数で測ることもある。この場合，規模の大きい機関や研究歴の長い研究者が有利になってしまう。また，量を測る際の留意点として共著に由来するカウントの問題もある。1本の論文を，各研究者や各機関に対して1と数えるのを整数カウント，共著者数で割った値を振り分けるのを分数カウントと呼ぶ。

そもそも量に関する問題の前提として，どの論文を対象とするのかという質の問題がある。論文の質は，最終的に研究分野ごとに専門家の判断に委ねることになるが，分野が細分化され，論文数が増加する現代においては，網羅的に論文を確認し，評価を下すことは難しい。

4．論文のインパクト評価

そこで，質的評価を手助けするものとして長らく用いられている評価軸が被引用である。引用を多く集めた論文が質の高い論文だとは必ずしも断言できないが，少なくとも後の研究に何らかのインパクトを与えたとは言えよう。

さらに近年，新たな評価軸としてソーシャルメディア等での反応もインパクトとして注目されている。このインパクトはオルトメトリクスの登場により計量可能となった。

数多の論文評価指標は，こうした論文発表後のインパクトを測るもので，上述の2種類の評価軸のどちらか，もしくは組み合わせに基づいている。以下，これらの評価軸に沿って論文評価指標を見ていこう。

Ⅲ．被引用回数：伝統的な論文評価指標

1．被引用回数の把握

そもそも被引用回数がどのように把握されるかを引用の現象面に立ち戻って考えてみたい。

場所の観点で考えると，引用は学生のレポートから小規模な研究会の発表資料，国際的な査読誌掲載論文に至るまで，様々な成果発表において現れる。

当然これらすべてを数え上げるのは不可能で，範囲を絞り込む必要がある。そこで用いられるのが，Web of Science[6] やScopus[7]（ともに有料）といった，文献間の引用を記録する引用文献データベースである。これらはそれぞれの収録基準で記録すべき文献を選別する。是非はともかく，これらの収録基準が論文評価の枠組みや前提となっている。収録文献は引用文献データベースによって異なるため，同じ論文でも被引用回数に違いが生じうる。そのため被引用回数を把握する際は，用いる引用文献データベースを決めておく必要がある。

次に時間の観点で考えてみる。ある論文が引用されるまでには，その論文がアクセスされ，読まれ，読者が研究成果を発表する際に情報源として記載される必要がある。信頼のおける出版物での発表となれば，さらに投稿，査読，出版を経て，上述のような引用文献データベースに収録される。最初の被引用が記録されるまでには通常，出版から数か月以上の時間がかかる。一方で，理論上引用されることに終わりはなく，引用文献データベースがオンラインで提供するデータは日々更新されるため，調査する日時も定める必要がある。

2．被引用回数の比較

そうして被引用回数が得られたとしても，その数値だけではインパクトの大小はわからない。それを知るためには，以下の3つの条件が同じである論文集合の中で比較する必要がある。

1）出版年

出版から時間が経過していない論文ほど引用される機会は少ない。同じ年でも1月と12月とでは時間差があるが，年を単位とするのが主流である。

2）分野

分野によって引用や論文出版の頻度に差があり，その多寡は被引用回数にも影響を及ぼす。そのため異なる分野間の比較ではなく，同一分野内での比較が妥当である。

3）論文の種類

論文の種類の違いも被引用回数に影響する。たとえばレビュー論文と原著論文とでは前者が引用されやすい。

なお，分野や論文の種類は引用文献データベースが索引付ける。分野については，一般にジャーナルに対して付与された分野が掲載論文にもそのまま付与される。複数の分野が付与される場合もあり，同じ論文でもどの分野の論文として測るかで評価が異なってくる。

さらに細かい話になるが，どこからの引用をカウントするか，たとえば会議録や専門書から受けた引用も対象となるか等も留意すべき点である。

さて，こうした論文集合内での位置を知るには，集合内の全論文の被引用回数を数え上げる必要があり，引用文献データベースからデータを抽出するだけでも大きな労力である。例外的にScopusは「被引用ベンチマーキング」指標において集合内での位置を数値化して表示している。しかし論文1本ずつ確認が必要で，一括抽出もできない。そこで，実際にはInCites[8] やSciVal[9]（ともに有料）等の引用文献データベースのデータに基づく研究分析ツールが用いられる。

なお，出版年，分野，論文の種類のうちひとつでも異なる論文間ではインパクトの大小について直接的な比較はできない。それでも調べる必要が生じた場合は，それらの被引用回数がそれぞれの論文集合内でどの位置にあるかを比較することで，間接的に求められる。

Ⅳ．被引用回数から派生した評価指標

被引用回数から様々な指標が派生する。ここでは代表的なものを3つ紹介する。

1．h-index

h-indexはHirschが提唱した，主に研究者に対して論文の量とそのインパクトとを同時に測る指標で，論文数

とそれらの被引用回数との両方を満たす最大の数値を指す[10]。たとえば，被引用回数が20，14，3，2，1の5本の論文を持つ研究者のh-indexは3（被引用回数3以上の論文が3本）で，同様に9，8，6，5，2の5本の場合はh-indexは4（被引用回数4以上の論文が4本）である。

指標の元となる論文数と被引用回数は引用文献データベースに依存するため，何で調べるかによってh-indexも異なる。Web of ScienceやScopusでは著者を同定すれば，機械的に数値が算出される。

h-indexの難点としては，論文発表の機会が少ない若手や引用頻度の少ない分野の研究者が不利となることが挙げられる。逆に，研究歴と分野が近しい研究者同士の比較，たとえば，ある研究ポストへの採用審査等においては，参考情報として活用されうる。

2．Top10%論文

出版年，分野，論文の種類が等しい論文集合内で被引用回数順に並べたときに上位10%に入る論文を「Top10%論文」等と呼ぶ。

この指標もどの引用文献データベースに基づくのか，さらには分野や論文の種類をどう線引きして集合を構成するかに依存する。また，被引用回数の増加により論文集合内の位置は常に変動しうる。

この指標は論文そのものの評価だけでなく，そうした論文の本数，あるいは総論文数に占める割合等を出すことで，機関や研究者個人の評価にも用いられる。

実際の評価や政策に用いられるものとしては，文部科学省科学技術・学術政策研究所が算出する「Top10%補正論文数」があり，これは量に関する指標である[11]。分析対象はWeb of Scienceのデータベース群のうち，医学を含む自然科学のそれ（Science Citation Index Expanded）に収録されたArticleとReviewで，被引用回数は年末時点の数値を用いている。分野はEssential Science Indicatorsの22分野分類をベースとしている。

3．Journal Impact Factor（JIF）

Journal Impact Factor（JIF）はClarivate Analytics社が提供するWeb of Scienceのデータを元にしたジャーナルに対する指標である。「インパクトファクター」や「IF」等と呼ばれ，読者にも馴染み深いかと思われる。

この指標は医学を含む自然科学系と社会科学系のジャーナルに対して，その掲載論文が平均してどれだけ引用されたかを示す。計算式は単純な割り算で，分母は算出対象となる年の1年前と2年前に掲載された論文数（論文の種類はArticleとReview）で，分子は算出対象年にそれらの論文がWeb of Science Core Collection収録文献（専門書や会議録を含む）から得た被引用回数である。最新版の数値が算出対象年の翌年6月頃に発表される。

JIFは過去のジャーナル掲載論文の被引用回数の平均値であり，個々の論文の被引用回数が均一して分布するとは限らない。にもかかわらず論文の評価に誤用されるケースが散見される。

従来この指標はJournal Citation Reports（JCR）[12]というデータベースを契約していないと把握できなかった。現在はウェブサイト（Web of Science Master Journal List）[13]に無料アカウントでログインすれば，各ジャーナルの「Journal Metrics」の項目において最新版のJIFを閲覧できる。いわゆるハゲタカジャーナルはJIFと紛らわしい指標を掲げることがあるので，JCR未契約機関の読者も実際に確認いただきたい。

なお，JCRでは過去のJIFや，同一分野内でのJIFの順位や四分位（Quartile）等の情報を参照できる。

Ⅴ．新しい技術を用いた被引用指標

近年，新しい技術を用いた被引用指標も登場している。引用の既存の枠組みや前提に切り込む一方，数値の検算が容易ではないため，評価のいわばブラックボックス化の懸念も考えられる。2例だけだが，紹介したい。

1．Relative Citation Ratio（RCR）

Relative Citation Ratio（RCR）は，米国国立衛生研究所のOffice of Portfolio Analysisが開発した，論文に対する指標で，PubMed収録文献からの被引用が分析対象である[14]。

先に紹介した指標群は，分野として引用文献データベースが付与したものを採用している。だが，実際の研究分野は最初から固定されてはおらず，研究活動の展開により確立されていくものであろう。RCRは固定した分野の代わりに，一緒に引用された論文群（共（被）引用グループ）を比較のまとまりとして用いる。引用されるたびに，その論文の共（被）引用グループは拡大する。

RCRは以下のように算出される。まず対象論文の1年あたりの被引用回数（Article Citation Rate：ACR）を求める。次にアルゴリズムを用いて共（被）引用グループにおける1年あたりの被引用数の期待値（Field Citation Rate：FCR）を求める。ACRをFCRで割ったものがRCRである。RCRが1であれば期待値どおりの，

2であれば期待値の2倍引用された論文となる。指標それ自体が比較可能な数値として表現される。

RCRはウェブサイト（iCite）[15]で無料で確認できる。

2．sciteによる引用分析

被引用回数の多さはインパクトの大きさを表すとしても，引用された論文の妥当性を示すわけではない。

米国国立科学財団と米国国立薬物乱用研究所から一部助成を受けたスタートアップであるsciteは，論文群を深層学習にかけて，論文内での引用が持つ意味をSupporting（支持），Mentioning（言及），Contradicting（否定）に分類し，それぞれ数値化している。引用の意味に踏み込んだ取り組みで，今後の展開が注目される。

これらの数値はsciteのウェブサイト[16]のほか，ChromeやFirefoxのブラウザ・プラグイン（無料）を使えば論文ページ上でも参照できる。

Ⅵ．オルトメトリクス：新たな評価軸の形成

1．オルトメトリクスの概要

近年，論文の電子化やウェブ技術の発展等を背景に，論文に対してウェブ上で様々な反応をすることが可能になった。Priemらによって提唱されたオルトメトリクス（Altmetrics）は，そうした反応，特にソーシャルメディアでの反応を計測，評価しようというものである[17]。「Alternative Metrics」に由来する造語で，既存の書誌（に含まれる引用）情報に基づく指標（Bibliometrics）では測定できないインパクトを顕在化させている。

2．論文に対してウェブ上で行われる反応

被引用という反応は論文出版の上に成り立ち，それを把握するには数か月以上の時間がかかる。一方で，現代では論文に引用して議論するというかたちを経ずとも，論文がオンラインで出版された直後にSNSやブログで言及できる。さらに言及にまで至らずとも，文献管理ソフトへの保存やソーシャルブックマークへの登録等，論文への注目を示す反応はすぐに行え，これらの件数は集計可能である。また，論文を出版するのは事実上研究者に限られるが，これらの反応は広く誰もが可能である。

3．オルトメトリクス指標に対する評価

論文に対してウェブ上で行われる反応は，多種多様であり，誰もが，そして場合によっては論文を読まずとも可能である。そのためオルトメトリクス指標の評価は難しい。被引用の場合，基本的に対象となるのは査読を経た論文からの引用であり，さらに引用文献データベースが論文の選別から分野や論文の種類に至るまで評価の枠組みを構成している。片やソーシャルメディアでの反応には一般に事前審査はなく，枠組みも定めがたい。オルトメトリクス指標については少なくとも被引用指標とはベクトルの異なるものと見るのが適切であろう。

なお，オルトメトリクスは論文に限らず，実験や観測，調査等で得られたデータセットや図表といった，多様な研究成果物を対象にできる。本稿ではオルトメトリクスを論文評価の面から扱うが，実際はさらに広い射程を持つ点にも留意しておきたい。

Ⅶ．代表的なオルトメトリクス指標

最後にオルトメトリクス指標を3つ紹介する。

1．Faculty OpinionsのArticle Recommendation

Faculty Opinions[18]は有料の論文推薦サービスで，医学・生物学分野の名声が確立した専門家からの推薦を得た論文を提示している。論文に対する反応の主体がそうした専門家に限られて広がりは乏しいものの，論文出版後，比較的早期に信頼のおける反応が得られる。指標としてExceptional，Very Good，Goodの3段階の推薦があり，それに加えてコメントも閲覧できる。推薦以外にDissent（異議）も付けられ，そのコメントも読める。

2．Altmetric

AltmetricはDigital Science社が無料で提供する指標で，ニュースやSNSでの言及，Wikipedia上での被引用等，様々なインパクト数値を収集し，種類ごとに重み付けのうえ合算している[19]。

Altmetricのドーナツ型のバナーはよく目にするかと思うが，同社が提供する学術情報プラットフォームDimensions[20]上でも無料版の機能として，被引用指標とともに参照できる。さらに無料のブックマークレット[21]を入手すれば，論文ページにアクセスした際に1クリックでその論文のAltmetricの数値を確認できる。

被引用回数と同様，数値だけでは多寡の評価が難しいが，出版年や出版物別に順位も表示している。

3．PlumX Metrics

PlumX Metricsは有料の引用文献データベースScopusを介して確認できる指標群である[22]。Usage（ダウンロード回数等），Captures（文献管理ソフトへの保存回数等），Mentions（ブログやニュースでの言及回数等），

Social Media（SNSでの言及回数等），そしてCitations（被引用回数）の5つのカテゴリーごとに数値を表示する。また，コードを埋め込むことでScopus外のウェブサイト経由での表示も可能である。

Ⅷ. おわりに

冒頭に述べたとおり，研究活動を支える図書館員はこれまで論文の入手やセルフアーカイブに努めてきた。一方で，近年のオープンサイエンスの潮流を踏まえて，仮に論文のオープンアクセスが実現した未来を想像すると，そうした業務の需要は減るだろう。アクセスの問題の代わりに，インパクトや評価の問題が前面に出てくる可能性がある。その場合，論文評価指標の知識が欠かせない。

欧米の事例ではあるが，近年Research Impact Librarianなる職務[23]の登場や，研究評価指標に関する図書館員向けのガイド[24]やサービス実例の報告書[25]の出版等，論文評価指標に取り組む図書館員の姿が見受けられる。こうした指標を理解してサービスに実装することで，図書館員は学術コミュニケーションにおいて新たな位置付けを獲得できるのではないだろうか。読者にとって本稿がその一助となれば幸いである。

参考文献

1）小野寺夏生，伊神正貫．研究計量に関するライデン声明について．STI Horizon．2016;2(4). [internet]．http://doi.org/10.15108/stih.00050 [accessed 2020-05-22]

2）孫媛．研究評価のための指標：その現状と展望．情報の科学と技術．2017;67(4):179-84. [internet]．https://doi.org/10.18919/jkg.67.4_179 [accessed 2020-05-22]

3）文部科学省．令和元年度成果を中心とした実績状況に基づく配分の仕組みについて[internet]．https://www.mext.go.jp/a_menu/koutou/houjin/1417264.htm [accessed 2020-05-22]

4）内閣府総合科学技術・イノベーション会議 科学技術イノベーション政策推進専門調査会．第5期科学技術基本計画における指標の活用について[internet]．https://www8.cao.go.jp/cstp/kihonkeikaku/sihyou/katuyou.pdf [accessed 2020-05-22]

5）Ozono H, Kato S, Sugita S, Suzuki M, Tsuchide I, Kuriyama M（ほか）訳．研究評価に関するサンフランシスコ宣言[internet]．https://sfdora.org/read/jp [accessed 2020-05-22]

6）Web of Science[internet]. http://webofknowledge.com/WOS [accessed 2020-05-22]

7）Scopus[internet]. https://www.scopus.com [accessed 2020-05-22]

8）InCites[internet]. https://incites.clarivate.com [accessed 2020-05-22]

9）SciVal[internet]. https://www.scival.com [accessed 2020-05-22]

10）Hirsch JE. An index to quantify an individual's scientific research output. PNAS. 2005;102(46):16569-72. [internet]. https://doi.org/10.1073/pnas.0507655102 [accessed 2020-05-22]

11）村上昭義，伊神正貫．研究論文に着目した日英独の大学ベンチマーキング2019：大学の個性を活かし，国全体としての水準を向上させるために（報告書全文）．文部科学省科学技術・学術政策研究所;2020.p.25-32. [internet]．http://doi.org/10.15108/rm288 [accessed 2020-05-22]

12）Journal Citation Reports[internet]. https://jcr.clarivate.com [accessed 2020-05-22]

13）Web of Science Master Journal List[internet]. https://mjl.clarivate.com [accessed 2020-05-22]

14）Hutchins BI, Yuan X, Anderson JM, Santangelo GM. Relative Citation Ratio (RCR): A New Metric That Uses Citation Rates to Measure Influence at the Article Level. PLOS Biology. 2016;14(9):e1002541. [internet]. https://doi.org/10.1371/journal.pbio.1002541 [accessed 2020-05-22]

15）iCite[internet]. https://icite.od.nih.gov [accessed 2020-05-22]

16）scite[internet]. https://scite.ai [accessed 2020-05-22]

17）Priem J, Taraborelli D, Groth P, Neylon C. Altmetrics: A manifesto. 2010. [internet]. http://altmetrics.org/manifesto [accessed 2020-05-22]

18）Faculty Opinions[internet]. https://facultyopinions.com [accessed 2020-05-22]

19）Altmetric[internet]. https://www.altmetric.com [accessed 2020-05-22]

20）Dimensions[internet]. https://www.dimensions.ai [accessed 2020-05-22]

21）Bookmarklet - Altmetric[internet]. https://www.altmetric.com/products/free-tools/bookmarklet [accessed 2020-05-22]

22）Plum Analytics[internet]. https://plumanalytics.com [accessed 2020-05-22]

23）Skinker L. New research impact librarian tracks scholarship. Virginia Tech Daily. 2018. [internet]. https://vtnews.vt.edu/articles/2018/11/univ-lib-researchimpactlibrarian.html [accessed 2020-05-22]

24）Roemer RC, Borchardt R. Meaningful metrics: A 21st century librarian's guide to bibliometrics, altmetrics, and research impact. Chicago:Association of College and Research Libraries;2015. [internet]. http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/booksanddigitalresources/digital/9780838987568_metrics_OA.pdf [accessed 2020-05-22]

25）Kirchner M. Profiles of Academic Library Efforts in Bibliometrics & Altmetrics. New York:Primary Research Group;2019.