

ОРИГИНАЛЬНАЯ СТАТЬЯ

УДК 004.91: 005.513

doi: 10.26907/2541-7746.2024.4.485-498

ЖУРНАЛЬНЫЙ ИМПАКТ-ФАКТОР: МАТЕМАТИЧЕСКИЕ И СТАТИСТИЧЕСКИЕ СВОЙСТВА

А.М. Елизаров¹, В.В. Писляков²

¹*Казанский (Приволжский) федеральный университет, г. Казань, 420008, Россия*

²*Национальный исследовательский университет «Высшая школа экономики»,
г. Москва, 109028, Россия*

Аннотация

Дано определение журнального импакт-фактора в классической форме и исследованы различные аспекты его применения в области математического моделирования научной деятельности. Изучены свойства импакт-фактора, связанные с особенностями дисциплинарной структуры науки, ограниченным набором источников в существующих базах данных, статистической достоверностью, чувствительностью к типам публикуемых документов, влиянием выбросов. Кроме того, обсуждается вопрос об искажениях, которые может внести самоцитирование журнала.

Ключевые слова: библиометрические индикаторы, импакт-фактор, Web of Science, Journal Citation Reports, журнальное самоцитирование

Введение

Как известно, при математическом моделировании процессов развития науки используют специальные индикаторы для характеристики деятельности различных участников научного процесса – ученых, лабораторий, факультетов, организаций, стран. Те показатели, которые связаны со статистикой опубликованных работ, называются библиометрическими показателями. Они применяются также для оценки отдельных научных изданий, в первую очередь – журналов. Среди журнальных библиометрических характеристик хронологически первым и наиболее известным является импакт-фактор журнала.

Цель настоящей работы – подробно изучить различные математические и статистические свойства импакт-фактора, проявляющиеся в эмпирических условиях, т. е. на множестве реальных мировых научных журналов. Мы будем использовать критический подход. Что имеется в виду? По смыслу создания импакт-фактор должен позволять сравнивать «качество» журналов между собой. Однако в реальности есть ряд «недостатков» данного индикатора в смысле отклонения от выполнения данной функции. Исследуя эти недостатки, отклонения и несовершенства, можно более наглядно и отчетливо пронаблюдать эмпирические свойства импакт-фактора. Однако прежде всего необходимо кратко охарактеризовать то множество, «поле» журналов, на котором применяется импакт-фактор (как и иные журнальные индикаторы), – библиометрические базы данных.

1. Базы данных

Каждый библиометрический показатель, в том числе импакт-фактор научного журнала, рассчитывается по той или иной базе данных научного цитирования. Поскольку всего в мире имеется не менее 400 тыс. периодических изданий [1], из которых около 50 тыс. – академические журналы [2], невозможно подсчитать ссылки по всей мировой совокупности изданий, полученные журналом/статьей/автором. Поэтому опубликование каждого индикатора, помимо методики, сопровождается указанием инструмента, базы данных, на основании которого он рассчитан. Значения показателя, рассчитанные по идентичной методике, но по разным базам данных, конечно же, будут отличаться. Наиболее часто используемыми и упоминаемыми в литературе являются базы данных Web of Science (WoS) компании Clarivate и Scopus компании Elsevier. Они включают в себя ок. 22000 и ок. 27000 журналов соответственно. Кроме того, эти ресурсы содержат книги (ок. 150 тыс. в WoS и ок. 350 тыс. в Scopus), а также труды конференций. Цитирования из книг и трудов конференций также индексируются и учитываются. В России существует собственная библиометрическая база данных – Российский индекс научного цитирования (РИНЦ) [3], https://elibrary.ru/project_risc.asp. Хотя говорить мы будем об импакт-факторе, который рассчитывается по WoS, стоит упомянуть, что в базе данных Scopus также есть методологически очень похожий показатель – CiteScore.

Во всех упомянутых базах данных научного цитирования приводятся сведения о числе ссылок, полученных отдельными статьями. Кроме того, эти данные объединены по авторам («авторские профили») и организациям («профили организаций»), в результате чего можно посмотреть суммарную статистику публикационной активности и цитируемости ученых или организаций. В случае Scopus и РИНЦ в стандартном интерфейсе этих платформ можно также получить индикаторы журналов как целого. В WoS эти индикаторы вынесены в специальную базу данных, агрегирующую журнальные показатели, – Journal Citation Reports (JCR).

2. Импакт-фактор: краткая история и определение¹

Импакт-фактор показывает среднюю цитируемость одной статьи в журнале. Термин впервые был введен Юджином Гарфилдом в 1955 г., но для обозначения цитирования отдельных статей [5] без привязки к журналу. Позже, в 1963 г., Ю. Гарфилд и И. Шер [6] использовали термин в известной сейчас форме – в отношении цитирования статей конкретного журнала.

Отметим малоизвестных предшественников импакт-фактора. В пионерской работе [7] авторы подсчитывали суммарное число ссылок на тот или иной журнал из ведущего журнала в области химии «Journal of the American Chemical Society», однако без нормализации по объему цитируемых изданий. В 1936 г. И. Хак [8] уже использовал нормализацию, но не на число отдельных статей, а на число вышедших томов журнала. В 1960 г. еще один журнальный показатель был предложен Л. Рейзигом [9] под названием «индекс реализованного исследовательского потенциала» – подсчитывалась доля хотя бы один раз процитированных статей в журнале. Здесь отличие от импакт-фактора состоит в том, что не учитывалось, сколько именно ссылок получила статья, – важен только факт цитирования статьи (было/не было)².

¹См. также статью «Импакт-фактор» в Большой российской энциклопедии [4].

²Обратный этому показатель сейчас называют «коэффициент нецитируемости».

Классический импакт-фактор IF_J в году Y для журнала J задаётся выражением

$$IF_J(Y) = \frac{\sum_{i=1}^2 CIT_J(Y, Y-i)}{\sum_{i=1}^2 PUB_J(Y-i)} = \frac{CIT_J(Y, Y-1) + CIT_J(Y, Y-2)}{PUB_J(Y-1) + PUB_J(Y-2)}, \quad (1)$$

где $CIT_J(Y1, Y2)$ – число цитирований (ссылок), полученных в году $Y1$ статьями журнала J , вышедшими в году $Y2$; $PUB_J(Y)$ – число статей, вышедших в журнале J в году Y . Нормировка на число статей позволяет убрать преимущество объемных журналов перед небольшими, а часто выходящих (например, ежемесячных) – перед редко выходящими (например, ежегодниками). Таким образом, классический импакт-фактор журнала – это среднее (в расчете на одну статью) число цитирований, которое получают в отчетном году статьи, вышедшие за два предыдущих года. Поскольку подсчитываются ссылки, полученные в одном (отчетном) году документами, опубликованными в течение двух лет, говорят об однолетнем «окне цитирования» и двухлетнем «публикационном окне».

Отметим, что в импакт-факторе не учитываются ссылки, полученные публикациями в год их выхода. Для этого в системе индикаторов JCR есть отдельный показатель – «индекс оперативности» (immediacy index), который учитывает только такие цитирования. Кроме того, отметим, что среди ссылок, учитываемых в (1), есть ссылки журнала на самого себя, т. н. «журнальное самоцитирование». Пока уточним это как методологический момент, позднее самоцитирование будет исследовано подробнее.

3. Методологические несовершенства и проблемы в использовании импакт-фактора

Перейдем, как было заявлено во введении, к критике импакт-фактора, т. е. к исследованию сложностей при его практическом использовании. Это позволит в деталях рассмотреть различные характеристики этого статистического индикатора.

3.1. Дисциплинарные различия. Серьезной, наиболее значимой проблемой импакт-фактора является принципиальная невозможность прямого сравнения его значений у журналов, относящихся к различным дисциплинам. Диапазоны и медианные журнальные импакт-факторы в разных областях науки отличаются практически на порядок, как показано в табл. 1. Этот эффект наблюдается как в естественных/технических/медицинских, так и общественных науках.

Табл. 1

Медианные импакт-факторы журналов по различным научным дисциплинам (JCR-2021)

Дисциплина	Медианный импакт-фактор
Естественные науки	
Cell Biology	4.7
Physics, Applied	2.7
Mathematics	1.0
Общественные науки	
Management	4.5
Sociology	2.1
History	0.6

Для примера: математический журнал с таким же импакт-фактором, какой имеет медианный, средний журнал по клеточной биологии (4.7), занимал бы очень высокую позицию в рейтинге из всех 333 математических журналов, представленных в JCR-2021, – он был бы на третьем месте.

Табл. 1 показывает, что разница в уровне цитируемости отдельных научных направлений наблюдается как в естественных (hard sciences), так и в общественных науках (social sciences). Более того, если взять эти конгломераты наук как целое, то и на таком «метауровне», объединяющем тысячи журналов, существует тот же эффект. В [10, с. 187] вычислены показатели из JCR-2018: медиана импакт-факторов журналов по всем естественным наукам (точнее, объединению естественных, технических и медицинских наук) равна 1.91, а по общественным наукам – 1.40.

3.2. Дисциплинарные границы, устанавливаемые на основе журналов. Как следует из предыдущего раздела, напрямую сравнивать журналы по их импакт-факторам имеет смысл только в рамках одной научной области. В базах данных WoS/JCR есть деление статей по научным направлениям, однако дисциплина присваивается документам только на основании журнала, в котором они опубликованы, особенности индивидуальных статей не учитываются. Поэтому к той или иной научной дисциплине приписывается журнал как целое (при этом журнал может быть отнесен сразу к нескольким, как правило, 2–3 дисциплинарным категориям), и опубликованные в нем статьи автоматически получают ту же самую научную классификацию.

Часто это оказывается некорректным, например, статья, посвященная библиометрическому анализу журнала по химии [11] и, по сути, относящаяся в первую очередь к наукометрии/библиометрии, приписана в WoS/JCR к категории «Chemistry, Multidisciplinary», к которой относится опубликовавший ее журнал «Успехи химии». В подобных случаях журналы, публикующие такие мультидисциплинарные статьи, получают преимущество, если они «захватывают» ссылки из областей науки, в которых наблюдается более высокая средняя цитируемость, и наоборот, если затрагивается менее цитируемая дисциплина.

3.3. Зависимость от контента базы данных. Это слабое место импакт-фактора связано с особенностями дисциплин, которые обсуждались в п. 3.1, однако иногда оно проявляется и на уровне отдельных журналов в рамках одного научного направления. Контент WoS ограничен и не может включать в себя все мировые источники. Среди типов научной литературы наиболее полно представлены журналы, в то время как покрытие других типов (книги, диссертации и т. д.) во многом недостаточно, если оценивать процент проиндексированных в WoS источников от всех существующих в мире качественных научных изданий.

В результате дисциплины или отдельные журналы, которые получают больше ссылок «извне» базы данных WoS, например из книг, препринтов, диссертаций, патентов, официальных документов и др., оказываются в худшем положении, чем те, которые цитируются преимущественно в журнальной литературе. Более того, эта проблема существует и для изданий, которые получают ссылки из журналов, но регионального, локального значения – тех, что в худшей мере охвачены WoS/JCR.

3.4. Произвольный выбор публикационного окна и окна цитирования. В методике расчета импакт-фактора выбор двухлетнего публикационного окна и однолетнего окна цитирования, в принципе, является произвольным и не

имеющим специального обоснования. О случайностях и их последствиях при конструировании используемой сейчас формулы импакта см. [12].

В известной мере этот выбор обусловлен тем, что в самом начале указатели цитирования (которые в дальнейшем стали базой данных WoS) были созданы Ю. Гарфилдом для биомедицинских наук, в которых научное знание достаточно быстро устаревает. В таком случае имеет смысл брать небольшие временные интервалы выхода публикаций и их цитирования, иначе не получится оценить текущий, актуальный уровень журнала, который со временем может изменяться.

Однако для других, менее динамичных отраслей естественных наук и в особенности наук общественных двухлетний интервал оценки цитируемости может оказаться недостаточным. Его не хватает, чтобы профессиональное сообщество смогло должным образом оценить выходящие в свет публикации, статьи «не успевают набрать цитирования». Показательно здесь исследование В. Гленцля и Х. Мута [13], которые собрали данные по цитируемости с течением времени однолетнего массива статей из социологического и медицинского журналов. Табл. 2 показывает, что если в год выхода (1980) медицинская статья получила в среднем в три раза больше ссылок, чем статья по социологии, то на интервале, примерно соответствующем импакт-фактору, 1980–1982, показатели двух журналов сравниваются. Если же «отойти» на десять лет от года выхода статей (1980–1989), то окажется, что социологическая статья в сумме получила в полтора раза больше цитирований, чем статья по медицине.

Табл. 2

Среднее число ссылок на статью в социологическом и медицинском журналах при варьировании окна цитирования (источник: [13])

Окно цитирования (время наблюдения)	Среднее число ссылок, полученное одной статьей, вышедшей в 1980 г.	
	American Sociological Review	Lancet
1980	0.2	0.6
1980–1981	1.8	2.4
1980–1982	4.3	4.5
1980–1985	12.1	9.7
1980–1989	20.9	14.0

Эффект долгого «старения знания» и, соответственно, необходимости долгой аккумуляции ссылок для оценки статей и журналов продемонстрирован в табл. 2 на примере общественно-научного журнала. Однако и в сфере естественных и технических наук также существуют медленно устаревающие дисциплины (например, зоология и геология). Более того, наиболее динамичные в этом плане области все равно выходят за рамки двух лет, отмеренных импакт-фактором: минимальная медиана хронологического распределения ссылок, сделанных в журналах, согласно JCR-2018, наблюдается в категориях «Электрохимия» и «Нанонауки и нанотехнологии», при этом она равна 5.7 лет [10, с. 216], т. е. половина литературы, цитируемой даже в этих науках, старше пяти лет, если считать от момента выхода цитирующих публикаций³.

³Интересно, что именно эта оценка наблюдалась самим Гарфилдом еще на заре эпохи указателей научных ссылок: «Наши исследования подтверждают, что учёные часто обращаются к источникам, вышедшим достаточно давно. В индексе 1961 г. более половины ссылок ведут на материалы старше пяти лет» [14].

3.5. Погрешность при вычислении импакт-фактора. Следующая уязвимость относится к статистическим характеристикам импакт-фактора. Значения индикатора исходно публикуются в JCR и в дальнейшем анализируются, как правило, без указания статистической погрешности. Создатель импакт-фактора в соавторстве с А. Пудовкиным отстаивал утверждение, что, поскольку набор журналов в базу данных JCR не является случайной выборкой из всех мировых журналов, статистической погрешности в импакт-факторе не существует в принципе [15]. Тем не менее сам процесс цитирования и получения/неполучения ссылки тем или иным документом считается многими исследователями стохастическим процессом, который они моделируют, в частности, отрицательным биномиальным распределением ссылок [16], распределением Пуассона [17, 18] или нормальным распределением [19]⁴. В итоге при построении значений импакт-факторов в рамках выделенной дисциплины с отметкой доверительного интервала, а также диапазона рангов, которые могли бы занимать журналы с учетом статистической погрешности, получаются графики, подобные рис. 1, взятому из [17]. Обычно определенность присутствует в самых верхних и самых нижних позициях, остальные значения импакта и места журналов в получающемся ранжировании в значительной мере размыты, и доверительные интервалы сильно перекрывают друг друга (рис. 1).

Это также приводит к оспариванию точности, с которой в JCR публикуются значения импакт-фактора. Многие годы этот показатель размещался в базе данных с точностью до третьей цифры после запятой, что рядом специалистов считалось избыточным и вводящим в заблуждение [20–22]. Под влиянием дискуссии компания Clarivate с 2023 г. стала публиковать значения импактов с одним знаком после запятой [23].

3.6. Зависимость от выбросов. Проблема среднего значения. Поскольку импакт-фактор – это среднее значение, среднее число ссылок, полученных одной статьей, то на него, как и на любой средний показатель, сильно влияют выбросы. Неординарно большое число ссылок, полученное всего одной-двумя статьями журнала, могут кардинально повысить журнальный импакт-фактор.

В принципе, как уже упоминалось, распределение цитирований по отдельным статьям журнала далеко от нормального, оно сильно асимметричное. Обычно существуют малый процент статей в журнале, которые получили много ссылок, и большая доля низкоцитируемых или вообще не процитированных статей. Такие закономерности в библиометрии объединяют термином «гиперболические распределения» [24, с. 71 и далее]. Таким образом, по импакт-фактору журнала трудно судить о цитируемости каждой отдельной статьи в нем.

Однако бывают крайне рельефные случаи, как в показанном на рис. 2 примере журнала «Acta Crystallographica A». Благодаря высокой цитируемости одной-единственной статьи, вышедшей в 2008 г., этот журнал на короткое время (пока статья попадала в публикационное окно вычисления импакт-фактора) вошел в верхушку рейтинга всей базы JCR.

Рис. 2 наглядно показывает, что высокий импакт-фактор не является следствием какого-то значительного улучшения качества и повышения научного уровня рассматриваемого журнала: когда статья, получившая на данный момент (2024) более 75 тысяч ссылок, перестала попадать в интервал расчета, импакт-фактор издания резко упал до прежнего скромного уровня.

⁴Последний вариант, увидевший свет в журнале по экономике, скорее является ошибочным, поскольку распределение ссылок по статьям существенно негауссовое, асимметричное.

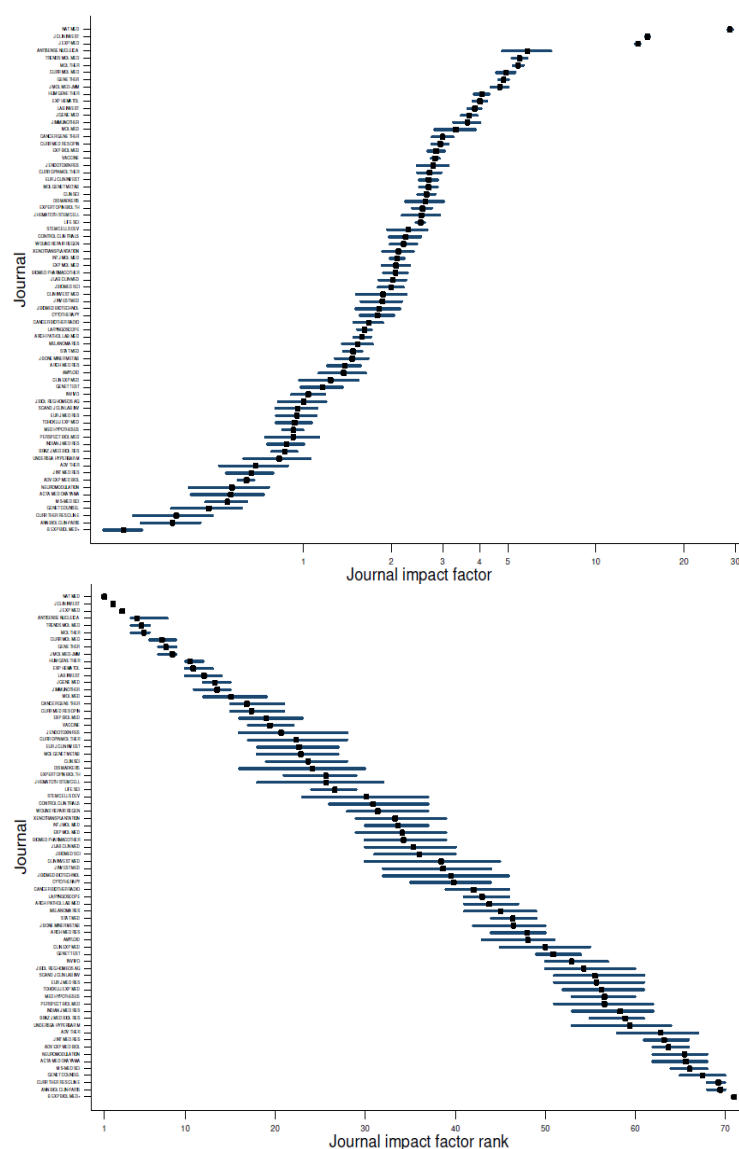


Рис. 1. Журналы тематической категории «Medicine, Research, and Experimental». Вверху: значения импакт-факторов с доверительными интервалами. Внизу: позиции в ранжировании по импакт-фактору с доверительными интервалами. Источник: [17]

3.7. Зависимость от типа публикаций. Если детально рассмотреть методику вычисления импакт-фактора, в реальности используемую компанией Clarivate, то обнаружится различие в учете типов документов в числителе и знаменателе формулы (1). В числителе, количестве полученных ссылок, учитываются цитирования, полученные всеми возможными публикациями оцениваемого журнала, вне зависимости от их типа. Однако в знаменателе подсчитываются только так, как их обозначают в JCR, – «документы, которые могут быть процитированы» (citable items). Это исследовательские статьи (Article) и научные обзоры (Review). Остальные виды публикаций, например редакционные статьи, письма,

обзоры конференций и др., исключены [23]. Такое различие в подходе к типам документа числителя и знаменателя продолжительное время отстаивается владельцем JCR [25], однако иногда продолжает считаться слабым местом методики расчета импакт-фактора.

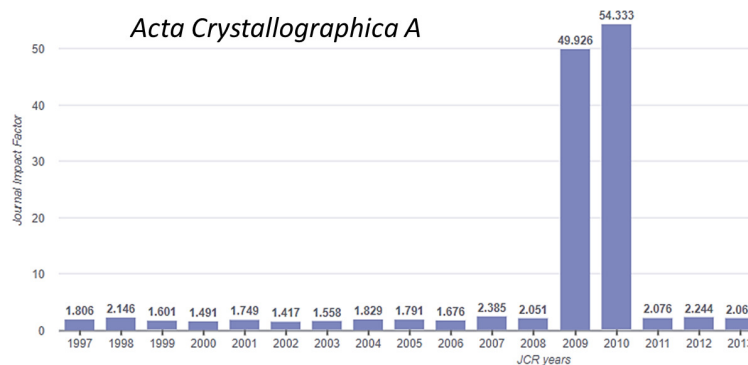


Рис. 2. Динамика импакт-фактора журнала «Acta Crystallographica A», 1997–2013. Источник: JCR

Вероятно, еще более серьезной проблемой является то, что разные типы документов цитируются по-разному. Многие исследования показали, что в среднем научные обзоры цитируются в несколько раз активнее, чем оригинальные исследовательские статьи (разумеется, при сравнении в рамках одной и той же области науки) [26, 27]. Это означает, что журнал, публикующий много обзоров (или исключительно обзоры, как, например, «Успехи химии» или журналы «Annual Review of ...»), получает только на этом основании преимущество в импакт-факторе перед иными изданиями. Это, вероятно, несправедливо и вводит в заблуждение, поскольку не имеет отношения к научному уровню публикуемых материалов.

3.8. Нечувствительность к научному уровню цитирующих источников. Еще одним недостатком методики расчета импакт-фактора можно считать его «безразличие» к тому, откуда оцениваемый журнал получает ссылки. Все приходящие ссылки одинаковы, импакт-фактор «не замечает», из какого именно источника получено то или иное цитирование. Каждое из них считается как «+1» в числителе индикатора, замена одного цитирующего журнала на другой, который дает столько же ссылок, никак не повлияет на результат. Нередко такой подход подвергается критике: должно ли упоминание журнала в мировых лидерах, например «Nature», «Reviews of Modern Physics», «Acta Numerica» и т. д., быть равнозначным ссылке из ординарного и лишь локально известного издания? В частности, в таком случае сложно отличить лидеров на мировой арене от звезд регионального масштаба, получающих ссылки из узкого круга знающих их авторов (см. также следующий пункт про самоцитирование).

3.9. Зависимость от журнального самоцитирования⁵. Эта сложность при использовании импакт-фактора для оценки уровня журналов связана не с его методологическими дефектами, а с теоретической возможностью достаточно легко

⁵Здесь термином «журнальное самоцитирование» мы обозначаем библиографическую ссылку из публикации в некотором журнале на другую публикацию из того же самого журнала. Речь не идет (как иногда бывает, когда применяют это понятие) о заимствовании текста автором из своих предыдущих публикаций.

манипулировать им. Как уже вкратце упоминалось, в определении формулы (1) в числитель входят все ссылки, полученные статьями журнала, включая цитирования из самого оцениваемого журнала. Любая публикация в журнале, цитирующая его же статью одно- или двухлетней давности, увеличивает импакт-фактор этого журнала. Предполагается, что этим могут манипулировать, например, редакции журналов, чтобы увеличить библиометрические показатели своих изданий. Иногда такие попытки наблюдались [28, 29].

Надо отметить, что общий вывод исследований журнального самоцитирования, как правило, таков: самоцитированием можно ощутимо изменить позицию слабого журнала, находящегося в «хвосте» ранжирования по импакт-фактору в своей дисциплине. Но, за крайне редкими исключениями, невозможно пробиться в топ рейтинга или существенно улучшить свое положение в нем. Тем не менее существует разумная рекомендация обращать внимание на роль самоцитирования в импакт-факторе и при возможности сопоставлять данные по метрикам с учетом и без учета ссылок журнала на самого себя. Подробнее см. обзор исследований журнального самоцитирования и «картелей» журналов, взаимно цитирующих друг друга, а также политики базы данных JCR по отношению к этим явлениям в [30].

Заключение

Как ответ на явные и часто серьезные недостатки импакт-фактора, перечисленные в разделе 3, профессиональное наукометрическое сообщество непрерывно работает над новыми, улучшенными метриками и уже выдвинуло множество интересных решений: пятилетний импакт-фактор, CiteScore, взвешенные индикаторы и индикаторы, построенные на методиках постатейного, а не пожурнального классификатора.

К сожалению, в сфере библиометрических журнальных индикаторов недостаточно предложить ту или иную методику, «идеальную формулу». Для полного ее принятия и внедрения в повседневные практики необходимы мощности и структуры, производящие массовые и регулярные расчеты нового индикатора по огромному массиву мировых научных журналов. Необходима «вычислительная фабрика», из года в год производящая результаты согласно вновь созданному алгоритму. Из сотен показателей, предложенных в теории, научных журналах и книгах по библиометрии, лишь часть получила такую поддержку от ведущих наукометрических компаний, которые теперь публикуют их на регулярной основе в тех или иных базах данных научного цитирования. Тем не менее сотрудничество ученых-наукомеров и производителей баз данных продолжается и становится всё более продуктивным.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Литература

1. ExLibris. Ulrichsweb: General FAQ. 2018.
[https://knowledge.exlibrisgroup.com/Ulrich's/Product_Documentation/Frequently_Asked_Questions_\(FAQs\)/Ulrichsweb](https://knowledge.exlibrisgroup.com/Ulrich's/Product_Documentation/Frequently_Asked_Questions_(FAQs)/Ulrichsweb)
2. Curcic D. Number of Academic Papers Published Per Year. 2023.
<https://wordrated.com/number-of-academic-papers-published-per-year/>
3. Moskaleva O., Pislyakov V., Sterligov I., Akoev M., Shabanova S. Russian Index of Science Citation: Overview and review // *Scientometrics*. 2018. V. 116, No 1. P. 449–462.
<https://doi.org/10.1007/s11192-018-2758-y>.

4. *Писляков В.В.* Импакт-фактор // Большая российская энциклопедия: научно-образовательный портал. URL: <https://bigenc.ru/c/impakt-faktor-c87b20/?v=8838250>. Дата публикации: 30.10.2023.
5. *Garfield E.* Citation indexes for science. A new dimension in documentation through association of ideas // *Science*. 1955. V. 122, No 3159. P. 108–111. <https://doi.org/10.1126/science.122.3159.108>.
6. *Garfield E., Sher I.H.* New factors in the evaluation of scientific literature through citation indexing // *Am. Doc.* 1963. V. 14, No 3. P. 195–201. <https://doi.org/10.1002/asi.5090140304>
7. *Gross P.L.K., Gross E.M.* College libraries and chemical education // *Science*. 1927. V. 66, No 1713. P. 385–389. <https://doi.org/10.1126/science.66.1713.385>.
8. *Hackh I.* The periodicals useful in the dental library // *Bull. Med. Libr. Assoc.* 1936, Sept. V. 25, Nos 1–2. P. 109–112.
9. *Raisig L.M.* Mathematical evaluation of the scientific serial // *Science*. 1960. V. 131, No 3411. P. 1417–1419. <https://doi.org/10.1126/science.131.3411.1417>.
10. *Писляков В.В.* Библиометрические индикаторы в ресурсах Clarivate. В кн.: Акеев М.А., Маркусова В.А., Москалева О.В., Писляков В.В. «Руководство по наукометрии: индикаторы развития науки и технологии». 2-е изд. Екатеринбург: Изд-во Уральск. ун-та, 2021. С. 177–220. <https://doi.org/10.15826/B978-5-7996-3154-3>.
11. *Зибарева И.В., Теплова Т.Н., Нефёдов О.М.* Библиометрический анализ журнала «Успехи химии» // *Усп. хим.* 2007. Т. 76, № 8. С. 747–751.
12. *Archambault É., Larivière V.* History of the journal impact factor: Contingencies and consequences // *Scientometrics*. 2009. V. 79, No 3. P. 635–649. <https://doi.org/10.1007/s11192-007-2036-x>.
13. *Glänzel W., Moed H.F.* Journal impact measures in bibliometric research // *Scientometrics*. 2002. V. 53, No 2. P. 171–193. <https://doi.org/10.1023/A:1014848323806>.
14. *Garfield E., Sher I.H.* Genetics Citation Index: Experimental Citation Indexes to Genetics with Special Emphasis on Human Genetics. Philadelphia, PA: Inst. Sci. Inf., 1963. xxvii, 864 p.
15. *Pudovkin A.I., Garfield E.* Rank normalization of impact factors will resolve Vancly's dilemma with TRIF // *Scientometrics*. 2012. V. 92, No 2. P. 409–412. <https://doi.org/10.1007/s11192-012-0634-8>.
16. *Schubert A., Glänzel W.* Statistical reliability of comparisons based on the citation impact of scientific publications // *Scientometrics*. 1983. V. 5, No 1. P. 59–73. <https://doi.org/10.1007/BF02097178>.
17. *Greenwood D.C.* Reliability of journal impact factor rankings // *BMC Med. Res. Methodol.* 2007. V. 7. Art. 48. <https://doi.org/10.1186/1471-2288-7-48>.
18. *Bornmann L.* Confidence intervals for Journal Impact Factors // *Scientometrics*. 2017. Vol. 111, No 3. P. 1869–1871. <https://doi.org/10.1007/s11192-017-2365-3>.
19. *Stern D.I.* Uncertainty measures for economics journal impact factors // *J. Econ. Lit.* 2013. V. 51, No 1. P. 173–189. <https://doi.org/10.1257/jel.51.1.173>.
20. *Vancly J.K.* Impact factor: Outdated artefact or stepping-stone to journal certification? // *Scientometrics*. 2012. V. 92, No 2. P. 211–238. <https://doi.org/10.1007/s11192-011-0561-0>.
21. *Zitt M.* The journal impact factor: Angel, devil, or scapegoat? A comment on J.K. Vancly's article 2011 // *Scientometrics*. 2012. V. 92, No 2. P. 485–503. <https://doi.org/10.1007/s11192-012-0697-6>.
22. *Bar-Ilan J.* Journal report card // *Scientometrics*. 2012. V. 92, No 2. P. 249–260. <https://doi.org/10.1007/s11192-012-0671-3>.

23. Clarivate. Journal Citation Reports™ Reference Guide. 2023. URL: https://clarivate.com/wp-content/uploads/dlm_uploads/2023/08/JCR-Reference-Guide-2023-August-update-1.pdf.
24. Яблонский А.И. Модели и методы исследования науки. М.: Эдиториал УРСС, 2001. 398 с.
25. McVeigh M.E, Mann S.J. The journal impact factor denominator: Defining citable (counted) items // JAMA. 2009. V. 302, No 10. P. 1107–1109. <https://doi.org/10.1001/jama.2009.1301>.
26. Miranda R., Garcia-Carpintero E. Overcitation and overrepresentation of review papers in the most cited papers // J. Inform. 2018. V. 12, No 4. P. 1015–1030. <https://doi.org/10.1016/j.joi.2018.08.006>.
27. Aksnes D.W. Characteristics of highly cited papers // Res. Eval. 2003. V. 12, No 3. P. 159–170. <https://doi.org/10.3152/147154403781776645>.
28. Smith R. Journal accused of manipulating impact factor // BMJ. 1997. V. 314. Art. 461. <https://doi.org/10.1136/bmj.314.7079.461d>.
29. Macdonald S. The gaming of citation and authorship in academic journals: a warning from medicine // Soc. Sci. Inf. V. 61, No 4. P. 457–480. <https://doi.org/10.1177/05390184221142218>.
30. Писляков В.В. Самоцитирование и его влияние на оценку научной деятельности: обзор литературы. Часть II // Научн. и техн. библ. 2022. № 3. С. 85–104. <https://doi.org/10.33186/1027-3689-2022-3-85-104>.

Поступила в редакцию 4.08.2024

Принята к публикации 24.09.2024

Елизаров Александр Михайлович, доктор физико-математических наук, профессор кафедры цифровой аналитики и технологий искусственного интеллекта Института информационных технологий и искусственного интеллекта

Казанский (Приволжский) федеральный университет

ул. Кремлевская, д. 18, г. Казань, 420008, Россия

E-mail: amelizarov@gmail.com

Писляков Владимир Владимирович, кандидат физико-математических наук, заместитель директора библиотеки

Национальный исследовательский университет «Высшая школа экономики»

Покровский б-р, д. 11, г. Москва, 109028, Россия

E-mail: pislyakov@hse.ru

ORIGINAL ARTICLE

doi: 10.26907/2541-7746.2024.4.485-498

Journal Impact Factor: Mathematical and Statistical Properties*A.M. Elizarov^{a*}, V.V. Pislyakov^{b**}*^a*Kazan Federal University, Kazan, 420008 Russia*^b*HSE University, Moscow, 109028 Russia*E-mail: **amelizarov@gmail.com*, ***pislyakov@hse.ru*

Received August 4, 2024; Accepted September 24, 2024

Abstract

Using a commonly agreed definition of the journal impact factor, its various applications in the mathematical modeling of scientific research were investigated. Special attention was paid to the impact factor's properties related to the disciplinary structure of science, the limitations of source data in existing databases, statistical reliability, the relevance of document types, and the effect of outliers. The potential bias resulting from journal self-citation practices was discussed.

Keywords: bibliometric indicators, impact factor, Web of Science, Journal Citation Reports, journal self-citation

Conflicts of Interest. The authors declare no conflicts of interest.

Figure Captions

Fig. 1. Journals in the “Medicine, Research and Experimental” category. Above: journal impact factors with confidence intervals. Below: journal rankings based on the journal impact factors with confidence intervals. Source: [17].

Fig. 2. Impact factor trend for “Acta Crystallographica A”, 1997–2013. Source: JCR.

References

1. ExLibris. Ulrichsweb: General FAQ. 2018. URL: [https://knowledge.exlibrisgroup.com/ Ulrich's/Product_Documentation/Frequently_Asked_Questions_\(FAQs\)/Ulrichsweb](https://knowledge.exlibrisgroup.com/ Ulrich's/Product_Documentation/Frequently_Asked_Questions_(FAQs)/Ulrichsweb).
2. Curcic D. Number of Academic Papers Published Per Year. 2023. URL: <https://wordrated.com/number-of-academic-papers-published-per-year/>.
3. Moskaleva O., Pislyakov V., Sterligov I., Akoev M., Shabanova S. Russian Index of Science Citation: Overview and review. *Scientometrics*, 2018, vol. 116, no. 1, pp. 449–462. <https://doi.org/10.1007/s11192-018-2758-y>.
4. Pislyakov V.V. Impact factor. The Great Russian Encyclopedia: A Science and Education Website. URL: <https://bigenc.ru/c/impakt-faktor-c87b20/?v=8838250>. (In Russian)
5. Garfield E. Citation indexes for science: A new dimension in documentation through association of ideas. *Science*, 1955, vol. 122, no. 3159, pp. 108–111. <https://doi.org/10.1126/science.122.3159.108>.

6. Garfield E., Sher I.H. New factors in the evaluation of scientific literature through citation indexing. *Am. Doc.*, 1963, vol. 14, no. 3, pp. 195–201.
<https://doi.org/10.1002/asi.5090140304>.
7. Gross P.L.K., Gross E.M. College libraries and chemical education. *Science*, 1927, vol. 66, no. 1713, pp. 385–389. <https://doi.org/10.1126/science.66.1713.385>.
8. Hackh I. The periodicals useful in the dental library. *Bull. Med. Libr. Assoc.*, 1936, Sept., vol. 25, nos. 1–2, pp. 109–112.
9. Raisig L.M. Mathematical evaluation of the scientific serial. *Science*, 1960, vol. 131, no. 3411, pp. 1417–1419. <https://doi.org/10.1126/science.131.3411.1417>.
10. Pislyakov V.V. Bibliometric indicators in the Clarivate databases. In: Akoev M.A., Markusova V.A., Moskaleva O.V., Pislyakov V.V. *Rukovodstvo po naukometrii: indikatory razvitiya nauki i tekhnologii* [Handbook for Scientometrics: Science and Technology Development Indicators]. 2 ed. Yekaterinburg, Izd. Ural. Univ., 2021, pp. 177–220.
<https://doi.org/10.15826/B978-5-7996-3154-3>. (In Russian)
11. Zibareva I.V., Pislyakov V.V., Teplova T.N., Nefedov O.M. Bibliometric analysis of the journal *Uspekhi Khimii* (*Russian Chemical Reviews*). *Herald Russ. Acad. Sci.*, 2008, vol. 78, no. 3, pp. 247–256. <https://doi.org/10.1134/S101933160803012X>. (In Russian)
12. Archambault É., Larivière V. History of the journal impact factor: Contingencies and consequences. *Scientometrics*, 2009, vol. 79, no. 3, pp. 635–649.
<https://doi.org/10.1007/s11192-007-2036-x>.
13. Glänzel W., Moed H.F. Journal impact measures in bibliometric research. *Scientometrics*, 2002, vol. 53, no. 2, pp. 171–193. <https://doi.org/10.1023/A:1014848323806>.
14. Garfield E., Sher I.H. *Genetics Citation Index: Experimental Citation Indexes to Genetics with Special Emphasis on Human Genetics*. Philadelphia, PA, Inst. Sci. Inf., 1963. xxvii, 864 p.
15. Pudovkin A.I., Garfield E. Rank normalization of impact factors will resolve Vanclay's dilemma with TRIF. *Scientometrics*, 2012, vol. 92, no. 2, pp. 409–412.
<https://doi.org/10.1007/s11192-012-0634-8>.
16. Schubert A., Glänzel W. Statistical reliability of comparisons based on the citation impact of scientific publications. *Scientometrics*, 1983, vol. 5, no. 1, pp. 59–73.
<https://doi.org/10.1007/BF02097178>.
17. Greenwood D.C. Reliability of journal impact factor rankings. *BMC Med. Res. Methodol.*, 2007, vol. 7, art. 48. <https://doi.org/10.1186/1471-2288-7-48>.
18. Bornmann L. Confidence intervals for Journal Impact Factors. *Scientometrics*, 2017, vol. 111, no. 3, pp. 1869–1871. <https://doi.org/10.1007/s11192-017-2365-3>.
19. Stern D.I. Uncertainty measures for economics journal impact factors. *J. Econ. Lit.*, 2013, vol. 51, no. 1, pp. 173–189. <https://doi.org/10.1257/jel.51.1.173>.
20. Vanclay J.K. Impact factor: Outdated artefact or stepping-stone to journal certification? *Scientometrics*, 2012, vol. 92, no. 2, pp. 211–238.
<https://doi.org/10.1007/s11192-011-0561-0>.
21. Zitt M. The journal impact factor: Angel, devil, or scapegoat? A comment on J.K. Vanclay's article 2011. *Scientometrics*, 2012, vol. 92, no. 2, pp. 485–503.
<https://doi.org/10.1007/s11192-012-0697-6>.
22. Bar-Ilan J. Journal report card. *Scientometrics*, 2012, vol. 92, no. 2, pp. 249–260.
<https://doi.org/10.1007/s11192-012-0671-3>.
23. Clarivate. Journal Citation Reports™ Reference Guide. 2023. URL:
https://clarivate.com/wp-content/uploads/dlm_uploads/2023/08/JCR-Reference-Guide-2023-August-update-1.pdf.
24. Yablonsky A.I. *Modeli i metody issledovaniya nauki* [Models and Methods for the Study of Science]. Moscow, Editorial URSS, 2001. 398 p. (In Russian)

25. McVeigh M.E., Mann S.J. The journal impact factor denominator: Defining citable (counted) items. *JAMA*, 2009, vol. 302, no. 10, pp. 1107–1109.
<https://doi.org/10.1001/jama.2009.1301>.
26. Miranda R., Garcia-Carpintero E. Overcitation and overrepresentation of review papers in the most cited papers. *J. Inform.*, 2018, vol. 12, no. 4, pp. 1015–1030.
<https://doi.org/10.1016/j.joi.2018.08.006>.
27. Aksnes D.W. Characteristics of highly cited papers. *Res. Eval.*, 2003, vol. 12, no. 3, pp. 159–170. <https://doi.org/10.3152/147154403781776645>.
28. Smith R. Journal accused of manipulating impact factor. *BMJ*, 1997, vol. 314, art. 461.
<https://doi.org/10.1136/bmj.314.7079.461d>.
29. Macdonald S. The gaming of citation and authorship in academic journals: A warning from medicine. *Soc. Sci. Inf.*, 2022, vol. 61, no. 4, pp. 457–480.
<https://doi.org/10.1177/05390184221142218>.
30. Pislyakov V.V. Self-citation and its impact on research evaluation: Literature review. Part II. *Nauchn. Tekh. Bibl.*, 2022, no. 3, pp. 85–104.
<https://doi.org/10.33186/1027-3689-2022-3-85-104>. (In Russian)

Для цитирования: Елизаров А.М., Писляков В.В. Журнальный импакт-фактор: математические и статистические свойства // Учен. зап. Казан. ун-та. Сер. Физ.-матем. науки. 2024. Т. 166, кн. 4. С. 485–498.
<https://doi.org/10.26907/2541-7746.2024.4.485-498>.

For citation: Elizarov A.M., Pislyakov V.V. Journal impact factor: Mathematical and statistical properties. *Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-Matematicheskie Nauki*, 2024, vol. 166, no. 4, pp. 485–498.
<https://doi.org/10.26907/2541-7746.2024.4.485-498>. (In Russian)