

ИГРА В ЦЫФИРЬ,
или как теперь оценивают труд ученого
(сборник статей о библиометрике)

Москва
Издательство МЦНМО
2011

УДК 01
ББК 78.5
И27

И27 **Игра в цифры**, или как теперь оценивают труд ученого (сборник статей о библиометрике). — М.: МЦНМО, 2011. — 72 с.: ил.

ISBN 978-5-94057-771-3

Настоящий сборник содержит статьи математиков, статистиков и биологов, содержащие критическую оценку широко распространенных теперь библиометрических методов оценки научной деятельности. Авторы показывают, к каким неверным выводам может приводить использование таких численных показателей, как импакт-фактор и индекс цитируемости. Основное место в сборнике занимает отчет «Статистики цитирования», подготовленный по инициативе Международного математического союза. К нему добавлен ряд других статей на эту же тему.

ББК 78.5

Электронная версия книги доступна по адресу
<http://www.mccme.ru/free-books/bibliometric.pdf>

ISBN 978-5-94057-771-3

Оглавление

От Международного математического союза	4
Р. Адлер, Дж. Эвинг, П. Тейлор. Статистики цитирования	6
П. А. Лоуренс. Потерянное при публикации: как измерение вредит науке	39
Ф. Кемпбелл. Бегство от импакт-фактора	46
Д. Арнольд, К. Фаулер. Гнусные цифры	52
Л. Д. Фаддеев. Из интервью журналу «Эксперт»	63
Приложение. Из ТИПОВОЙ МЕТОДИКИ оценки результативности деятельности научных организаций	65
Дополнительная литература	68
Сведения об авторах	70

От Международного математического союза

Уважаемые коллеги,

Перед вами первый специальный выпуск IMU-Net. Сегодня Международный математический союз (IMU) публикует важный документ «Статистики цитирования», на который мы хотим обратить ваше внимание.

В IMU-Net 24 (июль 2007 г.) было объявлено о создании комитета по «Количественной оценке исследований». Комитету было предложено исследовать различные стороны импакт-факторов и аналогичных статистик, основанных на цитированиях. Создание комитета было совместным действием Исполкома Международного математического союза, Международного совета по промышленной и прикладной математике (ICIAM) и Института математической статистики (IMS). Комитет состоял из следующих членов:

- Джон Эвинг (Провиденс, США), председатель, предложен IMU;
- Роберт Адлер (Хайфа, Израиль), предложен IMS;
- Питер Тейлор (Мельбурн, Австралия), предложен ICIAM.

Задачи, поставленные перед комитетом, изложены в Интернете по адресу* <http://www.mathunion.org/Publications/2007/Charge-ComOnQuantAssessmRes070521.pdf>

Источник: IMU-Net 29b: June 2008, <http://www.mathunion.org/imu-net/archive/2008/imu-net-29b/>. Специальный выпуск, посвященный «Статистикам цитирования» — отчету, анализирующему импакт-факторы и аналогичные статистики, основанные на цитированиях.

* В частности, предлагалось следующее.

1. Выяснить, до какой степени импакт-фактор, предложенный ISI (Institute of Science Information. О нем см. в тексте отчета. — *Прим. ред.*) является значимым индикатором качества исследователя, отдела или какого-либо учреждения в области статистики или математики.

2. Определить, до какой степени ISI-импакт-фактор может быть использован для сравнения качества исследований в математике с качеством исследований в других дисциплинах.

3. Выяснить, имеются ли в нахождении ISI-импакт-фактора какие-либо предположения по отношению к языку, месту, объему, происхождению или разделу науки публикуемого исследования, а также междисциплинарности работы.

4. Рассмотреть эти же вопросы для некоторой выборки других недавно предложенных методов измерения, основанных на данных цитирования.

5. Предложить подходящие замены для таких методов, основанных на данных цитирования.

6. Указать на возможные опасности или преимущества того влияния, которое широкое распространение использования импакт-факторов и схожих методов измерения может иметь на поведение, связанное с публикациями, принятие на работу, баланс между научными дисциплинами и т. п.

Предполагалось, что комитет будет рассматривать эти задачи с помощью обзора и обсуждения практик научного исследования, принятых в настоящее время, а также широкого круга литературы об использовании цитирований для оценки исследований. Отчет комитета, написанный с точки зрения математической статистики, был представлен исполкомам IMU, ICIAM и IMS, и все они одобрили его. Эти три организации делают сегодня отчет «Статистики цитирования» достоянием общественности.

Отчет размещен по следующему адресу*: <http://www.mathunion.org/Publications/Report/CitationStatistics>.

Заявление для прессы было сегодня распространено среди журналистов.

Предпринятые усилия были вызваны многочисленными запросами от стран — членов IMU, математических обществ, крупных математических учреждений и отдельных ученых, которые сообщали о возрастающем использовании импакт-факторов и других индикаторов, основанных на цитировании (а также и о злоупотреблениях ими), для измерения качества исследований отдельных ученых, департаментов или целых организаций.

IMU предлагает читателям IMU-Net не только прочитать этот отчет, но и распространить его среди администраторов и лиц, принимающих решения, которые вовлечены в процесс оценки качества исследований, чтобы довести до них точку зрения математической науки. IMU, ICIAM и IMS пришли к согласию, что для целей возможно более широкого распространения те журналы, бюллетени и аналогичные издания, которые заинтересованы в публикации этого отчета, получают не эксклюзивное право опубликовать его в одном из своих выпусков. Обратитесь, пожалуйста, к журналам/бюллетеням, с которыми вы связаны, и предложите им публикацию отчета «Статистики цитирования».

Все три организации, представляющие мировое сообщество теоретической, прикладной и промышленной математики и статистики, надеются, что тщательный анализ и рекомендации этого отчета будут рассмотрены лицами, принимающими решения и использующими для оценки исследований данные цитирований.

С наилучшими пожеланиями,

Л. Ловаш,
Президент IMU

* См. русский перевод в наст. изд., с. 6—38. — Прим. ред.

Статистики цитирования

Роберт Адлер, Джон Эвинг, Питер Тейлор

Доклад Международного математического союза (IMU) в сотрудничестве с Международным советом промышленной и прикладной математики (ICIAM) и Институтом математической статистики (IMS)

Настоящий доклад посвящен использованию и злоупотреблениям данными цитирований при оценке научных исследований. Сейчас все более распространенной становится идея о том, что оценка исследовательской деятельности должна осуществляться с помощью «простых и объективных» методов. Сами эти «простые и объективные» методы обычно интерпретируются как *библиометрические*, то есть как данные цитирований и связанные с ними статистики. Существует убеждение, что статистики цитирований по сути своей более точны, поскольку они оперируют числами, а не сложными суждениями, и, следовательно, позволяют обходить субъективность экспертной оценки. Но это убеждение является необоснованным.

- Опора на статистические данные не является более точной, если эти данные неправильно используются. Данные статистики и в самом деле могут вводить в заблуждение, если их неправильно используют или неправильно понимают. Представляется, что интерпретация и обоснованность статистик цитирований в современной библиометрике основывается по большей части на субъективном опыте и интуиции.

- Числа вроде бы «объективны», но их объективность может быть иллюзорной. Оценка смысла цитирования может быть даже более субъективной, чем экспертная оценка. Поскольку в области цитирований субъективный характер оценки менее очевиден, те, кто использует данные цитирований, имеют меньше шансов осознать их ограниченность.

- Использование только данных цитирований дает в лучшем случае неполное, а зачастую поверхностное понимание научного исследования — понимание, пригодное только тогда, когда оно под-

крепляется другими оценками. Числа по сути отнюдь не лучше, чем разумное суждение.

Использование данных цитирований для оценки научных исследований в конечном счете означает использование основанных на цитатах статистик для ранжирования объектов — журналов, статей, людей, программ и дисциплин. Статистические инструменты, используемые для ранжирования этих объектов, часто неправильно понимаются и неправильно используются.

- Для ранжирования журналов наиболее часто используется импакт-фактор. Это простое среднее, выводимое из распределения цитирований некоторой совокупности статей в журнале. Среднее улавливает только небольшое количество информации об этом распределении и является довольно грубой статистикой. Кроме того, при оценке журналов с помощью цитирований обнаруживается много сопутствующих факторов, так что любое сравнение журналов по импакт-факторам требует большой осторожности. Использовать лишь только импакт-фактор при оценке журнала — это все равно, что при оценке здоровья человека учитывать только его вес.

- Что касается статей, то вместо того, чтобы при сравнении отдельных статей полагаться на фактическое количество ссылок на них, люди часто заменяют его импакт-фактором журналов, в которых опубликованы статьи. Они считают, что более высокий импакт-фактор должен означать более высокий уровень цитирования. Но зачастую это совсем не так! Это широко распространенное неправильное использование статистических данных, с которым следует бороться, когда и где бы оно ни встретилось.

- При сопоставлении отдельных ученых сравнивать полные списки цитирований бывает затруднительно. Как следствие, были предприняты попытки найти простые статистики, которые выражали бы всю сложность списка цитирований ученого одним числом. Наиболее заметным из них является h -индекс, который, похоже, становится все более популярным. Но даже эпизодические проверки h -индекса и его вариантов показывают, что все они являются наивной попыткой решить сложную проблему анализа списков цитирований. Они охватывают лишь небольшой объем информации о распределении цитирований ученого, упуская из внимания важнейшие данные, определяющие оценку научного исследования.

Обоснованность таких статистик, как импакт-фактор и h -индекс, не является ни хорошо понятой, ни хорошо изученной. Связь этих статистик с качеством научных исследований иногда устанавли-

ливаются на основе «опыта». Оправданием для тех, кто полагается на них, является их «доступность». Немногочисленные попытки исследовать эти статистики были направлены исключительно на демонстрацию их корреляции с некоторыми другими мерами качества, а не на определение того, как лучше всего можно получить полезную информацию из данных цитирований. Мы не отбрасываем статистики цитирований как инструмент оценки качества научных исследований: данные цитирований и статистики могут нести вполне ценную информацию. Мы допускаем, что оценки должны быть практичными, и по этой причине легко получаемые статистики цитирований почти наверняка будут частью всего процесса. Вместе с тем, данные цитирований дают лишь ограниченное и неполное представление о качестве научных исследований, а статистики, полученные на основе данных цитирований, порой неправильно понимаются и используются. Научные исследования слишком важны, чтобы измерять их ценность только одним грубым инструментом.

Мы надеемся, что те, кто связан с оценками, познакомятся с комментариями и деталями этого доклада, чтобы не только осознать ограниченность статистик цитирований, но и понять, как их лучше использовать. Если мы устанавливаем высокие стандарты проведения научных исследований, то мы, безусловно, должны установить столь же высокие стандарты для оценки их качества.

Объединенный комитет по количественной оценке исследований

*Роберт Адлер, Израильский технологический институт (Technion)
Джон Эвинг (председатель), Американское математическое общество
Питер Тейлор, Университет Мельбурна*

Из задач комитета

Стремление к большей прозрачности и подотчетности в академическом мире создало «культуру чисел», когда учреждения и отдельные лица полагают, что справедливые решения могут достигаться путем алгоритмической оценки некоторых статистических данных; будучи не в состоянии измерить качество (что является конечной целью), лица, принимающие решения, заменяют качество числами, которые они измерить могут. Эта тенденция требует комментариев от тех, кто профессионально «имеет дело с числами», — математиков и статистиков.

Введение

Научные исследования имеют важное значение. Они в значительной мере определяют прогресс в нашем современном мире и дают надежду на то, что мы можем решить некоторые из кажущихся неразрешимыми проблем, стоящих перед человечеством, — от окружающей среды до роста народонаселения. В связи с этим правительства и учреждения по всему миру осуществляют значительную финансовую поддержку научных исследований. Естественно, они хотят знать, разумно ли инвестируются их деньги; они хотят оценить качество научных исследований, за которые они платят, чтобы принимать обоснованные решения относительно будущих инвестиций.

Это явление далеко не ново: люди оценивали научные исследования на протяжении многих лет. Новым является убеждение, что хорошая оценка должна быть «простой и объективной» и что это может быть достигнуто прежде всего на основе подсчетов (статистик) по данным цитирований, а не на основе сочетания разных методов, включающих экспертные оценки самих ученых. Эта точка зрения предельно ясно выражена в начальном абзаце одного из недавних докладов:

«После следующего, завершающегося в 2008 году, цикла правительство намерено заменить нынешний метод определения качества университетских научных исследований — UK Research Assessment Exercise (RAE). В центре внимания новой системы будут подсчеты, а не экспертные оценки, и ожидается, что основной индекс качества в этой системе будет библиометрическим (использующим подсчет журнальных статей и их цитирований)» [Evidence Report 2007, с. 3].

Те, кто ратует за объективность в такой простой форме, полагают, что научные исследования слишком важны, чтобы полагаться на субъективные суждения. Они считают, что основанные на цитированиях подсчеты вносят ясность в процесс ранжирования и устраняют неясности, присущие оценкам другого вида. Они полагают, что тщательно подобранные метрики (= системы подсчетов) являются независимыми и свободны от систематических ошибок. Но самое главное, они считают, что такие метрики позволяют сравнивать все составляющие научных исследований — журналы, статьи, людей, программы и даже целые дисциплины — просто, эффективно и без использования субъективных экспертных оценок.

Но такая вера в точность, независимость и эффективность подсчетов неуместна.

- Во-первых, точность этих показателей является иллюзорной. Общеизвестно, что статистики могут лгать, когда они неправильно используются. Злоупотребления статистиками цитирований получили широкое распространение и являются вопиющими. Несмотря на неоднократные попытки предостеречь от таких злоупотреблений (например, от злоупотребления импакт-фактором), правительства, учреждения, да и сами ученые продолжают делать необоснованные или даже ложные заключения из неправильно использованных статистик цитирований.

- Во-вторых, использование метрик, основанных исключительно на цитированиях, попросту заменяет один вид оценок другим. Вместо субъективной экспертной оценки мы получаем субъективную интерпретацию смысла цитирования. Те, кто призывают полагаться исключительно на подсчеты, основанные на цитированиях, неявно предполагают, что каждое цитирование означает одно и то же в отношении цитируемого научного исследования — его «импакт» (влияние). Это предположение не доказано и, весьма вероятно, неверно.

- В-третьих, хотя статистики дают нам ценную информацию для понимания мира, в котором мы живем, они обеспечивают только его частичное понимание. В современном мире иногда модно провозглашать мистическую веру в то, что численные измерения превосходят другие формы понимания. Те, кто пропагандирует использование статистик цитирований в качестве замены более полного понимания научного исследования, неявно придерживаются таких убеждений. Мы должны не только использовать статистики *правильно* — мы должны использовать их *разумно*.

Мы спорим не с попыткой оценивать исследование, но, скорее, с требованием, что такие оценки должны опираться преимущественно на «простой и объективный» показатель, основанный на цитировании. Это требование часто интерпретируется как необходимость поиска легких для подсчета чисел, которые выстроят в ряд публикации, или людей, или программы. Исследование обычно имеет множество целей, как непосредственных, так и более отдаленных, и поэтому его значимость, естественно, должна оцениваться при помощи не одного, а нескольких критериев. Математики знают, что имеется много предметов, как реальных, так и абстрактных, которые

не могут быть просто упорядочены в том смысле, что каждые два из них можно сравнить. Сравнение часто требует более сложного анализа, который иногда так и не позволяет решить, какой из двух предметов «лучше». Правильным ответом на вопрос «Что лучше?» иногда является: «Смотря как посмотреть!»

Предложение использовать для оценки качества исследования сразу несколько методов делалось и ранее (см., например, [Martin 1996] или [Carey—Cowling—Taylor 2007]). Публикации могут быть оценены множеством способов, не только при помощи цитирований. Мерилом профессионального качества исследователя могут быть такие знаки оценки его заслуг, как количество приглашений, членство в редакциях и награды. В некоторых дисциплинах и в некоторых странах определенную роль может играть субсидирование грантами. Важной составляющей общей оценки может быть и экспертная оценка, мнение коллег-ученых. (Мы не должны отказываться от экспертной оценки только потому, что она иногда страдает от предвзятости; во всяком случае, у нас не меньше оснований отказаться от статистических данных, связанных с цитированием, если учесть случаи их неправильного применения.) Это небольшой пример комбинации критериев, на основании которых может быть сделана оценка. Много путей ведет к хорошей оценке, и их относительная значимость меняется в зависимости от дисциплины. Несмотря на это, «объективные» статистические данные, основанные на цитированиях, часто становятся наиболее предпочтительным методом оценки. Соблазн простого способа и несложных чисел (еще лучше — одного числа), кажется, побеждает здравый смысл и трезвый расчет.

Данный доклад написан учеными-математиками, чтобы указать на неправильное применение статистических данных в оценке научного исследования. Естественно, неправильное применение иногда касается оценки исследований в самой математике, и это — одна из причин для написания этого доклада. Особая традиция цитирования в математике, состоящая в низком количестве ссылок на журналы, статьи и авторов, делает эту науку особенно уязвимой перед злоупотреблениями статистиками, основанными на цитированиях. Мы полагаем, однако, что *все* ученые, так же как и широкая

Исследование обычно имеет множество целей, как непосредственных, так и более отдаленных, и потому его значимость, естественно, должна оцениваться при помощи не одного, а нескольких критериев.

общественность, должны стремиться использовать правильные научные методы при оценке исследований.

Кое-кто в научном сообществе обошелся бы вообще без статистик, основанных на цитированиях, цинично реагируя на прошлые злоупотребления, но это будет означать отказ от ценного инструмента. Статистики, основанные на цитировании, *могут* играть роль в оценке исследования при условии, что они используются правильно, интерпретируются с осторожностью и составляют только часть процесса. Ссылки дают информацию о журналах, статьях и людях. Мы не хотим спрятать эту информацию; мы хотим пролить на нее свет.

Именно в этом и состоит цель настоящего доклада. Первые три раздела посвящены способам, при помощи которых данные цитирований могут быть использованы (и неправильно использованы) для оценки журналов, статей и людей. В следующем разделе обсуждаются разнообразные значения ссылок и вытекающие из них ограничения на статистики, основанные на цитированиях. Последний раздел рекомендует, как разумно использовать статистики, и призывает при получении оценок наряду со статистиками, основанными на цитировании, использовать и другие данные, даже если от этого оценки становятся менее простыми.

«Все должно быть сделано как можно более простым, но не чересчур простым», — сказал однажды Альберт Эйнштейн.¹⁾ Этот совет одного из выдающихся ученых мира особенно уместен при оценке научных исследований.

Ранжирование журналов: импакт-фактор²⁾

Импакт-фактор был создан в 1960-х годах как способ измерять ценность журналов путем вычисления среднего числа цитирований на статью за определенный промежуток времени [Garfield 2005]. Среднее число вычисляется на основе данных, собранных *Thomson Scientific* (бывший Институт научной информации) и публикующихся в *Journal Citation Reports*. *Thomson Scientific* извлекает ссылки более чем из 9 000 журналов, добавляя информацию о каждой статье и ее ссылках в собственную базу данных каждый год [THOMSON: SELECTION]. Используя данную информацию, можно рассчитать, как часто конкретная публикация цитируется в последующих статьях, которые были опубликованы в журналах, включенных в перечень индексируемых публикаций. (Заметим, что *Thomson Scientific* индек-

сирует менее половины математических журналов, охватываемых *Mathematical Reviews* и *Zentralblatt*, двумя ведущими реферативными журналами по математике.³⁾

Величина импакт-фактора для данного журнала в данном году вычисляется путем расчета среднего количества ссылок на статьи в этом журнале, опубликованные в течение двух предыдущих лет, причем принимаются во внимание все ссылки в статьях, опубликованные в этом конкретном году (строго в рамках перечня журналов, индексируемых *Thomson Scientific*). Если импакт-фактор журнала составляет 1,5 в 2007 году, то это означает, что в среднем статьи, опубликованные в нем в течение 2005 и 2006 годов, были 1,5 раза процитированы в статьях 2007 года, изданных журналами из перечня индексируемых журналов.

Thomson Scientific сам использует импакт-фактор как один из факторов при отборе журналов, которые индексируются [THOMSON: SELECTION]. С другой стороны, он поддерживает более широкое использование импакт-фактора, чтобы сравнивать журналы.

«Импакт-фактор может быть учтен при решении вопроса комплектации библиотек, предоставляя администратору информацию о журналах, которые уже есть в наличии, и журналах, для которых решается вопрос об их приобретении. Эти данные должны рассматриваться наряду со сведениями о стоимости и тираже, для того чтобы принимать рациональные решения о покупке журналов» [THOMSON: IMPACT FACTOR].

Многие авторы отмечали, что нельзя судить об академической ценности журнала, используя одни только данные о ссылках, и авторы настоящей статьи полностью согласны с этим. В дополнение к этому общему наблюдению импакт-фактор подвергся критике и по другим причинам. (См. [Seglen 1997], [Amin—Mabe 2000], [Monastersky 2005], [Ewing 2006], [Adler 2007] и [Hall 2007].)

(i) Определение импакт-фактора как среднего числа не совсем правильно. Поскольку многие журналы содержат редко цитируемые вспомогательные публикации, такие как письма или редакционные статьи, эти публикации не учитываются в знаменателе импакт-фактора. С другой стороны, пусть и нечасто, но эти публикации все же цитируются, и эти ссылки *учитываются* в числителе. Следовательно, импакт-фактор — не совсем среднее количество ссылок на статью. Если в журналах помещается большое количество таких «вспомогательных» публикаций, это отклонение может быть значи-

тельным. Во многих областях, включая математику, это отклонение минимально.

(ii) Двухлетний период, используемый при определении импакт-фактора, предназначался для того, чтобы сделать статистический показатель современным [Garfield 2005]. Для некоторых областей, таких как биомедицинские науки, это целесообразно, потому что на большинство опубликованных статей ссылаются вскоре после публикации. В других областях, таких как математика, большинство ссылок выходит за рамки двухлетнего периода. Анализ совокупности из более 3 миллионов недавних ссылок в математических журналах (база данных *Math. Reviews*), показывает, что примерно 90 % ссылок на журнал выходит за пределы этого двухлетнего окна. Следовательно, импакт-фактор основывается всего лишь на 10 % ссылок и игнорирует подавляющее большинство ссылок.⁴⁾

Кривые цитирования

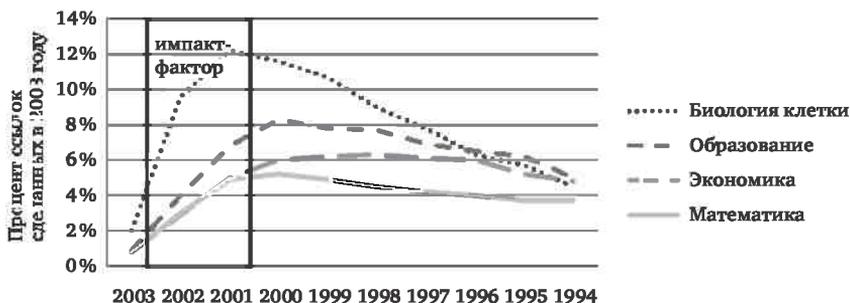
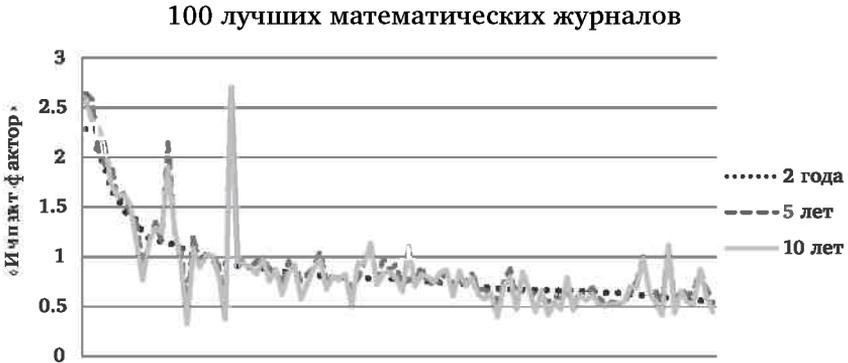


График показывает возраст ссылок из статей, опубликованных в 2003 году, и охватывает четыре различные области. Ссылки на статьи, опубликованные в 2001—2002 годах, — это те, которые вносят вклад в импакт-фактор, а все остальные ссылки не имеют отношения к импакт-фактору. Данные взяты из *Thomson Scientific*.

Ведет ли ли двухлетний интервал импакт-фактора к недоразумениям? Для математических журналов ответ на этот вопрос не так однозначен. *Thomson Scientific* вычисляет пятилетние импакт-факторы, которые, как он указывает, хорошо коррелируют с обычными (двухлетними) импакт-факторами [Garfield 1998]. Используя базу цитирований *Math. Reviews*, можно вычислить «импакт-факторы» (т. е. среднее число ссылок на 1 статью) для совокупности ста наи-

более цитируемых математических журналов, используя периоды в 2, 5 и 10 лет. Приведенная ниже диаграмма показывает, что 5- и 10-летние импакт-факторы в общем следуют 2-летнему импакт-фактору.



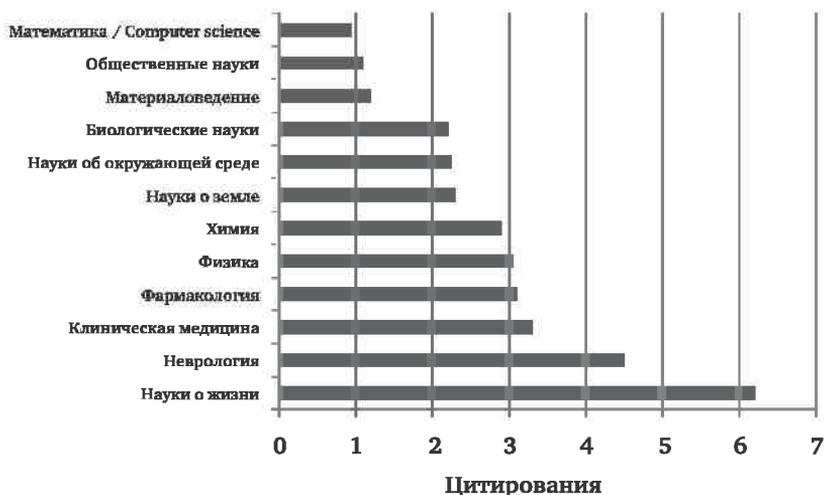
«Импакт-факторы» для 2, 5 и 10 лет для 100 математических журналов. Данные из базы цитирований Math. Reviews.

Единственным существенным выбросом является журнал, который в течение части этого времени не публиковал статьи. Менее значимые отклонения соответствуют журналам, которые публикуют относительно небольшое количество статей каждый год, и диаграмма просто отражает нормальную изменчивость в импакт-факторах для таких журналов. Очевидно, что изменение числа «заданных годов» при расчете импакт-фактора приводит к изменению рейтинга журналов, но эти изменения, как правило, умеренны, за исключением небольших журналов, где величина импакт-фактора также зависит от выбора «исходного года» (см. ниже).

(iii) Импакт-фактор значительно варьируется в зависимости от выбора дисциплины [Amin—Mabe 2000]. Это различие отчасти связано с наблюдением (ii): если в некоторых дисциплинах много ссылок случается за пределами двухлетнего окна, то импакт-факторы для журналов будут намного ниже. С другой стороны, различие отчасти возникает просто из-за того, что традиции цитирования отличаются от дисциплины к дисциплине и ученые ссылаются на статьи с различной частотой и по различным причинам. (Мы подробно остановимся на этом наблюдении позже, потому что смысл ссылки чрезвычайно важен.) Из этого следует, что, используя им-

пакт-факторы, невозможно никаким разумным способом сравнить два журнала, представляющие разные дисциплины.

Среднее количество ссылок на статью



Среднее количество ссылок на статью по различным дисциплинам, показывающее, что практики цитирования заметно различаются. Данные от Thomson Scientific [Amin—Mabe 2000].

(iv) Импакт-фактор может значительно изменяться от года к году, и эти изменения, как правило, больше для небольших журналов [Amin—Mabe 2000]. Например, для журналов, публикующих меньше чем 50 статей, среднее *изменение* импакт-фактора с 2002 по 2003 год составляло почти 50 %. Это, конечно, полностью ожидаемо, потому что объем выборки для небольших журналов весьма мал. С другой стороны, часто сравнивают журналы за фиксированный год, игнорируя высокую годовую вариабельность данных для небольших журналов.

(v) Журналы, которые публикуют статьи на других языках, кроме английского, скорее всего, получают меньше ссылок, поскольку большая часть научного сообщества не может читать (или не читает) их. Тем самым на импакт-фактор может влиять тип журнала, а не одно только качество. Журналы, которые публикуют, например, обзорные статьи, часто получают гораздо больше ссылок, чем журналы, которые этого не делают, и, следовательно, имеют бо-

лее высокие (иногда существенно более высокие) импакт-факторы [Amin—Mabe 2000].

(vi) Наиболее существенная критика импакт-фактора заключается в том, что не вполне ясен его смысл. При использовании импакт-фактора для сравнения двух журналов нет никакой априорной модели, в которой было бы исходно определено, что означает, что какой-то журнал «лучше». Единственная модель происходит от самого импакт-фактора — журнал с бóльшим импакт-фактором и считается лучше. При классическом статистическом подходе сначала определяется модель, затем формулируется гипотеза (неважно какая), а потом применяется статистическая процедура, в зависимости от результатов которой можно принять или опровергнуть выдвинутую гипотезу. Извлечение информации (а возможно, и построение модели) из данных является принятым способом статистического анализа, но в нашем случае не вполне ясно, как осмыслить полученную информацию. Каким образом импакт-фактор оценивает качество? Является ли он наилучшей статистической величиной для оценки качества? Что в точности измеряет импакт-фактор? (См. также наше дальнейшее обсуждение значения цитирования.) Отметим, что весьма немного известно о том, какой должна быть модель, описывающая качество журналов, и о том, как она могла бы быть связана с импакт-фактором.

Все шесть приведенных выше критических замечаний об импакт-факторе вполне валидны, но означают лишь то, что импакт-фактор недостаточно продуман, хотя и не бесполезен. Например, импакт-фактор может быть использован в качестве отправной точки для того, чтобы разбить все журналы на группы и затем с помощью других критериев составить более подробный рейтинг и убедиться в разумности деления на группы. Однако использование импакт-фактора для сравнительной оценки журналов требует большой осторожности. Например, импакт-фактор не может быть применен для сравнения журналов, относящихся к различным дисциплинам, а при их упорядочивании в соответствии с импакт-фактором необходимо учитывать тип самих журналов. Кроме того, следует уделять большое внимание годовым колебаниям импакт-фактора, особенно для небольших журналов, и иметь при этом в виду, что незначительные различия могут быть отнесены на счет чисто случайных колебаний. Важно осознавать, что в некоторых областях знания импакт-фактор может не полностью отражать все цитирования, так как, с одной стороны, учитываются не все журналы, а

с другой стороны, рассматриваемый период времени слишком короток. Статистические данные, основанные на более длительном периоде времени и большем количестве анализируемых журналов, могли бы повысить качество оценки. Наконец, цитируемость является лишь только одним способом оценки журналов и должна быть дополнена другой информацией (это и есть основная идея данной статьи).

Все эти соображения касаются, в сущности, любого упорядочивания на основе статистических данных. Необдуманная оценка журналов на основе их импакт-фактора за какой-либо конкретный год является некорректным использованием статистических данных. Следует отдать должное агентству *Thomson Scientific*, которое соглашается с этим утверждением и (в мягкой форме) предупреждает об этом всех тех, кто использует импакт-фактор:

«При оценке полезности журналов *Thomson Scientific* учитывает не только импакт-фактор, что рекомендуется делать и другим. Импакт-фактор не может быть использован без учета многочисленных показателей, влияющих на цитируемость, например среднего числа ссылок в одной статье. Импакт-фактор должен быть дополнен компетентной экспертной оценкой» [THOMSON: IMPACT FACTOR].

К сожалению, этот совет слишком часто игнорируется.

Ранжирование работ

Импакт-фактор и подобные ему статистические данные, связанные с цитируемостью, могут быть неверно истолкованы при определении рейтинга журналов. Однако существует более глубокое и более коварное заблуждение: использование импакт-фактора для сравнения отдельных работ, людей, исследовательских программ или даже целых областей знания. Эта проблема становится все более актуальной во многих странах и для многих дисциплин; она усугубляется принятыми недавно методами оценки национальных* исследований.

В определенном смысле, это явление не новое. К ученым часто обращаются с просьбой оценить чей-либо список публикаций, в ответ на что регулярно слышны такие комментарии, как «она публикуется в хороших журналах» или «большая часть его работ опубли-

* Выражение «национальный» (national) является общепринятым в англоязычной литературе по отношению к стране (государству) в целом. — *Прим. ред.*

кована в журналах низкого уровня». Такого рода оценки могут быть вполне разумными: качество журналов, в которых ученый обычно (или даже постоянно) публикует свои статьи, является одним из многочисленных показателей, которые могут быть использованы для общей оценки исследований данного ученого. Тем не менее, использование импакт-фактора усилило тенденцию приписывать свойства журнала каждой статье в нем (а также каждому автору статьи).

Thomson Scientific в неявном виде придерживается следующей позиции:

«Возможно, наиболее важным в недавно начавшемся использовании импакт-фактора является его применение для оценки научной деятельности. С помощью импакт-фактора можно составить довольно полное представление о престиже журналов, в которых публикуются ученые». [THOMSON: IMPACT FACTOR]

Вот некоторые примеры того, как интерпретируется эта рекомендация (по свидетельству математиков из разных частей света).

Пример 1. Недавно мой университет ввел новую классификацию журналов на основе Science Citation Index Core. Все журналы были разделены на три группы в зависимости от одного только их импакт-фактора. В список лидеров попало 30 журналов, среди которых нет ни одного математического. Второй список содержит 667 журналов, включая 21 математический журнал. Публикация в журнале из первого списка повышает поддержку исследований со стороны университета в три раза; публикация в журналах второго списка — в два раза. Публикация в основном списке приносит 15 баллов; публикация в любом журнале перечня Thomson Scientific приносит 10 баллов. Для повышения в должности требуется определенное минимальное количество баллов.

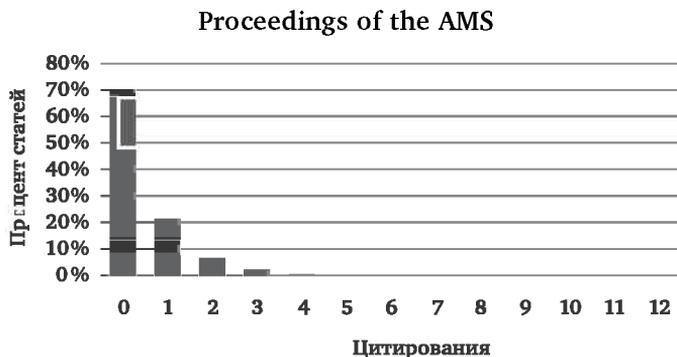
Пример 2. В моей стране постоянные сотрудники университетов аттестуются раз в шесть лет. Несколько подряд успешно пройденных аттестаций открывают возможность для академического роста. Помимо резюме, важнейшим фактором при оценке является рейтинг пяти опубликованных статей. В последние годы за каждую статью дается 3 очка, если она опубликована в журнале из первой трети списка Thomson Scientific, 2 очка, если из второй трети, и 1 очко, если статья опубликована в журнале из последней трети (деление журналов на эти три группы происходит на основе импакт-фактора).

Пример 3. Каждый сотрудник нашего факультета оценивается при помощи формулы, включающей в себя число статей (с поправками в случае соавторства), умноженное на импакт-фактор журналов, в которых они опубликованы. На этой формуле частично основано повышение в должности и прием на работу.

В этих примерах, как и во многих других рассказанных нам случаях, импакт-фактор, явным или неявным образом, используется для сравнения отдельных статей, а также их авторов: если импакт-фактор у журнала А больше, чем у журнала В, то, безусловно, статья из А должна быть лучше статьи из В, а автор статьи из А лучше автора статьи из В. В некоторых случаях это рассуждение распространяется и на оценку факультетов или даже целых областей знания.

Давно известно, что распределение числа цитирований по статьям в одном журнале крайне неравномерно и близко к так называемому степенному закону ([Seglen 1996], [Garfield 1987]). Последствия этого можно продемонстрировать на конкретном примере.

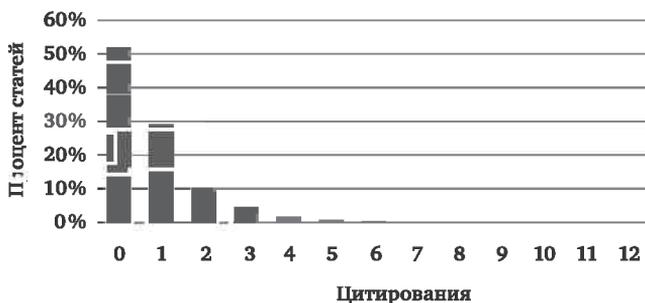
Ниже приводится распределение цитируемости по статьям в журнале *Proceedings of the American Mathematical Society* за период 2000—2004 г. *Proceedings* публикует короткие статьи, как правило, менее десяти страниц каждая. За этот период журналом была опубликована 2381 работа (около 15 000 страниц). На основе данных *Math. Reviews* о цитировании за 2005 год среднее количество цитирований одной статьи (то есть импакт-фактор) равно 0,434.



Журнал *Transactions of the AMS* публикует более длинные статьи, которые, как правило, значительно больше как по объему, так и по содержанию, чем статьи из *Proceedings*. За тот же период време-

ни *Transactions* опубликовал 1165 работ (более 25 000 страниц), с числом цитирований от 0 до 12. Среднее количество цитирований одной статьи равно 0,846, что примерно в два раза больше, чем для *Proceedings*.

Transactions of the AMS



Теперь предположим, что из двух математиков один опубликовал статью в *Proceedings*, а другой — в *Transactions*. При упомянутых выше подходах второй математик должен быть оценен выше первого, поскольку статья второго математика опубликована в журнале с большим импакт-фактором (причем в два раза большим!). Но насколько правомерно такое мнение? Действительно ли статьи из *Transactions of the AMS* в два раза лучше статей из *Proceedings*?

Когда мы утверждаем, что отдельные статьи из *Transactions* лучше (в смысле цитируемости) отдельной статьи из *Proceedings*, мы должны сравнивать не средние статьи в этих журналах, а соответствующие вероятности: какова вероятность того, что мы не правы? То есть какова вероятность того, что у случайно выбранной статьи из *Proceedings* по крайней мере столько же цитирований, сколько у случайно выбранной статьи из *Transactions*?

Элементарное вычисление дает величину 62%. Это означает, что мы неправы в 62% случаев, и случайно выбранная статья из *Proceedings* будет по крайней мере не хуже случайно выбранной статьи из *Transactions*, несмотря на то что импакт-фактор у *Proceedings* равен всего лишь половине импакт-фактора у *Transactions*! Таким образом, мы чаще ошибаемся, чем оказываемся правы. Большинству людей это кажется удивительным, хотя это всего лишь следствие

сильно асимметричного распределения цитируемости по статьям и слишком короткого интервала времени, используемого для вычисления импакт-фактора (что является причиной высокого процента нецитируемых работ)⁵⁾. Это показывает важность точного статистического анализа по сравнению с интуитивным представлением.

Описанная ситуация характерна для всех журналов; два выбранных нами журнала не являются какими-то особенными. (Например, *Journal of the AMS* за тот же период имеет импакт-фактор 2,63, что в шесть раз больше, чем у *Proceedings*. Однако в 32% случаев выбранная наугад статья из *Proceedings* в смысле цитирования ничуть не хуже статьи из *Journal*.)

Таким образом, хотя и неверно было бы утверждать, что импакт-фактор совсем ничего не говорит об отдельных статьях в журнале, он дает весьма расплывчатую информацию и может ввести в глубокое заблуждение.

Можно заключить, что мало разумного в рассуждениях, подобных представленным в трех приведенных выше примерах и использующих импакт-фактор как приближение к реальному количеству

Хотя и неверно было бы утверждать, что импакт-фактор совсем ничего не говорит об отдельных статьях в журнале, он дает весьма расплывчатую информацию и может ввести в глубокое заблуждение.

цитирований одной отдельной статьи. Высказывание утверждений, которые неверны в более чем половине (или хотя бы трети) случаев, безусловно, является некорректным методом проведения оценки.

Хотя и неверно было бы утверждать, что импакт-фактор совсем ничего не говорит об отдельных статьях в журнале, он дает весьма расплывчатую информацию и может ввести в глубокое заблуждение.

Коль скоро мы осознаем бессмысленность использовать импакт-фактор для оценки количества цитирований

отдельной статьи, не имеет смысла использовать импакт-фактор и для оценки авторов статей, исследовательских программ, которыми они занимаются, и тем более областей знаний, в которых они работают. Импакт-фактор, как и вообще средние показатели, является не слишком пригодной величиной для того, чтобы делать содержательные сравнения без привлечения дополнительной информации.

Конечно же, ранжирование людей совсем не то же самое, что ранжирование их статей. Однако если требуется оценить чьи-либо статьи, принимая цитируемость как меру качества, то следует начать с рассмотрения числа цитирований каждой статьи. Импакт-фактор журнала, в котором опубликована статья, не является адекватной заменой этого числа.

Ранжирование ученых

Хотя импакт-фактор — наиболее известная статистика, основанная на цитировании, недавно появились и другие показатели, которые сейчас активно популяризируются.

Вот три из них, предназначенные для оценки людей.

h-индекс: *h*-индекс данного ученого является наибольшим таким числом n , что у ученого есть n статей, на каждую из которых существует по крайней мере n ссылок. Это самая популярная из упомянутых выше статистических величин. Ее ввел Дж. Е. Хирш [Hirsch 2006] в попытке оценить числом научную продукцию ученого, приняв за основу «хвост» кривой распределения цитируемости его статей. Целью было ввести единую величину для оценки и количества цитирований, и их распределения.

m-индекс: *m*-индекс ученого определяется как отношение его *h*-индекса к числу лет, прошедших после первой публикации этого ученого. Этот показатель был предложен тем же Хиршем в упомянутой выше статье. Его идея заключается в том, чтобы компенсировать молодым ученым недостаток времени на то, чтобы опубликовать много работ и набрать большое количество цитирований.

g-индекс: *g*-индекс ученого определяется как наибольшее n , для которого n наиболее цитируемых работ автора в общей сложности цитируются по крайней мере n^2 раз. Этот показатель предложил Лео Эгг в 2006 году [Egghe 2006]. *h*-индекс не принимает во внимание тот факт, что некоторые из n наиболее цитируемых статей могут быть очень высоко цитируемыми. *g*-индекс призван исправить этот недостаток.

Существуют и другие индексы, и их много. В их числе модификации названных выше, в которых учитывается возраст работ или количество авторов ([Batista—Campitelli—Kinouchi—Martinez 2005], [Batista—Campitelli—Kinouchi 2006], [Sidiropoulos—Katsaros—Manolopoulos 2006]).

В своей работе Хирш пишет, что он предлагает *h*-индекс как «легко вычисляемую величину, которая дает оценку важности, значимости и масштаба влияния совокупного вклада ученого в науку» [Hirsch 2005, с. 5]. Далее он добавляет, что «этот показатель может служить полезным средством для сравнения различных лиц, конкурирующих за один и тот же ресурс, когда критерием оценки являются научные достижения».

Ни одно из этих утверждений не подтверждается убедительными доказательствами. В поддержку своего утверждения, что h -индекс является мерой важности и значимости совокупного вклада ученого в науку, Хирш приводит данные анализа h -индекса для выборки, состоящей из лауреатов Нобелевской премии (и, отдельно, членов Национальной академии). Он показывает, что члены этих групп, как правило, имеют высокий h -индекс. Можно заключить, что коль скоро ученый является лауреатом Нобелевской премии, он имеет высокий h -индекс. И все же без дополнительной информации мы вряд ли способны сделать заключение о том, что некто станет Нобелевским лауреатом или членом Национальной академии, если известно только то, что у него высокий h -индекс. А именно такая информация нужна, чтобы установить состоятельность h -индекса.

В своей статье Хирш также утверждает, что можно использовать h -индекс для сравнения двух ученых:

«Я утверждаю, что два человека, имеющие одинаковую величину h , сопоставимы с точки зрения их научного вклада, даже если их общее число статей или их общее число цитирований очень разные. И наоборот, из двух людей (одного и того же научного возраста) с одинаковым количеством работ или одинаковым количеством цитирований и сильно различающимися h -значениями тот, у которого h больше, вероятно, является ученым более высокого уровня» [Hirsch 2005, с. 1].

Эти утверждения не в ладах со здравым смыслом. (Представьте себе двух ученых, у каждого из которых есть по 10 работ и на каждую из них есть по 10 ссылок, но при этом у одного из них есть еще 90 работ, на каждую из которых ссылаются по 9 раз. Или, например, у одного ровно 10 работ, и на каждую из них ссылаются ровно по 10 раз, а у другого — ровно 10 работ, и на каждую из них ссылаются по 100 раз. Неужели кто-то и вправду будет считать этих людей эквивалентными?)⁶⁾

Хирш превозносит достоинства h -индекса, утверждая, что « h -индекс предпочтительнее других численных критериев, которые обычно используются для оценки научной деятельности исследователей...» [Hirsch 2005, с. 1], но не определяет смысл слова «предпочтительнее», равно как и не объясняет, почему *хотят* найти именно «критерий с одним числом».

Хотя данный подход и подвергался критике, но до настоящего момента ему не доставало серьезного анализа. Большая часть анализа состояла в демонстрации «сходящейся валидности», т. е. того,

что h -индекс хорошо коррелирует с другими измерениями публикаций и цитирования, такими как количество опубликованных работ или общее число цитирований. Такая корреляция не очень впечатляет, так как все эти величины являются функциями одной и той же основной сущности — самих публикаций. В одной интересной статье об h -индексе [Lehmann—Lackson—Lautrup 2006] проводится более тщательный анализ и показывается, что h -индекс (точнее говоря, m -индекс) «хуже», чем простое среднее число ссылок на работы. Отметим, что даже здесь авторы не в состоянии адекватно определить, что значит «хуже». При применении классической статистической парадигмы [Lehmann—Jackson—Lautrup 2006] h -индекс оказывается менее надежным, чем другие показатели.

Были разработаны разные варианты h -индекса, приспособленные для сравнения качества исследователей не только в рамках одной дисциплины, но также для сравнения качества исследований в различных областях ([Batista—Campitelli—Kinouchi 2006], [Molinan—Molinan 2008]). Некоторые полагают, что h -индекс может быть использован для сравнения различных институтов и их подразделений [Kinney 2007]. В основном все это — неимоверно наивные попытки описать весь сложный процесс цитирования одним числом. Действительно, основное преимущество этих новых показателей над обычными гистограммами цитирований состоит в том, что индекс избавляет нас от излишних деталей, что делает любых двух исследователей сравнимыми между собой. Однако даже на простых примерах видно, что отброшенная информация необходима, чтобы понимать сущность исследования. Несомненно, что именно понимание должно быть целью оценки исследований, а не только обеспечение того, чтобы любые два человека могли быть сравнимы.

Именно понимание должно быть целью оценки исследований, а не только обеспечение того, чтобы любые два человека могли быть сравнимы.

В некоторых случаях h -индекс или его варианты используются государственными органами, оценивающими исследования. Это не что иное, как использование данных не по назначению. К сожалению, слишком велик соблазн определить рейтинг каждого ученого, оперируя всего одним числом; может стать, что этим соблазнится и широкая общественность, которая и в более простых случаях оказывается неспособна понять, что может и что не может следовать из тех или иных статистических величин.

Смысл цитирования

Те, кто продвигает мысль об использовании статистики цитирований в качестве основной меры качества исследований, не дают ответа на главный вопрос: что означает цитирование? Они собирают большое количество данных о цитировании, обрабатывают их, получают из них статистические показатели, а затем утверждают, что сделанные на этой основе оценки «объективны». Однако мы можем делать заключения лишь на основании *интерпретации* статистических показателей, а интерпретация основана на понимании *смысла* цитирования, которое само по себе достаточно субъективно.

Как ни странно, в литературе, написанной в поддержку такого подхода, очень трудно найти четкие заявления о смысле цитирования.

«Идея, стоящая за индексацией цитирований, принципиально проста. Если признать, что ценность информации определяется теми, кто ее использует, что может быть лучше для оценки качества работы, чем измерение влияния, которое она производит на сообщество в целом. В этом случае наибольшее количество членов научного сообщества (т. е. тех, кто использует или цитирует исходный материал) определяет влияние или воздействие идеи и ее автора на наш объем знаний». [THOMSON: HISTORY]

«Хотя измерить числом качество отдельных ученых затруднительно, общее мнение состоит в том, что лучше публиковать больше, чем меньше, и что количество ссылок на работу (с учетом общих традиций цитирования в данной области) является адекватной мерой ее качества» [Lehman—Jackson—Lautrup 2006, с. 1003].

«Частота цитирования отражает качество журнала и пользу, которую из него извлекают...» [Garfield 1972, с. 535].

«Когда врач или специалист в области биомедицины цитирует журнальную статью, это указывает, что цитируемый журнал каким-то образом повлиял на него». [Garfield 1987, с. 7]

«Цитаты — признание интеллектуального долга» [THOMSON: FIFTY YEARS].

Важные для нас термины — «качество», «значение», «влияние», и «интеллектуальный долг». Термин «импакт» стал родовым понятием, описывающим придание цитированию какого-либо веса или смысла. Термин впервые возник в короткой статье, которую напи-

сал в 1955 году Е. Гарфилд, продвигая идею индекса цитирования. Он пишет:

«Таким образом, в случае весьма значимых работ индекс цитирования имеет количественное значение, поскольку он может помочь историку в измерении степени влияния работы, т. е. ее „импакт-фактора“» [Garfield 1955, с. 3]

Довольно ясно, что здесь, как во всех других случаях, термин «импакт-фактор», или «степень влияния», предполагает, что цитирующая работа «построена» на цитируемой, т. е. цитирование является механизмом, посредством которого исследования продвигаются вперед.

Существует обширная литература о фактическом смысле цитирований, и из нее следует, что процесс цитирования является более сложным, чем можно было предположить, опираясь на эти расплывчатые заявления. Например, в работе 1983 года об оценке исследований Мартин и Ирвайн пишут:

«В основе всех этих проблем с использованием цитирования в качестве меры качества лежит наше непонимание причин того, почему авторы ссылаются на одни работы, а не на другие. (...) Простой анализ цитирований предполагает опору на весьма рациональную модель цитирования, согласно которой цитирование происходит по преимуществу с учетом научной признанности предыдущей работы, имеющей высокое качество или важность, и у данной работы существует равная вероятность быть процитированной разными авторами...» [Martin—Irvine 1983, с. 69].

В своей работе 1988 года о смысле цитирования [Cozzens 1989] Коззенс утверждает, что цитирование является результатом одной из двух линий поведения при оформлении научных публикаций — «признательной» и «риторической».

Цитирования первого типа несут в себе смысл, который мы чаще всего ожидаем в таких случаях, — признание того, что цитирующая работа имеет «интеллектуальный долг» по отношению к цитируемой. Цитиро-

вания второго типа имеют совершенно другой смысл: это просто ссылка на предшествующую работу, объясняющую какой-то результат, быть может, даже вовсе не принадлежащий цитируемому автору. Такие «риторические» цитирования являются просто средством

Смысл цитирования не так прост, и статистические показатели, основанные на цитировании, далеко не так «объективны», как утверждают их сторонники.

ведения научного обсуждения, а не признания интеллектуального долга. Конечно, в некоторых случаях цитирование может нести в себе оба этих смысла.

По наблюдениям Коззенс, большинство цитирований имеют «риторический» характер.

Это подтверждается опытом большинства практикующих математиков. (Например, в базе данных *Math. Reviews* из более чем 3 миллионов ссылок почти 30% приходится на книги, а не на исследовательские статьи в журналах.) Почему это важно? Потому что в отличие от «признательных» ссылок, которые, как правило, относятся к фундаментальным работам, выбор того, какие работы цитировать «риторически», зависит от многих факторов: известности цитируемого автора (эффект «гало»), отношения между цитирующим и цитируемым, доступности соответствующего журнала (являются ли журналы, имеющиеся в открытом доступе, более цитируемыми?), желания для удобства сослаться на несколько результатов, содержащихся в одной работе, и так далее. Немногие из этих факторов имеют непосредственное отношение к «качеству» цитируемой работы.

Даже когда цитирование носит «признательный» характер, оно может быть обусловлено различными мотивами, в том числе «степенью известности, отрицательной репутацией, рабочим характером информации, убедительностью, положительной репутацией, желанием предупредить читателя или общественным мнением» [Brooks 1996]. В большинстве случаев цитирование бывает обусловлено более чем одним из них. Некоторые заметные результаты могут пострадать из-за эффекта «стирания», если сразу же включаются в другие работы, которые в свою очередь служат основой для дальнейших ссылок.

Некоторые ссылки не просто не являются выражением «признательности» за выдающиеся исследования, а служат скорее предупреждением об ошибочных результатах. Можно привести много примеров таких «предупреждающих» цитирований.

Социология цитирования является сложным предметом, который выходит за рамки настоящего доклада. Однако даже наше беглое обсуждение показывает, что цитирование имеет сложный смысл и статистические показатели, основанные на цитировании, далеко не так «объективны», как утверждают их сторонники.

Некоторые могут возразить, что поиск смысла цитирований не играет особой роли, поскольку статистические показатели, осно-

ванные на цитировании, хорошо коррелируют с некоторыми другими мерами качества исследований (скажем, экспертной оценкой).

Например, упоминавшееся ранее сообщение в *Evidence* утверждает, что основанные на цитировании статистические показатели могут (и должны) заменить другие формы оценки из-за этой корреляции:

«Есть свидетельства, что библиометрические методы могут давать показатели качества исследований, приводящие примерно к тем же результатам, что и оценка другими исследователями» [Evidence Report 2007, с. 9].

Судя по всему, из этого делается заключение, что основанные на цитировании статистические показатели, независимо от их точного смысла, должны заменить другие методы оценки, потому что они с ними хорошо коррелируют. Даже если не принимать во внимание порочный круг в этом рассуждении, нелепость такого предложения всякому очевидна.

Разумное использование статистики

Ревностные попытки чрезмерно довериться объективным метрикам (статистическим показателям) для оценки исследований не являются ни новым, ни единичным явлением. Оно красноречиво описано в вышедшей в 2001 году популярной книге «Damned lies and statistics» («Проклятая ложь и статистика»), написанной социологом Дж. Бестом:

«Есть культуры, представители которых верят, что некоторые объекты имеют магическую силу; антропологи называют эти объекты фетишами. В нашем обществе своего рода фетишем является статистика. Мы склонны придавать статистическим показателям магический смысл, как будто они что-то большее, чем просто числа. Мы относимся к ним как к образцу истины, как если бы они перерабатывали сложность и запутанность реальности в простоту и ясность. Мы используем их для того, чтобы свести сложные социальные проблемы к более понятным оценкам, процентам и отношениям.

Статистика дает направление решению проблем, она показывает нам, о чем и насколько сильно мы должны беспокоиться. В некотором смысле социальная проблема становится статистической, и, поскольку мы относимся к статистике как к истине в последней инстанции, она становится своего рода фетишем, имеющим магический контроль над тем,

как мы рассматриваем социальные проблемы. Мы считаем, что статистика — это факты, которые мы открыли, а не просто выдуманные нами самими числа» [Best 2001, с. 160].

Эту мистическую веру в волшебную силу статистики цитирований можно найти в документации для программ, как национальных, так и ведомственных, оценки качества исследовательской работы. Можно найти ее и в работах, использующих *h*-индекс и его модификации.

Это отношение проявляется также в современных попытках улучшить импакт-фактор при анализе цитат путем использования более сложных математических алгоритмов, в том числе алгоритмов классификации страниц ([Bergstrom 2007], [Stringer—Sales-Pardo—Nunes 2008]). Сторонники этого подхода заявляют о его эффективности, что не оправдывается анализом и плохо поддается оценке. Так как они основаны на более сложных расчетах, предположения (зачастую скрытые), которые стоят за ними, большинству людей не так просто разглядеть.⁷⁾ Предполагается, что мы должны относиться к

Если в практической медицине принято консультироваться с врачами, то в области статистики, безусловно, следует советоваться со статистиками.

цифрам и рейтингам с благоговением — как к истине, а не как к тому, что создано нами.

Научные исследования — это не первый вид публично финансируемой деятельности, подвергающейся тщательному рассмотрению: за последние десятилетия люди пытались произвести количественные оценки производительности всего, начиная с систем образования (школ) и до здравоохранения (больниц и даже отдельных хирургов).

В некоторых случаях к делу подключались специалисты по статистике, чтобы проконсультировать тех, кто занимается оценками, по поводу разумных подсчетов и правильного использования статистических данных. Если в практической медицине принято консультироваться с врачами, то в области статистики, безусловно, следует советоваться со статистиками (и прислушиваться к их советам). Два прекрасных примера можно найти в работах [Bird 2005] и [Goldstein—Spiegelhalter 1996]. Хотя каждый из них имеет дело с оценкой производительности, а не с научными исследованиями (мониторинг производительности государственного сектора в первом и в здравоохранения/образования во втором), каждый

дает представление о разумном использовании статистики в оценке научных исследований.

Работа Гольдштейна и Шпигельхайтера, в частности, имеет дело с использованием League Tables (рейтингов*) на основе простейших числовых показателей (например, достижений студентов или медицинских результатов), и в особенности это относится к оценке научных исследований с помощью рейтинга журналов, статей или авторов с использованием статистики цитирований. В своей работе авторы описывают состоящую из трех частей основу для любой оценки эффективности.

Данные

Каков бы ни был мыслимый объем собранных статистических данных, он не сможет преодолеть ключевых сомнений относительно *пригодности* или *полноты* этих данных» [Goldstein—Spiegelhalter 1996, с. 389].

Это важное замечание для оценок производительности, основанных на цитировании.

Импакт-фактор, например, основан только на тех данных, которые получены по журналам из перечня *Thomson Scientific*. (Заметим, что импакт-фактор сам по себе является основным критерия отбора журналов в этот перечень.) Некоторые ставят под сомнение целостность таких данных [Rossner—VanEpps—Hill 2007]. Другие указывают на то, что иные наборы данных могут быть более полными [Meho—Yang 2007]. Некоторые группы продвинули идею использования Google Scholar для внедрения статистики, основанной на цитировании, например *h*-индекса, но данные, содержащиеся в Google Scholar, часто неточны (так как, например, как имена авторов, автоматически извлекаются из публикаций в Сети). Индекс цитирования отдельных ученых иногда трудно получить, поскольку в некоторых ситуациях и в некоторых странах бывает непросто однозначно определить авторство, и это это может стать огромным препятствием для сбора точных данных по цитатам. Зачастую не принимаются во внимание некоторые данные, которые относятся к анализу цитирования. А из статистики, основанной на ошибочных данных, делаются ошибочные выводы.

* Специфический вид рейтинга. См. [http://en.wikipedia.org/wiki/League table](http://en.wikipedia.org/wiki/League_table). —Прим. ред.

Статистический анализ и представление информации

«Мы обратим особое внимание на выбор подходящей статистической модели, на досадную *неопределенность* в представлении всех результатов, на методы коррекции результатов с учетом возмущающих факторов и, наконец, на то, до какой степени мы можем доверять имеющимся *рейтингам*» [Goldstein—Spiegelhalter 1996, с. 390].

Как мы уже писали ранее, в большинстве случаев, в которых статистики цитирования используются для оценки работ, людей и программ, никакой конкретной модели заранее не определяется. Сами данные подводят к идее модели, которая часто оказывается весьма расплывчатой. Образуется порочный круг, когда объекты оцениваются выше, потому что у них (в базе данных) более высокий рейтинг. Неопределенности *любого* из этих рейтингов уделяется недостаточное внимание, и мало анализируется, как эта неопределенность (например, ежегодные изменения импакт-фактора) будет влиять на рейтинги. Наконец, второстепенные факторы (например, конкретная дисциплина, тип статей, которые публикует журнал, является ли конкретный ученый экспериментатором или теоретиком), часто игнорируются в таких рейтингах, особенно если они осуществляются в оценке результатов в национальном масштабе.

Интерпретация и влияние

«Сравнения, которые обсуждаются в данной работе, вызывают большой общественный интерес, и, очевидно, это та область, где особое внимание к недостаткам является одновременно жизненно важным и наиболее вероятно игнорируется. Являются ли скорректированные результаты в какой-либо степени справедливой мерой „качества“ какого-либо учреждения — это один вопрос; с другой стороны, аналитики должны также знать о потенциальном влиянии результатов в отношении будущего изменения поведения учреждений и отдельных лиц, желающих улучшить свой последующий „рейтинг“» [Goldstein—Spiegelhalter 1996, с. 390].

Оценка научных исследований *также* вызывает большой общественный интерес. Результаты оценки отдельного ученого могут оказать серьезное и долгосрочное влияние на его карьеру; оценка отдела может изменить шансы на успех в далеком будущем, а что касается дисциплин, то набор оценок может разделить их на процветающие и «застойные». Поскольку дело столь важно, надо, конечно,

отдавать себе отчет в применимости и ограничениях используемых инструментов.

В какой степени цитирование измеряет качество исследований? Кажется, что количество цитирований коррелирует с качеством, и есть интуитивное понимание того, что высококачественные статьи много цитируются. Но, как говорилось выше, некоторые статьи, особенно в некоторых дисциплинах, много цитируются по причинам, отличным от высокого качества, так что из того, что статьи много цитируются, не следует, что они обязательно высокого качества. Необходимо лучше понять точную природу интерпретации рейтингов на основе статистик цитирований. Кроме того, если статистики цитирований играют центральную роль в оценке научных исследований, ясно, что авторы, редакторы и даже издатели найдут способы управлять системой в своих интересах [Macdonald—Kam 2007]. Долгосрочные последствия этого неясны и не изучены.

Статью Гольдштейна и Шпигельхальтера очень ценно прочитать сегодня, потому что из нее становится ясно, что чрезмерная зависимость от простой статистики в оценке научных исследований не является изолированной проблемой. Правительства, учреждения и частные лица боролись с подобными проблемами в прошлом и в других контекстах, и они нашли способы лучше понять статистические инструменты и дополнить их другими средствами оценки. Гольдштейн и Шпигельхальтер заканчивают свою работу в позитивном оптимистическом ключе:

«Наконец, хотя мы в целом настроены критично ко многим нынешним попыткам составить мнения об учреждениях, мы не хотим создать впечатление, что мы считаем, что все такие сравнения не имеют смысла. Нам кажется, что сравнение учреждений и попытка понять, почему они отличаются, являются чрезвычайно важным направлением деятельности и лучше всего проводить его в духе сотрудничества, а не конфронтации.

Пожалуй, это единственный надежный способ получения объективной информации, которая может привести к пониманию и, в конечном итоге, к совершенствованию процедуры. *Реальной проблемой с упрощенными процедурами, которые мы подвергли критике, является то, что они отвлекают от этой достойной цели как внимание, так и ресурсы»* [Goldstein—Spiegelhalter 1996, с. 406].

Было бы трудно найти лучшие слова, чтобы сформулировать цели, которые должны ставить перед собой все те, кто причастен к организации оценки исследований.

Цитированная литература*

- Adler R., *The impact of impact factors*, IMS Bulletin 36 (2007), no. 5, 4. [www](#)
- Amin M., Mabe M., *Impact factor: use and abuse*, Perspectives in Publishing 1 (2000), 1—6. [www](#)
- Batista P. D., Campiteli M. G., Kinouchi O., Martinez A. S., *Universal behavior of a research productivity index*, preprint, 2005, available at: [physics/0510142v1](#).
- Batista P. D., Campiteli M. G., Kinouchi O., *Is it possible to compare researchers with different scientific interests?*, Scientometrics 68 (2006), no. 1, 179—189. [www](#)
- Bergstrom C., *Eigenfactor: measuring the value and prestige of scholarly journals*, College & Research Libraries News 68 (2007), no. 5. (See also <http://www.eigenfactor.org/methods.pdf>.) [www](#)
- Best J., *Damned lies and statistics: untangling the numbers from the media, politicians, and activists*. University of California Press, Berkeley, 2001.
- Bird S. et al., *Performance indicators: good, bad, and ugly; Report of a working party on performance monitoring in the public services*, J. R. Statist. Soc. A 168 (2005), part 1, 1—27. [www](#)
- Brooks T., *Evidence of complex citer motivations*, Journal of the American Society for Information Science 37 (1986), no. 1, 34—36. [www](#)
- Carey A. L., Cowling M. G., Taylor P. G., *Assessing research in the mathematical sciences*, Gazette of the Australian Math Society 34 (2007) no. 2, 84—89. [www](#)
- Cozzens S. E. *What do citations count? The rhetoric-first model*, Scientometrics, 15 (1989), no. 5—6, 437—447. [www](#)
- Egghe L., *Theory and practice of the g-index*, Scientometrics 69 (2006), no. 1, 131—152. [www](#)
- Evidence Report, 2007, *The use of bibliometrics to measure research quality in the UK higher education system*. (A report produced for the Research Policy Committee of Universities, UK, by Evidence Ltd., a company specializing in research performance analysis and interpretation. Evidence Ltd. has «strategic alliance» with Thomson Scientific.) <http://bookshop.universitiesuk.ac.uk/downloads/bibliometrics.pdf>
- Ewing J., *Measuring journals*, Notices of the AMS 53 (2006), no. 9, 1049—1053. [www](#)
- Garfield E., *Citation indexes for science: A new dimension in documentation through association of ideas*, Science 122(3159) (1955), 108—111. [www](#)
- _____, *Citation analysis as a tool in journal evaluation*, Science 178(4060) (1972), 471—479. [www](#)
- _____, *Why are the impacts of the leading medical journals so similar and yet so different?*, Current Comments, No. 2, January 12, 1987. [www](#)
- _____, *Long-term vs. short-term journal impact (part II)*, The Scientist, 12(14) (1998), 12. [www](#)
- _____, *Agony and the ecstasy—the history and meaning of the journal impact factor*, presented at the International Congress on Peer Review and Bibliomedical Publication, Chicago, September 16, 2005. <http://garfield.library.upenn.edu/papers/jifchicago2005.pdf>
- Goldstein H., Spiegelhalter D. J., *League tables and their limitations: Statistical issues in comparisons of institutional performance*, J. R. Statist. Soc. A 159 (1996), no. 3, 385—443. [www](#), [www](#)
- Hall P., *Measuring research performance in the mathematical sciences in Australian universities*, The Australian Mathematical Society Gazette 34 (2007), no. 1, 26—30. [www](#)

* Интернет-адреса всех приведенных здесь источников содержатся в электронной версии данного сборника.

- Hirsch J. E., *An index to quantify an individual's scientific research output*, Proc Natl Acad Sci USA **102** (2006), no. 46, 16569—16573. [www](#)
- Kinney A. L., *National scientific facilities and their science impact on nonbiomedical research*, Proc Natl Acad Sci USA **104** (2007), no. 46, 17943—17947. [www](#)
- Lehmann S., Jackson A. D., Lautrup B. E., *Measures for measures*, Nature **444** (2006), no. 21, 1003—1004. [www](#)
- Macdonald S., Kam J., Aardvark et al.: *quality journals and gamesmanship in management studies*, Journal of Information Science **33** (2007), 702—717. [www](#)
- Martin B. R., *The use of multiple indicators in the assessment of basic research*, Scientometrics **36** (1996), no. 3, 343—362. [www](#)
- Martin B. R., Irvine J., *Assessing basic research*, Research Policy **12** (1983), 61—90. [www](#)
- Meho L., Yang K., *Impact of data sources on citation counts and rankings of LIS faculty: Web of Science vs. Scopus and Google Scholar*, Journal of the American Society for Information Science and Technology **58** (2007), no. 13, 2105—2125. [www](#)
- Molinari J. F., Molinari A., *A new methodology for ranking scientific institutions*, Scientometrics **75** (2008), no. 1, 163—174. [www](#)
- Monastersky R., *The number that's devouring science*, Chronicle Higher Ed. **52** (2005), no. 8. [www](#)
- Rossner M., Van Epps H., Hill E., *Show me the data*, Journal of Cell Biology **179** (2007), no. 6, 1091—1092. [www](#)
- Seglen, P. O., *Why the impact factor for journals should not be used for evaluating research*, BMJ **314** (1997), 497. [www](#)
- Sidiropoulos A., Katsaros D., Manolopoulos Y., *Generalized h-index for disclosing latent facts in citation networks*. VI, preprint, 2006, available at [cs/0607066v1](#).
- Stringer M. J., Sales-Pardo M., Nunes Amaral L. A., *Effectiveness of journal ranking schemes as a tool for locating information*, PLoS ONE **3**(2) (2008), e1683. [www](#)
- THOMSON: JOURNAL CITATION REPORTS. <http://scientific.thomson.com/products/jcr/>
- THOMSON: SELECTION. <http://scientific.thomson.com/free/essays/selectionofmaterial/journalselection/>
- THOMSON: IMPACT FACTOR. <http://scientific.thomson.com/free/essays/journalcitationreports/impactfactor/>
- THOMSON: HISTORY. <http://scientific.thomson.com/free/essays/citationindexing/history/>
- THOMSON: FIFTY YEARS. <http://scientific.thomson.com/free/essays/citationindexing/50y-citationindexing/>

Примечания

¹⁾ В *Reader's Digest* от октября 1977 года эта цитата была приписана Эйнштейну. Похоже, что она была извлечена из следующей настоящей его цитаты: «Едва ли можно отрицать, что высшей целью всех теорий является то, чтобы сделать основные элементы как можно более простыми и немногочисленными, не поступаясь при этом адекватным представлением ни единого опытного данного». Из Спенсеровской лекции «О методе теоретической физики», прочитанной Эйнштейном в Оксфорде (10 июня 1933 г.); опубликовано также в *Philosophy of Science* **1** (1934), no. 2, 163—169.

²⁾ Хотя в этом разделе мы сосредоточились на импакт-факторе *Thomson Scientific*, отметим, что *Thomson* продвигает также использование и двух других статистик. Добавим, что аналогичные статистики, основанные на подсчете среднего числа цитирований в журналах, могут быть получены из других баз данных, включая Scopus, Spires, Google Scholar и (по математике) базу данных *Math. Reviews*. Последняя включает цитаты из более чем 400 математических журналов за период с 2000 г. по настоящее время. Эти те журналы, которые реферируются в *Math. Reviews* с 1940 года. База содержит более 3 миллионов ссылок.

³⁾ *Thomson Scientific* указывает (март 2008 г.), что рассматривает журналы в следующих категориях:

- математика (217);
- прикладная математика (177);
- междисциплинарная математика (76);
- математическая физика (444);
- вероятность и статистика (96).

Категории перекрываются, а общее число журналов составляет около 400.

В отличие от этого, *Math. Reviews* каждый год анализирует более 1200 журналов, из которых более 800 журналов составляют «ядро» (в том смысле, что каждая статья в журнале включена в *Math. Reviews*). *Zentralblatt* охватывает примерно столько же математических журналов.

⁴⁾ База данных цитирований в *Math. Reviews* включает (по состоянию на март 2008 г.) более чем 3 миллиона ссылок приблизительно из 400 журналов, изданных с 2000 г. по настоящее время. Ссылки согласованы с разделами в базе данных *Math. Reviews* и охватывают период в несколько десятков лет. В отличие от *Science Citation Index*, сюда включено цитирование и из книг, и из журналов. Любопытно, что примерно 50 % цитирований относится к публикациям предыдущего десятилетия, 25 % ссылок приходится на статьи, появившиеся в десятилетие до него, 12,5 % — в предшествующее десятилетие и так далее. Для каждой дисциплины характерна, конечно, своя картина.

⁵⁾ Высокая неравномерность распределения цитируемости по статьям в сочетании с коротким интервалом времени (рассматриваются ссылки на статьи, опубликованные за последние пять лет, в статьях, опубликованных за последний год) означает, что много статей имеют либо мало ссылок, либо не имеют их вообще. Инту-

итивно понятно, что это приводит к слишком частому равенству количества цитирований у двух произвольно выбранных статей.

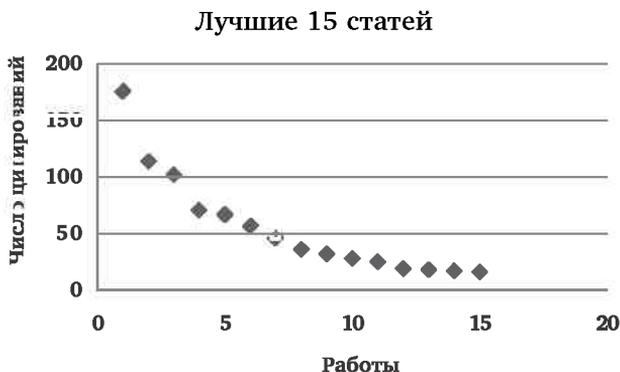
Тот факт, что многие статьи не имеют цитирований (или их число очень мало), также является следствием большого временного интервала, характерного для цитирования в математических статьях: очень часто перед появлением первого цитирования проходит много лет. Если рассматривать более длительный период времени как для источников цитирований, так и для самих статей, то количество цитирований заметно возрастет и станет легче различать журналы по этому показателю. Именно такой подход используется в [Stringer et al. 2008]. Там показано, что для достаточно длительного периода времени распределение цитирований по статьям оказывается лог-нормальным. Таким образом, появляется механизм, позволяющий сравнивать журналы в терминах соответствующих распределений, что, безусловно, более продумано, чем использование импакт-фактора. И тем не менее, при этом опять-таки учитывается только цитируемость и ничего более.

⁶⁾ Чтобы показать, как много информации теряется при использовании только h -индекса, приведем реальный пример одного уважаемого математика, находящегося в середине своей карьеры, который опубликовал 84 научных работы. Распределение цитирований выглядит следующим образом:



Обратите внимание, что чуть менее 20% публикаций имеют 15 или больше цитирований. Распределение действительных под-

счетов цитирований для этих 15 работ таково:



В анализе Хирша, однако, вся эта информация выброшена. Он лишь констатирует, что h -индекс равен 15, т. е. что 15 лучших работ имеют 15 или больше цитирований.

⁷⁾ Алгоритм из работы [Bergstrom 2007] использует алгоритм страничного ранжирования, который придает каждой цитате вес, а затем вычисляет «импакт-фактор», используя средневзвешенные для цитирований. Сильной стороной алгоритмов страничного ранжирования является то, что они принимают во внимание «ценность» цитат.

С другой стороны, их сложность может быть опасной, потому что окончательные результаты труднее понять. В этом случае все «самоцитирования» отбрасываются, т. е. отбрасываются все цитаты из статей в данном журнале J на статьи, опубликованные в том же J в течение предыдущих пяти лет. По сути это нельзя назвать «самоцитированием», но рассмотрение некоторых данных базы цитат *Math. Reviews* позволяет предположить, что это приводит к выбрасыванию примерно одной трети всех цитирований.

Алгоритм из работы [Springer et al. 2008] интересен отчасти потому, что он пытается рассмотреть различные временные рамки цитирований, а также вопрос о сравнении случайно выбранных статей в одном журнале со статьями в другом. В этом случае сложность алгоритмов также затрудняет для большинства людей получение оценки имеющихся результатов. На с. 2 этой работы принимается примечательная гипотеза: «Наше первое предположение, что статьи, опубликованные в журнале J , имеют нормальное распределение „качества“...» Это кажется противоречащим обычному опыту.

Потерянное при публикации: как измерение вредит науке

Питер А. Лоуренс*

Измерение научной продукции (производительности) является делом трудным. Используемые измерения (импакт-фактор журнала, степень цитирования оцениваемой работы) грубы. Но эти измерения сейчас повсеместно приняты, что определяют большинство значимых вещей: иметь постоянную должность или быть безработным, получить постдокторский грант или нет, быть успешным или провальным. В результате ученых стали вынуждать отойти от общепринятых целей научного исследования, заменив стремление совершать открытия на желание публиковать как можно больше статей, пытаясь при этом помещать их непременно в журналах с высоким импакт-фактором. Как следствие, научная деятельность деформировалась, а полезность, качество и объективность статей ухудшились. Срочно требуется изменить способы оценки ученых, и в настоящей работе я предлагаю некоторые из них.

«Are all thy conquests, glories, triumphs, spoils,
Shrunk to this little measure»**

William Shakespeare: *Julius Caesar*, Act III, scene I

«What category? Minor writer? ...Significant minority writer? Major minor writer? ...never minor major writer? What did he say to you?»***

Brian Friel, 1997: *Give Me Your Answer, Do!* (p. 48)

«Время от времени мне наносят визиты серьезные мужчины и женщины, вооруженные опросниками и магнитофонами, которые хотят узнать, что же сделало лабораторию молекулярной биологии в Кембридже столь замечательно творче-

Оригинал: Peter A. Lawrence, *Lost in publication: how measurement harms science*, *Ethics in Science and Environmental Politics* 8 (2008), 9—11. © Inter-Research, 2008. В своем письме автор указывает на следующий смысл первой половины заглавия его статьи: «Оно неявно отсылает к фразе «Lost in Translation», которая говорит о том, что все многообразие смысловых значений может навсегда исчезнуть в процессе перевода. Заглавие указывает на то, что, концентрируясь на самом факте публикации, а не на содержании статьи, мы теряем нечто значимое». Перевод с английского под редакцией А. П. Васильевича.

* Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK, and MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK. Email: pal@mrc-lmb.cam.ac.uk

** Ужели слава всех побед, триумфов // Здесь уместилась? (пер. М. Зенкевича).

*** Какая категория? Второстепенный автор?...Заметный автор из числа второстепенных? Первый из второстепенных авторов? ...первостепенный автор, которому никогда не грозит стать второстепенным? Что он Вам сказал?

ской. Они приходят из общественных наук, чтобы... найти свой Священный Грааль в междисциплинарной организации. Я испытываю искушение обратить их внимание на Флоренцию XV века, которая при населении менее 50 000 человек дала Леонардо, Микеланджело, Джиберти... и других великих художников. Выясняли ли мои собеседники, не могли ли правители Флоренции создать междисциплинарную организацию скульпторов, архитекторов и поэтов, чтобы воплотить в жизнь этот расцвет великого искусства?... Мои вопросы не столь уж абсурдны как могут показаться, поскольку творчество в науке, как и в искусстве, не может быть организовано. Оно возникает спонтанно из индивидуального таланта. Хорошо управляемые лаборатории могут способствовать этому, но иерархическая организация, негибкие бюрократические правила и груды бесполезных документов могут убить это. Открытия не могут быть запланированы; они появляются, как эльфы, в неожиданных местах».

Макс Перуц: Предисловие к *I Wish I'd Made You Angry Earlier: Essays on Science, Scientists, and Humanity*

Независимо от того, какие измерения разработаны, чтобы оценивать людей, независимо от того, насколько они неадекватны, в человеческой природе заложено стремление приспособиться к ним и получить как можно большую оценку. Поэтому не должно удивлять, что попытки количественно определить качество научных публикаций привели ко многим поведенческим и политическим изменениям в мире науки.

Для исследователей в науке всегда было крайне важным опубликовать свою работу. Всегда имелись три цели: во-первых, распространить новую информацию, чтобы другие могли усвоить ее; во-вторых, чтобы другие ученые могли повторить исследования или основываться на них, используя дополнительные наблюдения или эксперименты; и лишь в-третьих, чтобы поддержка ученого, финансовая или иная, могла быть оправдана в глазах заинтересованных сторон. Эта третья причина была вспомогательной, но более это не так; опубликование стало главной целью, потому что это — путь выживания ученого [Lawrence 2003]. Эта кардинальная смена приоритетов нанесла вред научной практике, трансформировала мотивацию исследователей, изменила способ представления результатов и уменьшила точность и доступность научной литературы.

Позвольте мне объяснять: если вы публикуетесь, чтобы получить талон на питание, вам следует это делать, когда вы голодны, а не тогда, когда работа завершена, окончательно прояснена и достой-

на публикации. Это объясняет, например, почему многие диссертации на степень PhD издаются в неудовлетворительном виде до конца обучения по программе PhD: если студент должен остаться в науке после получения степени, он будет нуждаться в гранте, а большинство заявок на грант требует публикаций. Эти публикации требуют времени на подготовку и на принятие к опубликованию в журнале, и это время должно быть добавлено к задержке между начальной подачей заявки на грант и моментом поступления денег — а это означает, что многие студенты должны начать готовить бумаги так рано, когда они еще едва знают, чем они занимаются! Постдоки должны публиковать, чтобы получить работу, и более старшие ученые должны постоянно добавлять к своему списку публикаций новые наименования, чтобы возобновлять гранты или получить постоянную работу. Снова и снова работы представляются в журналы преждевременно, а проекты либо срочно слепляются друг с другом, либо специально конструируются, чтобы «сделать статьи» до истечения сроков. Умелые карьерные ученые учатся правильно упаковывать свои работы, чтобы они могли быть поданы сначала в журналы, а затем в утробы грантовых агентств.

Далее, имеется вопрос, в какой журнал послать работу. Так как ученые теперь оцениваются не столько по их силе, интересам или качеству самой работы, сколько по импакт-факторам журналов [Steel et al. 2006], многие, если не большинство ученых, тратят слишком много времени и усилий на обдумывание и хлопоты по поводу стратегии опубликования. Что следует делать, если статья отклонена первоначально выбранным журналом? Стоит ли произвести еще один заход в надежде на изменение решения редакции или было бы более благоразумным немедленно попробовать другой журнал? Если другой журнал, то следует ли снова посылать статью неизменной или же попытаться облачить результаты в иной маскарадный костюм? Возможно, в статью стоит добавить другие находки, чтобы она выглядела более современной и пробуждала к себе интерес. Возможно, статья нуждается в связи с чем-то недавно изданным и модным, «крюке», как это называют журналисты, или же в увязывании с идущими сейчас спорами: ученые выучили, что редакторы должны поддержать импакт-фактор своих журналов, а споры могут привести к цитированию. Есть столько уловок, которые можно попробовать, но все они отнимают время, и если они терпят неудачу, как это часто происходит, то месяцы оказываются потраченными впустую и приоритет может быть потерян.

Эти уловки нацелены прежде всего на редакторов, поскольку именно они будут решать, должна ли работа быть отрецензирована. Однако, когда этот барьер преодолен, следующем препятствием является процесс рецензирования, а рецензенты представляют различные опасности для квалифицированного автора статей. Здесь нужно действовать очень тщательно. Нужно позаботиться о том, чтобы процитировать всех «лидеров» области. Также может быть безопаснее преуменьшить любые результаты, не соответствующие восприятию вероятных рецензентов, и неблагоразумно открыто противопоставлять расходящиеся результаты. В настоящее время нормальная стратегия должна состоять в игнорировании или скрытии конфликтов. Их открытое рассмотрение может быть опасным, так как, возможно, какой-то из рецензентов может быть раздражен, ибо он участвовал в получении этих результатов (или он даже может быть ответственен за них). Это особенно относится к журналам с высоким импакт-фактором, поскольку редакторы этих журналов постоянно настаивают, чтобы статью похвалил «с энтузиазмом» каждый из рецензентов, вследствие чего одна двусмысленная рецензия может погубить статью. Однако у всех этих уловок есть своя обратная сторона, поскольку они обычно уменьшают объективность и ясность аргументов. Политика обессиливает науку.

По моему мнению, растущее использование h -индекса [Hirsch 2005] — который, будучи основанным на числах цитирований, пытается дать количественное выражение сразу и производительности, и влиятельности ученого — приведет к зависимости от цитирования и к одержимости им. Я предсказываю что «вылавливание цитирований» (citation-fishing) и «цитатный обмен» (citation-bartering) станут обычной практикой [Lawrence 2007]. Отмечу, что большинство журналов теперь устанавливает предел для числа ссылок — либо напрямую, либо ограничивая число слов в статьях, включая ссылки. Авторы таким образом поощряются делать неадекватные или несоответствующие цитирования. Например, ссылка на единственный обзор вместо нескольких оригинальных статей, сообщающих об открытии, требует меньшего количества слов, и, поскольку обычно имеется нескольких различных соответствующих обзоров, выбор того, какой именно процитировать, легко может иметь скорее политический, чем научный характер. Сколь неадекватны ни бывают цитирования, ученые будут оценивать по ним собственную работу, и это приводит к сомнительной практике. Например, ка-

кие-то ученые будут утверждать свое превосходство над другими, если они имеют больше цитирований, и это будет одобрено счетоводами повсюду. Как следствие, тот, кто работает в небольших областях (где имеется склонность к меньшему цитированию), окажется в невыгодном положении.

Дэвид Колкухун изучал ситуацию в Империял Колледже, Лондон. Он проанализировал последствия охвата научной жизни такими измерениями (www.dcs-science.net). В этом несомненно дремучем колледже ожидается, что каждый ученый выдаст ежегодно по крайней мере 3 статьи, из которых по крайней мере одна должна быть в журнале с импакт-фактором выше 5. Колкухун убедительно доказывает, что эти требования не имеют никакого смысла, указывая, например, что если бы современные измерения были применены в прошлом, то некоторые из самых лучших ученых, включая Нобелевских лауреатов, время от времени терпели бы карьерную неудачу.

Имеются и другие последствия использования численных измерений: поскольку такое использование поощряет агрессивное, стяжательское и эксплуататорское поведение [Lawrence 2002, Lawrence 2003, Montgomerie—Birkhead 2005], оно отберет среди ученых тех, кто обладает этими качествами. Я обосновал в другом месте, почему творческому открытию не помогают измерения, которые отбирают жестких борцов и направлены против более склонных к размышлениям скромных людей [Lawrence 2006]. Например, Эд Льюис, получивший нобелевскую премию 1995 г. по физиологии или медицине и бывший в течение своей жизни преданным и скромным ученым, а также редким и нерегулярным автором статей, имел очень низкий *h*-индекс и не смог бы выиграть гранты в сегодняшнем мире. Кроме того, в среднем, женщины менее агрессивны, чем мужчины [Baron-Cohen 2003, Babcock—Laschever 2004, Lawrence 2006, Symonds et al. 2006] и, таким образом, хотя эти измерения дискриминируют более деликатных ученых обоих полов, женщин это касается в большей мере. Статистические данные не оставляют сомнений, что женщины наиболее часто выпадают из биологической науки [ETAN 2000], а многие из них очень талантливы. Наиболее часто озвучиваемые причины таковы: (1) «Я не хочу тратить оставшуюся часть моей жизни на конкуренцию и хвастовство», (2) «Я не хочу получать приз в борьбе за получение гранта», и (3) «Современная наука превратилась в борьбу за выживание, а я предпочитаю делать что-нибудь другое».

Что следует сделать, чтобы улучшить положение вещей? По моему мнению, грантовые агентства должны использовать другие способы оценки качества работы и объявить об этом. Тогда те, кого они поддерживают, могли бы тратить меньше времени на интриги по опубликованию своих статей в журналах тщеславия и больше времени на исследование и решение научных проблем [Keller 2007]. По моему ограниченному опыту, заявки на грант не описывают того, что вы фактически делаете, а являются в действительности тестом на изобретательность и осведомленность, в котором честность имеет малую цену; они сводятся к попытке продемонстрировать, что вы знаете что делаете, можете предугадать результаты экспериментов и оценить, что опасно раскрывать. Вам требуется пронизательность в понимании того, что могло бы впечатлить оценивающие комиссии. Я не уверен, сколь хорошо этот нелепый процесс служит времени или науке. Чтобы уменьшить давление на ученых, которые проводят так много времени, придумывая заявки на грант, и сократить количество нудных текстов, подлежащих оценке и сравнению, я предлагаю использовать альтернативное средство заявки на грант, которое бы целиком основывалось на качестве работы последних 5 лет, как это имеет место в Канаде (CNSRC). Все, что претенденты, которые использовали бы этот метод, должны были бы сделать, — это перечислить, скажем, 3 статьи из тех, что они считают своими лучшими работами за последние 5 лет, и кратко описать тех, кто было ответственен за эти работы. Преимущество оценки лишь небольшого количества работ имеет, помимо очевидного аспекта — уменьшения объема чтения, еще и тот, что авторы поощрялись бы писать малое число работ, означающих или сообщающих что-то важное, вместо большого числа работ, написанных только ради того, чтобы быть подсчитанными. Воодушевляет, что в США и, конкретно, в Медицинском институте Говарда Хьюза оценка ограничивается 5 статьями за последние 5 лет, выбранными претендентами. Возможно, это первый шаг в том, что, на мой взгляд, было бы полезной тенденцией. Но как следует оценивать эти статьи? Только просто читая их: никакого лучшего пути нет. Этот метод был использован в прошлом; он не «объективен», но служит попыткой получить то, что имеет значение; это, я думаю, лучше, чем полагаться на точное измерение того, что не имеет значения. Исследование должно быть оценено в отношении строгости, новизны и значимости, яркости, экономической и эвристической ценности; эти качества может быть трудно оценить, но мы должны пытаться.

Цитированная литература

- Babcock S., Laschever S., *Women don't ask: negotiation and the gender divide*, Princeton University Press, Hoboken, NJ, 2004.
- Baron-Cohen S., *The essential difference. Men, women and the extreme male brain*, Allen Lane, London, 2003.
- ETAN (European Technology Assessment Network) on women and science, *Science policies in the European Union: promoting excellence through mainstreaming gender equality*, European Commission, Brussels, 2000
- Friel B., *Give me your answer, do!*, The Gallery Press, Loughcrew, 1997.
- Hirsch J. E., *An index to quantify an individual's scientific research output*, Proc Natl Acad Sci USA **102** (2005), 16569—16572.
- Keller L., *Uncovering the biodiversity of genetic and reproductive systems: time for a more open approach*, Am Nat **169** (2007), 1—8.
- Lawrence P. A., *Rank injustice*, Nature **415** (2002), 835—836.
- Lawrence P. A., *The politics of publication*, Nature **422** (2003), 259—261.
- Lawrence P. A., *Men, women, and ghosts in science*, PLoS Biol **4** (2006), e19. [www](#)
- Lawrence P. A., *The mismeasurement of science*, Curr. Biol **17** (2007), R583—R585.
- Montgomerie B., Birkhead T., *A beginner's guide to scientific misconduct*, Int Soc Behav Ecol **17** (2005), 16—24.
- Perutz M., *I wish I'd made you angry earlier: essays on science, scientists, and humanity*, Cold Spring Harbor Laboratory Press, Woodbury, NY, 2003.
- Steele C., Butler L., Kingsley D., *The publishing imperative: the pervasive influence of publication metrics*, Learned Publishing **19** (2006), 277—290.
- Symonds M. R. E., Gemmell N. J., Braisher T. L., Gorringer K. L., Elgar M. A. *Gender differences in publication output: towards an unbiased metric of research performance*, PLoS One **1** (2006), e127. [www](#)

Бегство от импакт-фактора

Филипп Кемпбелл*

Меня, как главного редактора журнала Nature, волнует имеющаяся среди академической администрации тенденция сосредотачиваться на импакт-факторе журнала при оценке значимости научного вклада исследователя, влияющая на его продвижение, прием на работу и, в некоторых странах, на финансовые вознаграждения за каждую статью. Наше собственное внутреннее расследование показывает, что высокий импакт-фактор журнала может быть искажен вследствие многократного цитирования небольшого числа статей. Таким образом, он не отражает средний уровень большинства статей, что уменьшает его ценность как объективной меры отдельной статьи. Предлагаемые альтернативные индексы имеют свои собственные недостатки. Многие исследователи отмечают, что их важные работы были опубликованы в журналах с низким импакт-фактором. Сосредоточение на цитированиях отдельной статьи является более надежным индикатором индивидуального влияния. Положительным моментом является увеличивающаяся способность отследить индивидуальные вклады посредством оценки вкладов соавторов и, возможно, в будущем, цитируемости частей статьи, а не всей статьи в целом. Имеются попытки избежать обращения к иерархии журналов с высоким импакт-фактором посредством недифференцированных баз данных рецензируемых статей, таких как PLoS One. Все еще предстоит выяснить, поможет ли эта модель какой-то выдающейся работе достичь должного признания независимо от редакционного отбора. Хотя существующая система и может быть эффективной в оценке заслуг на национальном и институтском уровне, самый эффективный и справедливый анализ индивидуального вклада дает непосредственная оценка отдельных статей, независимо от того, где они были опубликованы.

Введение

В качестве редактора, вовлеченного в журнальную конкуренцию, я обращаю свое внимание (среди многого другого) на то, как возрастает значение того, что делают мои коллеги для авторов по улучшению качества их статей и для читателей в деле оценки важности этих статей. Цитирования являются одним из показателей нашего успеха, и наши статьи действительно много цитируются, но на «импакт-фактор» я обращаю как можно меньше внимания.

В течение многих лет я выражал беспокойство по поводу зависимости столь многого от импакт-фактора как меры оценивания.

Оригинал: Philip Campbell, *Escape from the impact factor*, *Ethics in Science and Environmental Politics* 8 (2008), 5–7. © Inter-Research, 2008. Перевод с английского К. Иванова.

*Nature, 4 Crinan Street, London N1 9XW, UK.

Рискуя повторить высказанное ранее, я все же еще раз изложу суть моих опасений, прежде чем более конструктивно обсудить другие пути продвижения вперед.

Проблемы

Основное мое беспокойство касается грубости оценки. В 2005 г. мы ее измерили. Импакт-фактор журнала в 2004 году был равен 32,2. Как сказано в редакционной статье (*Nature* 435 (2005), 1003—1004), мы проанализировали цитирования отдельных статей, опубликованных в *Nature* за соответствующий период (т. е. цитирования в 2004 г. работ, опубликованных в период с 2002 по 2003 г.), и обнаружили, что 89 % нашего импакт-фактора происходит лишь от 25 % наших статей.

Среди статей в *Nature*, опубликованных в 2002—2003 г., наиболее цитируемой в 2004 г. была статья о геноме мыши. Эта статья являлась кульминацией большого совокупного проекта и как таковая представляла собой удобную ссылку, а не изложение, глубоко проникающее в суть дела. Когда мы проанализировали статистику в 2005 г., на эту статью пришлось более 1000 цитирований. При оценке импакт-фактора лишь одного 2004 г. на нее пришлось 522 цитирования. Наша следующая по числу цитирований статья, опубликованная в период 2002—2003 г., касалась функциональной организации протеомы (proteome*) дрожжей. В 2004 г. на нее пришлось 351 цитирование. Лишь на 50 из примерно 1800 цитируемых и опубликованных в 2002—2003 г. статей приходится более 100 цитирований в 2004 г. На большинство же наших статей приходится менее 20 цитирований.

Некоторые администраторы рассматривают импакт-фактор как меру уровня цитирования, типичного для данного журнала, но, как показывают приведенные цифры, для многих журналов он «типичным» вовсе не является.

Почему для большинства статей уровень цитируемости был «низким» (по нашим стандартам)? Потому что они относились к дисциплинам либо с характерно низким показателем уровня цитируемости на статью — таким, например, как физика, либо с показателями уровня цитируемости, который, как правило, растет медленно, как, например, в науках о земле, либо же потому, что, хотя они и были

* Proteome — полный набор белков (в данном случае) организма. — Прим. ред.

превосходными (например, являлись провидческими), но не являлись «горячими».

У импакт-фактора есть другой аспект, который я не приемлю всей душой. В 2003 и 2004 г. мы опубликовали значительно больше статей, чем наш основной конкурент, журнал *Science*. Но в формуле импакт-фактора число ISI*-учитываемых цитирований статей в *Nature* за 2 года делится на число тех статей в *Nature* за тот же период, на которые «можно сослаться». Чем больше число статей, тем ниже импакт-фактор. Другими словами, забота о максимизации импакт-фактора превращает в бремя то, что многие посчитали бы за достоинство, — большое число хороших статей для чтения.

Добавляя соли на раны, скажем, что числа, на которых основывается величина импакт-фактора, очень сомнительны. Как мы ни пытались, мои коллеги и я не можем согласовать наши собственные подсчеты цитируемых публикаций в самом *Nature*, в некоторых других журналах *Nature*** и даже в *Science* с подсчетами, используемыми в ISI.

Некоторые коллеги придают большое значение такому конкурентному ранжированию, хотя и вполне понимают, насколько произвольным это число является. В результате оно действительно оказывает влияние на то, как научные работники воспринимают журналы. Я не могу отрицать, что наш собственный маркетинговый отдел в восторге от высокого рейтинга журналов *Nature*, но, хотите верьте, хотите нет, эта оценка интересует меня только в том отношении, что кто-то может быть затронут этими числами в такой степени, что это может неблагоприятно затронуть и нас. Редакторы *Nature* имеют полное право гордиться публикуемыми статьями и тем спектром уровня цитирований, которого они добились. Я склонен думать, что мы стремимся публиковать лучшие статьи, но слово «лучшие» ведет к вопросам и субъективным оценкам. До какой степени мы преуспеваем в этом — судить читателям в их собственных терминах, но лучше всего, чтобы оценка «лучшие» была независимой от импакт-фактора.

Некоторые могут считать это лицемерным позерством. В конце концов, скажут они, *Nature* находится в одном ряду с другими журналами и самим импакт-фактором, краеугольными камнями тирании академической подотчетности, где все чувствуют себя вы-

* Institute for Scientific Information — предшественник *Thomson Scientific* (см. статью «Статистики цитирования» в наст. изд., с. 12). — Прим. ред.

** Издаваемых *Nature Publishing Group*. — Прим. ред.

нужденными толкаться для публикации в журналы с высоким импакт-фактором. В ответ я говорю, что журналы делают свою работу по отбору сильных статей, а то, как другие эксплуатируют эту работу в своих целях, является их собственным делом.

Этот анализ можно расширить. Например, биолог Питер Лоуренс указал на поведенческие аспекты этой «тирании», включая рост неуместного почетного авторства, увеличение количества авторов, скрывающее их истинные вклады, и поездки на встречи — ненужные во всех других отношениях, — чтобы потрепаться (среди прочих) с редакторами журналов [Lawrence 2003, Lawrence 2007]. На мой взгляд, любой, кто верит тому, что эта последняя деятельность значительно увеличит вероятность опубликований их работ, вводит себя в заблуждение.

Любая система измерения порождает трюкачество с целью увеличить успех. Лоуренс тоскует по одной системе, которая устраняет это: по обсуждениям лицом к лицу между рецензентами, позволяющим дать действительно глубокую и хорошо сфокусированную оценку научной заслуги. Я тоскую по ней тоже. Позор, что этот метод так часто считают непрактичным.

Возможные альтернативы

Как можно заметить по страницам *Scientometrics* и других журналов, признавая необходимость количественных показателей в качестве входных данных при оценивании степени влияния, многие люди занимаются разработкой лучших измеряющих показателей. Статья в *Wikipedia* об *h*-индексе Хирша выдвигает на первый план его недостатки, хотя многие теперь применяют его. Другая работа, сравнивающая такие индексы, высветила достоинства подсчета среднего числа цитирований за статью при оценке степени индивидуального влияния ученого [Lehman et al., 2006].

Я заинтригован возможностью большей степени детализации в том, что касается публикаций. Это уже происходит в отношении вкладов соавторов в публикацию. В самом *Nature* и научно-исследовательских журналах *Nature* (после инициативы, взятой на себя — к их большой чести — медицинскими журналами) мы поощряем авторов включать краткие резюме, описывающие вклад каждого автора в публикуемую работу. Все больше авторов принимает это предложение. Это будет стимулироваться и дальше и, без сомнения,

станет более принятым, если финансирующие агентства и учреждения начнут явно отслеживать такую информацию.

Другое возможное изменение, ждущее своего появления, состоит в возможности цитирования подразделов статей. Для этого требуется, чтобы эти подразделы (или «частицы статей» как называет их один мой коллега) были индивидуально идентифицированы, например с помощью цифрового идентификатора объектов (Digital Object Identifier, DOI). Недавно два исследователя, один — эколог, другой — микробиолог, рассказали мне, что их наиболее высоко цитируемые статьи были опубликованы в малоизвестных журналах, а основанием их цитируемости был не столько основной вывод статьи, сколько технический аспект, касающийся используемых методов. Было бы, несомненно, здравым шагом предоставить людям возможность цитирования этих подразделов статей, особенно в предвидении литературы, ориентированной на компьютеры, когда последние извлекают такие ключевые разделы с помощью нетривиальных процедур поиска и агрегирования. То же самое должно иметь место в отношении публикаций в серой* литературе, таких как индивидуальные фрагменты в базах данных, которые оказываются имеющими большую ценность.

Такие усовершенствования системы не покончат с одним ключевым аспектом тирании подотчетности — сверхзависимостью, в поверхностных оценках, от журнальных брендов и, в особенности, от чрезмерного акцента на те из них, которые имеют очень высокие импакт-факторы. Один из путей вперед — снова за счет увеличения степени детализации доступной информации — это сделать статистику цитирований каждой статьи более доступной.

В настоящее время идет эксперимент, пытающийся полностью покончить с престижными брендами в литературе; это рецензируемый архив, в котором не предпринимается попыток отбирать работы по уровню совершенства. К такому типу относится *PLoS One*, сетевой журнал, издающийся Публичной научной библиотекой (the Public Library Science, www.plosone.org).

Аспектом, наиболее волнующим его основателей, является возможность комментирования опубликованных работ. Как ожидается, система основана на модели, в которой автор оплачивает публикацию. Обе эти особенности являются экспериментами сами по се-

* Род научной литературы, распространяемый вне обычной издательской практики (отчеты, препринты и т. п.). — *Прим. ред.*

бе и еще должны утвердить себя, соответственно, в отношении популярности или финансовой жизнеспособности. Здесь к делу больше всего относится то, что модель PLoS One сознательно избегает той миссии, с которой *Nature* еженедельно пытается обратиться к вниманию своих читателей, — миссии отбора наиболее интересного из того, что представлено к опубликованию. Вместо этого используется единственный критерий — то, что статья была оценена рецензентами как заслуживающая доверия в научном отношении.

Однако, поскольку никто не хочет пробираться через болото статей чрезвычайно смешанного качества, как же обнаруживать наиболее интересные статьи в таком архиве? Возможно, некоторую модель дает хорошо налаженный сервер препринтов arXiv (<http://arxiv.org>)*. Он содержит тысячи препринтов предварительных версий статей и не выносит суждений об их значении, но единственный способ, с помощью которого люди могут выделить статьи с особо интересующим их содержанием, состоит в том, чтобы ежедневно посещать этот сайт и видеть, что было добавлено к соответствующей предметной области.

Читатели не удивятся, что я отстаиваю, при всех ее недостатках, модель, в которой редакторы отбирают интересные статьи. Я указал здесь на некоторые пути, по которым использование соответствующей статистики могло бы быть улучшено.

Я уверен, что статистики цитирований большого количества отдельных статей могут отражать влияние публикаций на институтском или национальном уровне. Они могут дать полезную информацию и на индивидуальном уровне. Но для верной оценки человека нет замены тому, чтобы прочитать сами статьи, независимо от журнала, в котором они появились.

Цитированная литература

- Lawrence P., *The politics of publication*, *Nature* **422** (2003), 259—261.
Lawrence P., *The mismeasurement of science*, *Curr Biol* **17** (2007), R583—R585.
Lehmann S., Jackson A.D., Lautrup B.E., *Measures for measures*, *Nature* **444** (2006), 1003—1004.

* Содержит огромное количество текстов, в основном по физике и математике, за последние 20 лет. — *Прим. ред.*

Гнусные цифры

Дуглас Арнольд и Кристин Фаулер*

Из-за (неуместного) акцентирования важности импакт-факторов повышение значения этого показателя становится целью, к достижению которой стремятся редакторы и издатели журналов. В свою очередь, это становится новым источником проблем. Закон Гудхарта предупреждает нас, что «когда достижение некоторого показателя становится целью, он перестает быть хорошим показателем». В точности так обстоит дело с импакт-факторами. Их ограниченная полезность еще больше компрометируется манипулированием импакт-факторами, такими действиями над этим пресловутым показателем качества журналов, которые увеличивают значение показателя, но не повышают (а в действительности — понижают) качество журналов.

Введение

Широко распространено использование импакт-фактора в качестве показателя качества журнала. Импакт-факторы применяются в библиотеках при принятии решений о подписке на периодические издания; исследователями, решающими, где стоит публиковаться и что стоит читать; университетскими рекомендательными комитетами, работающими в предположении, что публикация в журнале с более высоким значением импакт-фактора лучше характеризует специалиста; редакторами и издателями как средство оценки и продвижения своих журналов. Импакт-фактор для журнала в заданном году вычисляется ISI (*Thomson Reuters*) как среднее число появившихся в этом году ссылок на статьи данного журнала, опубликованные за два предыдущих года. Этот показатель подвергается критике с самых разных позиций^{1), 2), 3), 4)}:

Распределение ссылок на журнал не определяет его качество.

Импакт-фактор — это грубая статистика, доносящая только часть информации о распределении ссылок.

Оригинал: Douglas N. Arnold, Kristine K. Fowler, *Nefarious Numbers*, Notices of the American Mathematical Society 58(3) (2011), 434—437. Перевод С. Кузнецова (http://www.ispras.ru/ru/preprints/archives/rep_22_2010.php). Перевод названия одобрен автором.

* Дуглас Арнольд является Макнайтовским президентским профессором математики в университете Миннесоты и был в 2009—2010 г. президентом Общества промышленной и прикладной математики (Society for Industrial and Applied Mathematics, SIAM). Кристин Фаулер — сотрудник математической библиотеки университета Миннесоты. Авторы выражают признательность Сюзан Лори, которая разработала и поддерживала базу данных, использованную в этом исследовании, а также Молли Уайт.

Эта статистика дефектна. С одной стороны, распределение ссылок по статьям является сильно скошенным, так что усреднение по всем статьям журнала обычно бывает недостоверным. С другой стороны, в импакт-факторе учитываются ссылки, появившиеся только за два первых года после публикации (недостаток, особенно существенный для математики, где около 90 % ссылок появляется позже).

Используемая база данных дефектна, содержит ошибки и включает предвзятую подборку журналов.

Игнорируются многие побочные факторы, например тип статьи (редакционные, обзорные статьи, письма в противопоставление оригинальным исследовательским статьям), наличие нескольких авторов, самоцитирование, язык публикации и т. д.

Несмотря на эти сложности, для многих оказывается неотразимой привлекательность импакт-фактора как отдельного, легко доступного числового показателя, не требующего сложного разбирательства или привлечения экспертов, но претендующего на характеристику качества журнала. В 2000 г. в статье для информационного бюллетеня для редакторов журналов Амин и Мейб писали⁵⁾, что «в последние годы импакт-фактор из невразумительного библиометрического показателя превратился в основной количественный показатель качества журнала, его исследовательских статей, исследователей, пишущих эти статьи, и даже организаций, в которых они работают». Обычным явлением для журналов стало распространение абсурдных объявлений, рекламирующих их импакт-факторы, подобных тому, которое было разослано по всему миру издательством World Scientific, издающим *International Journal of Algebra and Computation* (IJAC): «Импакт-фактор IJAC вырос с 0,414 в 2007 г. до 0,421 в 2008 г.! Поздравляем редакционный совет и авторов IJAC». В этом случае возрастание импакт-фактора на 1,7 % представляет появление одной дополнительной ссылки на одну из 145 статей, опубликованных в журнале за два предыдущих года.

Из-за (неуместного) акцентирования важности импакт-факторов повышение значения этого показателя становится целью, к достижению которой стремятся редакторы и издатели журналов. В свою очередь, это становится новым источником проблем. Закон Гудхарта предупреждает нас, что «когда достижение некоторого показателя становится целью, он перестает быть хорошим показателем»⁶⁾. В точности так обстоит дело с импакт-факторами. Их ограниченная полезность еще больше компрометируется манипулиро-

ванием импакт-факторами, такими действиями над этим пресловутым показателем качества журналов, которые увеличивают значение показателя, но не повышают (а в действительности — понижают) качество журналов.

Возможно множество форм манипулирования импакт-факторами. В 2007 г. в очерке о вредоносных воздействиях манипулирования импакт-факторами Макдональд и Кэм⁷⁾ редактор вырачивает кадры постоянных авторов, которые надежно способствуют повышению измеряемого качества журнала путем самоцитирования и цитирования друг друга». Широко распространены жалобы авторов рецензируемых рукописей, которых редакторы просят или вынуждают сослаться на другие статьи из того же журнала. Поскольку судьба публикации статьи автора зависит от решения редактора, эта практика граничит с вымогательством, даже если соответствующее пожелание подается в форме всего лишь намека. В большинстве случаев можно лишь догадываться о наличии подобного давления, но в 2005 г. уже поступали сообщения о явных случаях принуждения от Монастырского⁸⁾ в *Chronicle of Higher Education* и Бегли⁹⁾ в *Wall Street Journal*. Третий установившийся метод, за счет которого редакторы повышают значения импакт-факторов своих журналов, состоит в публикации обзорных заметок с большим числом ссылок на данный журнал. Например, главный редактор *Journal of Gerontology A* взял за правило писать и публиковать в каждом январском номере журнала обзорную статью, опирающуюся на публикации предыдущих двух лет; в 2004 г. 195 из 277 ссылок указывали на *Journal of Gerontology A*. Хотя ущерб, наносимый этими ненаучными методами научной литературе, время от времени вызывает тревогу, многие люди считают влияние подобных методов минимальным или полагают их настолько легко опознаваемыми, что оказываемым ими воздействием можно пренебречь. Для обоснования наличия причин для серьезного беспокойства следует привести контрпример.

Пример IJNSNS

В области прикладной математики имеется наглядный пример, позволяющий исследовать такое искажение импакт-фактора. В последние несколько лет в списках импакт-факторов категории «Mathematics, Applied» доминирует *International Journal of Nonlinear Sciences and Numerical Simulation* (IJNSNS). В 2006—2009 г. этот журнал

занимал первое место, обычно со значительным отрывом, а в 2005 г. находился на втором месте. Однако, как мы увидим, более тщательное изучение показывает, что IJNSNS никоим образом не является ведущим в данной области. Попробуем понять происхождение высокого значения его импакт-фактора.

В 2008 г. (этот год мы рассмотрим наиболее детально) импакт-фактор IJNSNS составлял 8,91, будучи, безусловно, наивысшим среди 175 журналов, относимых к категории прикладной математики в отчетах о цитируемости журналов (Journal Citation Report, JCR) ISI. Для контроля мы также взглянем на два журнала из той же категории со вторым и третьим значениями импакт-фактора: *Communications on Pure and Applied Mathematics* (CPAM) и *SIAM Review* (SIREV), у которых в 2008 г. значения импакт-фактора составляли 3,69 и 2,80 соответственно. CPAM тесно связан с Курантовским институтом математических наук (Courant Institute of Mathematical Sciences), а SIREV является ведущим журналом Общества промышленной и прикладной математики (Society for Industrial and Applied Mathematics, SIAM)¹⁰. У обоих журналов превосходная репутация.

Наилучшей альтернативой оценки качества журналов по показателям цитируемости является экспертная оценка. У метода тщательного ранжирования журналов экспертами имеются собственные потенциальные проблемы, но, скорее всего, он обеспечивает намного более точную и целостную оценку качества журнала, чем импакт-фактор и подобные ему показатели. В математике, как и во многих других областях, у исследователей близко совпадают мнения относительно того, какие журналы являются наилучшими в их специальности. Австралийский исследовательский совет (Australian Research Council) недавно опубликовал результаты такого сравнения, в котором качественно оценены 20000 рецензируемых журналов разных дисциплин. Этот список разрабатывался на основе широкомасштабного процесса рецензирования с участием научных академий (таких как Австралийская академия наук (Australian Academy of Science)), сообществ представителей отдельных дисциплин (таких как Австралийское математическое общество (Australian Mathematical Society)) и многих исследователей и экспертов¹¹. Этот рейтинг будет использоваться в 2010 г. в рамках инициативы повышения уровня научных исследований в Австралии (Excellence in Research Australia, ERA), и его принято называть списком журналов ERA 2010 (ERA 2010 Journal List). Рейтинг каче-

ства журнала, который оценивает его «общее качество», представляется одной из четырех оценок:

A*: один из лучших в своей области или подобласти;

A: очень высокое качество;

B: надежная, хотя и не выдающаяся репутация;

C: не отвечает критериям высокого уровня.

В список ERA вошли 170 из 175 журналов, которым в 2008 г. в JCR был назначен импакт-фактор в категории «Mathematics, Applied». На рис. 1 показаны импакт-факторы журналов для каждого из четырех уровней рейтинга.

Мы видим, что импакт-фактор слабо соответствует оценке экспертов. Имеется много примеров журналов с более высокими зна-

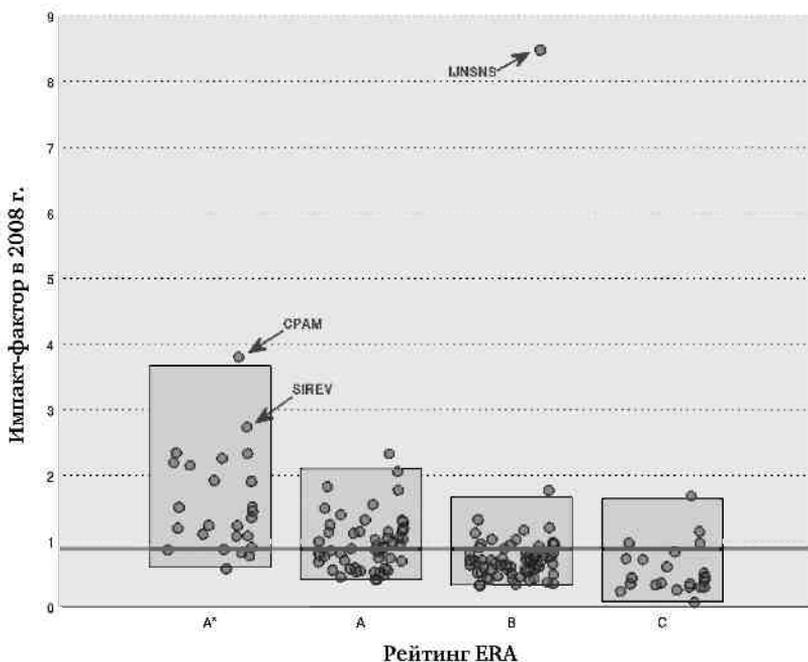


Рис. 1. Импакт-факторы в 2008 г. 170 журналов по прикладной математике, сгруппированные по своим уровням в рейтинге ERA. Для каждого уровня демонстрируется диапазон от 2,5-й до 97,5-й перцентили, что позволяет охарактеризовать средние 95%. Горизонтальные позиции точек данных внутри уровней выбраны случайным образом для улучшения наглядности. Толстая серая линия соответствует 20-й перцентили уровня A*.

чениями импакт-фактора, чем у других журналов, занимающих в рейтинге позиции на один, два и даже на три уровня выше. Под красной линией находится 20 % журналов уровня A*; примечательно, что более высокое значение импакт-фактора имеют 51 % журналов уровня A, 23 % журналов уровня B и даже 17 % журналов уровня C. Самый экстремальный выброс демонстрирует IJNSNS, который, несмотря на огромное значение импакт-фактора, находится не на первом или втором, а скорее на третьем уровне рейтинга. В рейтинге ERA наивысшую оценку A* получили 25 журналов. В этой группе представлено большинство журналов с наивысшими значениями импакт-фактора, включая SPAM и SIREV, но из 10 журналов с наивысшими значениями импакт-фактора два получили оценку A, и только IJNSNS попал в группу B. Всего имелось 53 журнала с рейтингом A и 69 журналов с рейтингом B. Если допустить, что IJNSNS является лучшим среди журналов группы B, то имелось бы 78 журналов с более высоким рейтингом ERA; если же он был бы худшим, то его рейтинг упал бы до 147. Коротко говоря, рейтинг ERA показывает, что IJNSNS не только не является лучшим журналом в области прикладной математики, но что его рейтинг находится где-то в диапазоне 75—150. Это удивительное несоответствие между репутацией журнала и его импакт-фактором нуждается в разъяснении.

Создание высокого импакт-фактора

На первом шаге на пути к пониманию происхождения высокого импакт-фактора IJNSNS мы взглянем на то, сколько авторов внесло существенный вклад в возрастание числа ссылок и кто такие эти авторы. Автором, который больше всего ссылался на IJNSNS в 2008 г., является главный редактор журнала Хэ Цзихуань ссылавшийся на журнал (на статьи, вышедшие в течение предыдущего двухлетнего окна) 243 раза. На втором месте со 114 ссылками находится Д. Д. Ганжи, который также является членом редакционного совета, и на третьем месте — региональный редактор Мохамед эль Наши с 58 ссылками. Совместно эти трое несут ответственность за 29 % ссылок, учтенных в импакт-факторе. Для сравнения, три автора, больше других ссылавшиеся на SIREV, сослались на журнал всего 7, 4 и 4 раза соответственно, что составляет менее 12 % учтенных ссылок, и ни один из этих авторов не участвовал в редактировании журнала. Для SPAM три автора, наиболее активно ссылавшиеся на журнал (9, 8 и 8 ссылок), привнесли около 7 % от общего числа ссылок, и опять

же никто из них не состоял в редакционном совете. Другое заметное явление представляет уровень концентрации ссылок на статьи, опубликованные в пределах двухлетнего окна, которое используется для подсчета импакт-фактора. Наш анализ ссылок, появившихся в 2008 г., на статьи, опубликованные после 2000 г., показывает, что 16 % ссылок на CPAM относится к статьям этого двухлетнего окна, а в случае SIREV таких ссылок всего 8 %. По сравнению с этим, 71,5 % ссылок на IJNSNS ведут на статьи, опубликованные в пределах двухлетнего окна.

В таблице 1 мы показываем импакт-факторы этих трех журналов в 2008 г., а также модифицированные значения импакт-фактора, соответствующие среднему числу ссылок в 2008 г. на статьи, опубликованные в журналах не в 2006—2007 г., а за предыдущие шесть лет. Поскольку время полураспада цитирования (время, за

Таблица 1

Журнал	Импакт-фактор в 2008 г. при использовании обычного окна 2006—2007 г.	Модифицированный «импакт-фактор» в 2008 г. при использовании окна 2000—2005 г.
IJNSNS	8,91	1,27
CPAM	3,69	3,46
SIREV	2,8	10,4

которое появляется половина от окончательного числа ссылок на статью) для прикладной математики составляет около 10 лет¹²⁾, этот показатель, по крайней мере, имеет не меньше смысла, чем импакт-фактор. Кроме того, он независим (в отличие от пятилетнего импакт-фактора JCR), поскольку его период времени не перекрывается с тем, который используется при вычислении двухлетнего импакт-фактора. Заметим, что при использовании другого окна цитирования импакт-фактор IJNSNS стремительно падает, уменьшаясь в семь раз. В отличие от этого, импакт-фактор CPAM остается почти неизменным, а импакт-фактор SIREV заметно возрастает. Можно просто заметить, что, в отличие от контрольных образцов, в ссылках на IJNSNS в 2008 г. наблюдается значительное предпочтение статей, опубликованных точно в течение тех двух лет, которые используются при вычислении импакт-фактора.

Дальнейшее прозрение приходит при анализе журналов, а не авторов с большим числом ссылок. Учет журнальных самоцитиро-

ваний в импакт-факторе часто подвергается критике, и в данном случае этот подход действительно используется. В 2008 г. в IJNSNS появились 102 ссылки на статьи этого же журнала, или 7% от общего числа ссылок, учитывавшихся при вычислении импакт-фактора. Соответствующими цифрами являются 1 ссылка (0,8%) для SIREV и 8 ссылок (2,4%) для SPAM. Расхождения для других недавних лет настолько же или еще более велики.

Однако наибольшее число ссылок на IJNSNS обеспечил *Journal of Physics: Conference Series*. В одном выпуске этого журнала содержались 294 ссылки на статьи, опубликованные в пределах окна подсчета импакт-фактора, что составляет более 20% от общего числа ссылок, на основе которых подсчитывался импакт-фактор. И что же это был за выпуск? Это были труды конференции, организованной главным редактором IJNSNS Хэ Цзихуань в его родном университете. Он отвечал за рецензирование статей этого выпуска. На втором месте по числу ссылок на IJNSNS находился журнал *Topological Methods in Nonlinear Analysis*, который обеспечил 206 ссылок (14%), и все ссылки были размещены в одном выпуске журнала. Это был специальный выпуск с Хэ Цзихуань в качестве приглашенного редактора; его соредактор Сюй Лань также является членом редакционного совета IJNSNS. Сам Хэ Цзихуань в этом выпуске опубликовал краткую статью из трех страниц текста и 30 ссылок. Из них 20 ссылок вели на статьи IJNSNS, опубликованные в пределах окна подсчета импакт-фактора. Среди оставшихся ссылок имелось 8 ссылок на Хэ Цзихуань и 2 на Сюй Лань.

Продолжая спускаться по списку журналов, активно ссылающихся на IJNSNS, можно обнаружить еще один подобный случай: 50 ссылок из одного выпуска *Journal of Polymer Engineering* (как и IJNSNS, он публикуется издательством Freund), приглашенными редакторами которого является та же пара Хэ Цзихуань и Сюй Лань. Однако третье место занимает журнал *Chaos, Solitons & Fractals* (CS&F) со 154 ссылками, разбросанными по разным номерам. Это опять такие ссылки, которые можно считать обусловленными влиянием или контролем редакторов. В 2008 г. Хэ Цзихуань состоял в редакционном совете CS&F, а главным редактором являлся Мохамед эль Наши, который также был соредактором IJNSNS. При обстоятельствах, получивших широкую огласку, весь редакционный совет CS&F недавно был заменен, но эль Наши остался соредактором IJNSNS.

Многие другие ссылки на IJNSNS появлялись в статьях, публикуемых в журналах, редактором которых являлся Хэ; например,

40 ссылок обеспечил журнал *Zeitschrift für Naturforschung A*. Таких журналов слишком много, чтобы можно было их здесь перечислить, поскольку Хэ исполнял редакторские функции более чем в 20 журналах (а теперь назван главным редактором еще четырех журналов недавно образованного издательства Asian Academic Publishers). Еще одним источником многочисленных ссылок являлись статьи, написанные другими редакторами IJNSNS. В общей сложности, совокупность таких родственных редакторам ссылок, для выявления которых требуется много времени, включает более 70 % ссылок, использованных для вычисления импакт-фактора IJNSNS.

Библиометрия для отдельных личностей

Библиометрия также используется для оценки отдельных людей, статей, институтов и даже государств. Наукометрическая база данных ключевых научных показателей (Essential Science Indicators), поддерживаемая компанией *Thomson Reuters*, рекламируется как средство для ранжирования «ведущих стран, журналов, ученых, статей и институтов в разных исследовательских областях». Однако эти показатели, главным образом, основываются на тех же данных о ссылках, которые используются для подсчета импакт-фактора, и такими показателями так же легко манипулировать (на самом деле, это делается одновременно). В специальном выпуске *Journal of Physics: Conference Series*, который редактировался Хэ и собрал 243 ссылки на его журнал, также содержатся 353 ссылки на самого Хэ. Он утверждает, что общее число ссылок на его статьи превышает 6800¹³⁾. Достижение даже половинного значения этого показателя считается весьма примечательным событием, о чем свидетельствует следующее сообщение, опубликованное на сайте ScienceWatch.com¹⁴⁾: «По данным недавнего анализа ключевых научных показателей от *Thomson Scientific*, профессор Хэ Цзихуань назван восходящей звездой в области компьютерной науки... В Web of Science зарегистрированы его 137 статей, которые к настоящему времени цитировались 3193 раза». В составе дюжины ученых из всех областей науки Хэ упоминался в списках ESI «Наиболее активные исследования 2007—2008 г.» (Hottest Research of 2007-8) и «Наиболее активные исследования 2009 г.».

Другим популярным показателем продуктивности исследователей, основанным на учете ссылок, является *h*-индекс (индекс Хир-

ша). h -индекс отдельного человека — это наибольшее число его (или ее) статей, каждая из которых цитировалась не менее того же числа раз. На этот показатель также действует закон Гудхарта. Хэ утверждает, что его h -индекс составляет 39, в то время как сам Хирш оценивает среднее значение своего индекса для Нобелевских лауреатов в области физики как 35¹⁵⁾. Оценки отдельных исследователей или журналов, основанные на учете ссылок, не заменяют экспертных оценок качества.

Заключительные размышления

Несмотря на многочисленные недостатки, импакт-фактор широко используется как показатель качества журналов и даже статей и авторов. Это порождает стимул для манипулирования данным показателем. Более того, можно значительно повысить значение импакт-фактора вообще без повышения качества журнала. Огромное влияние на импакт-фактор может иметь деятельность всего лишь нескольких заинтересованных людей, причем для разоблачения этой деятельности требуются значительные усилия. Мы, главным образом, обсуждали один экстремальный пример, но вряд ли стоит сомневаться в том, что подобные методы применяются во многих журналах, хотя и в меньшей степени (и из-за этого их труднее обнаружить). Совокупным результатом недостатков общего подхода и использования методов манипулирования является то, что импакт-фактор обеспечивает очень неточное представление о качестве журналов. В более общем смысле, ссылки, образующие основу импакт-фактора, и различные другие библиометрические показатели ненадежны по своей природе.

Эта плачевная ситуация имеет огромные последствия. Неверным способом назначаются награды, научная литература и предметные области деформируются, вокруг всего этого растет уровень скептицизма. Что же делать? Как и в самих научных исследованиях, нужно противиться искушению принятия упрощений, когда это может нанести серьезный ущерб точности. Порицаются ученые, поддающиеся искушению скрыть некоторые данные или исказить статистику, чтобы получить более понятные результаты. Мы должны придерживаться аналогичного уровня принципиальности при оценке результатов исследований. Администраторы, финансирующие организации, библиотекарки и все прочие, кому требуются такие оценки, должны отказаться от упрощенных подходов и прини-

мать важные решения на основе вдумчивости, здравого смысла и экспертизы.

Примечания

¹⁾ P. O. Seglen, *Why the impact factor of journals should not be used for evaluating research*, *BMJ* **314** (1997), 498—502.

²⁾ J. Ewing, *Measuring journals*, *Notices of the AMS* **53** (2006), 1049—1053.

³⁾ R. Golubic, M. Rudes, N. Kovacic, M. Marusic, and A. Marusic, *Calculating impact factor: how bibliographical classification of journal items affects the impact factor of large and small journals*, *Sci. Eng. Ethics* **14** (2008), 41—49.

⁴⁾ R. Adler, J. Ewing, P. Taylor, *Citation statistics*, *Statistical Sciences* **24** (2009), 1—14.

⁵⁾ M. Amin, M. Mabe, *Impact factors: use and abuse*, *Perspectives in Publishing* **1** (2000), 1—6.

⁶⁾ Эта сжатая формулировка взята из статьи M. Strathern, „*Improving Ratings*“: *Audit in the British University System*, *European Review* **5** (1997), 305—321.

⁷⁾ S. Macdonald, J. Kam, *Aardvark et al.: quality journals and gamesmanship in management studies*, *Journal of Information Science* **33** (2007), 702—717.

⁸⁾ R. Monastersky, *The number that's devouring science*, *Chronicle of Higher Education* **52** (2005).

⁹⁾ S. Begley, *Science journals artfully try to boost their rankings*, *Wall Street Journal*, 5 June 2006, B1.

¹⁰⁾ Первый автор данной статьи в 2009—2010 г. был президентом SIAM.

¹¹⁾ Australian Research Council, *Ranked Journal List Development*. http://www.arc.gov.au/era/journal_list_dev.htm.

¹²⁾ В 2010 г. в JCP категории «Mathematics, Applied» было установлено совокупное время полураспада цитирования в 9,5 лет.

¹³⁾ Это утверждение, а также то, что его *h*-индекс составляет 39, содержится в биографических сведениях к одной из его недавних статей (*Nonl. Sci. Letters* **1** (2010), 1).

¹⁴⁾ ScienceWatch.com, April 2008. <http://sciencewatch.com/inter/aut/2008/08-apr/08aprHe/>.

¹⁵⁾ J. Hirsch, *An index to quantify an individual's scientific research output*, *PNAS* **102** (2005), 16569—16572.

Из интервью журналу «Эксперт»

Л. Д. Фаддеев

А. М.: Проблема, которую постоянно обсуждают, о которой говорят, — как измерить эффективность затрат в фундаментальной науке. Многие рассматривают цитируемость и количество публикаций в качестве главного критерия.

Л. Ф.: Я с этим категорически не согласен. Чтобы никто не заподозрил меня в заинтересованности, сразу скажу, что мой индекс цитируемости хороший*. Но я утверждаю, что индекс цитирования — это только лишь один из многих показателей, по которым может оцениваться научная работа.

Международный математический союз посвятил вопросам статистической оценки труда ученых специальный доклад, который так и называется «Citation statistics», то есть «Статистики цитирования»**. Я вам приведу его основные выводы: «Мы не отказываемся от статистики цитирования как способа оценки качества исследований. Но если мы хотим, чтобы эта оценка приносила пользу, то индекс цитирования может быть только частью оценки. Потому что он представляет собой только ограниченный и неполный взгляд на качество исследований, и статистика, выведенная из данных цитирования, зачастую понимается и используется неверно. Научные исследования настолько сложная вещь, что их важность невозможно измерить только единственным грубым средством».

Надо также понимать, что индекс цитирования, а особенно импакт-параметр, то есть относительная частота цитирования «средней» статьи данного журнала в течение определенного периода времени, очень зависит от специальности. В этом же докладе Международного математического союза есть прекрасный график***, который это иллюстрирует.

А. М.: Некоторые наши работающие здесь ученые говорят: публиковаться за границей их вынуждает то обстоятельство, что наши журналы за рубежом читают, но не цитируют.

Источник: А. Механик, *Другой академии у нас нет*, Эксперт 5(691), 8 февраля 2010 г. © ЗАО «Группа Эксперт», 2010.

* Общее число цитирований более 17 тысяч, индекс Хирша не меньше 50. — *Прим. ред.*

** См. наст. изд., с. 6—38. — *Прим. ред.*

*** См. наст. изд., с. 16. — *Прим. ред.*

Л. Ф.: К сожалению, такое бывает. Когда некоторые заграничные ученые решили, что индекс цитирования важен для престижа и, главное, для получения грантов, они стали искусственно повышать свои индексы. Мы знаем примеры, когда целые лаборатории договариваются цитировать друг друга: мы цитируем вас, вы цитируете нас. Конечно, чужие им не нужны. Это конкуренция.

Приложение

Из ТИПОВОЙ МЕТОДИКИ

оценки результативности деятельности научных организаций, выполняющих научно-исследовательские, опытно-конструкторские и технологические работы гражданского назначения*

9. На основе показателей оценки результативности деятельности научной организации она может быть отнесена к одной из следующих категорий:

- а) 1 категория — научные организации-лидеры;
- б) 2 категория — стабильные научные организации, демонстрирующие удовлетворительную результативность;
- в) 3 категория — научные организации, утратившие научный профиль и перспективы развития.

Из приложения: ПОКАЗАТЕЛИ

оценки результативности деятельности научных организаций, подведомственных федеральным органам исполнительной власти или государственным академиям наук, выполняющих научно-исследовательские, опытно-конструкторские и технологические работы гражданского назначения (данные за последние пять лет с разбивкой по годам)

1. Научный потенциал и эффективность научных исследований

1.1. Общая характеристика научного потенциала

- Удельный вес внутренних затрат на исследования и разработки в общем объеме выполненных научной организацией работ, услуг (%)
- Внутренние затраты на исследования и разработки отнесенные к численности исследователей (тыс. рублей/чел.)
- Число защищенных за период оценки работниками научной организации докторских и кандидатских диссертаций, отнесенное к численности исследователей
- Перечень государственных и международных премий, призов, наград, почетных званий, полученных научной организацией или отдельными ее работниками за период оценки

Источник: http://www.edu.ru/db-mon/mo/Data/d_09/prm406-2.htm.

* Утверждено приказом Министерства образования и науки Российской Федерации от 14 октября 2009 г. №406.

1.2. Публикационная активность

– Число публикаций работников научной организации в Российском индексе научного цитирования (далее — РИНЦ), отнесенное к численности исследователей (за каждый год из последних пяти лет, начиная с года, предшествующего текущему)

– Цитируемость работников научной организации в РИНЦ (Общее число ссылок на публикации работников научной организации в РИНЦ (за каждый год из последних пяти лет, начиная с года, предшествующего текущему), отнесенное к численности исследователей научной организации в году, предшествующем текущему)

– Число публикаций работников научной организации в Web of Science, отнесенное к численности исследователей (Число публикаций за последние пять лет (сумма, начиная с года, предшествующего текущему). Численность исследователей в году, предшествующем текущему)

– Цитируемость работников научной организации в Web of Science (Общее число ссылок на публикации работников научной организации в Web of Science (сумма за последние пять лет, начиная с года предшествующего текущему), отнесенное к численности исследователей научной организации в году, предшествующем текущему)

– Импакт-фактор публикаций работников научной организации в Web of Science (Число публикаций работника научной организации в журнале, зарегистрированном в Web of Science (за каждый год из последних пяти лет, начиная с года, предшествующего текущему) умножается на импакт-фактор данного журнала в соответствующем году; сумма полученных значений (по годам, работникам и журналам) делится на общее (за последние пять лет) число публикаций работников научной организации в Web of Science)

– Число опубликованных докладов, тезисов докладов, представленных работниками научной организации на крупных конференциях, симпозиумах и чтениях, (более 150 участников), а также конференциях, организованных в соответствии с планами федеральных органов исполнительной власти, государственных академий наук или на средства российских и международных фондов (включая РФФИ и РГНФ), отнесенное к численности исследователей

1.3. Объекты интеллектуальной собственности

– Число монографий и глав в монографиях, учебников и глав в учебниках, отнесенное к численности исследователей

– Число охраняемых объектов интеллектуальной собственности, принадлежащих научной организации, отнесенное к численности исследователей

– Число отечественных и зарубежных патентов (свидетельств) на объекты интеллектуальной собственности, полученных научной организацией и ее работниками за период оценивания, отнесенное к численности исследователей

2. Вовлеченность научной организации в национальное и мировое научно-образовательное сообщество

3. Коммерциализация и прикладное значение результатов исследований

4. Кадровая обеспеченность научной организации

5. Ресурсная обеспеченность научной организации

6. Состояние финансовой деятельности научной организации

Дополнительная литература*

1. J. Ewing, *Judging Journals*, 2000.
<http://www.ams.org/ewing/Documents/Judgingjournals.pdf>
2. И. Маршакова-Шайкевич, *Вклад России в развитие мировой науки: библиометрическая оценка*, Отечественные записки 7 (2002).^{www}
3. L. Butler, *Modifying publication practices in response to funding formulas*, *Research Evaluation* 17:1 (2003), 39—46.^{www}
4. J. Ewing, *Why the AMS does not provide journal usage statistics* 2004. <http://www.ams.org/ewing/Documents/NoStatistics-43.pdf>
5. С. М. Гусейн-Заде, *Вокруг импакт-фактора*, *Успехи матем. наук*, 59:5 (2004), 186—188.^{www}
6. J. Ewing**, *Measuring Journals*, *Notices of the Amer. Math. Soc.* 53:9 (2006), 1049—1053.^{www}
7. Л. М. Гохберг, Г. С. Сагиева, *Российская наука: библиометрические индикаторы*, *Форсайт* 1 (2007), 44—53.^{www}
8. А. Б. Жижченко, А. Д. Изаак, *Информационная система Math-Net.Ru. Применение современных технологий в научной работе математика*, *Успехи матем. наук* 62:5(377) (2007), 107—132.^{www}
9. J. Glaser, G. Laudel, *Evaluation without Evaluators: The impact of funding formulae on Australian university research*, in: R. Whitley, K. Barker, J. Glaser (Hg.), *The Changing Governance of the Sciences: The Advent of Research Evaluation Systems*, *Yearbook Sociology of the Sciences*, 2007, Dordrecht.
10. *Ethics in Science and Environmental Politics* 8:1 (2008), ESEP theme section: *The use and misuse of bibliometric indices in evaluating scholarly performance* (цикл статей о библиометрике).
<http://www.int-res.com/abstracts/esep/v8/n1/>
11. В. Г. Горохов, Г. Бехманн, *Изменение роли науки в обществе: поиск новых идеалов в научной системе Германии*, *Вестник РАН* 80:3 (2010), 258—266.

* Здесь мы даем ссылки на дополнительную литературу и сайты по вопросам, затронутым в этом сборнике. В работах [3, 9, 11] изложено как преимущественное использование библиометрических показателей (числа публикаций) для оценки научных результатов привело к резкому понижению качества исследований в масштабе целой страны (Австралия). Мы признательны В. Г. Горохову, любезно предоставившему ссылки на эти работы. Интернет-адреса многих из приведенных здесь источников содержатся в электронной версии данного сборника.

** См. также другие его статьи на сайтах <http://www.ams.org/ewing/>, <http://mathematicsjournals.blogspot.com/>

12. Nature **465**, 17 June 2010 (цикл статей о библиометрике).
<http://www.nature.com/news/specials/metrics/index.html>
13. Н. Быкова, *Новое руководство РАМН: ориентация на индекс Хирша* 24.03.2011.
http://www.strf.ru/material.aspx?CatalogId=221&d_no=38162
14. Н. В. Мотрошилова, *Недоброкачественные сегменты наукометрии*, Вестник РАН **81**:2 (2011), 134—146.

Сведения об авторах

Адлер Роберт (Adler Robert) — профессор Израильского технологического института (Technion), Хайфа, основные работы по математической статистике, теории случайных процессов и их приложениям к геометрии.

Арнольд Дуглас (Arnold Douglas) — профессор университета Миннесоты (США), в 2009—2010 г. президент SIAM*, основные работы по численному анализу, дифференциальным уравнениям с частными производными, теории устойчивости в механике.

Кемпбелл Филипп (Campbell Philip) — физик по образованию, главный редактор журнала *Nature*.

Ловаш Ласло (Lovasz Laslo) — директор Математического института Будапештского университета, в 2007—2010 г. президент Международного математического союза**, основные работы по комбинаторике, теории алгоритмов, теории сложности, случайным процессам.

Лоуренс Питер А. (Lawrence Peter A.) — профессор Кембриджского университета (Великобритания), Департамент зоологии и Лаборатория молекулярной биологии, основные работы по биологии развития, возникновению и генетическому контролю формы и размера клеточных структур в процессе эмбриогенеза.

Тейлор Питер (Taylor Peter) — профессор университета Мельбурна (Австралия), основные работы по теории марковских процессов, моделированию биологических систем, стохастическому анализу логических сетей, телекоммуникационным системам.

Фаддеев Людвиг Дмитриевич — директор Эйлеровского международного математического института (Санкт-Петербург), академик-секретарь Отделения математических наук РАН, в 1987—1990 г.

*SIAM — Общество промышленной и прикладной математики. Возникло в 1951 г. Объединяет около 12 тысяч математиков-прикладников и представителей близких профессий, в основном из Северной Америки (www.siam.org). Организует многочисленные конференции. Издает 14 престижных журналов. См. http://en.wikipedia.org/wiki/Society_for_Industrial_and_Applied_Mathematics.

**Международный математический союз объединяет национальные математические сообщества 77 стран. Организует, среди прочего, каждые четыре года международные математические конгрессы, на которых присуждается наиболее известная из математических премий — медаль Филдса. См. www.mathunion.org.

президент Международного математического союза, основные работы по квантовой теории рассеяния, спектральной теории операторов, квантованию неабелевых калибровочных теорий поля, теории бесконечномерных интегрируемых систем, общей теории относительности.

Фаулер Кристин (Fowler Kristine) — сотрудник математической библиотеки университета Миннесоты.

Эвинг Джон (Ewing John) — в 1995—2008 г. исполнительный директор Американского математического общества*.

* Возникло в 1894 г. Объединяет более 30 тысяч как американских математиков, так и математиков других стран (см. www.ams.org). Публикует целый ряд престижных журналов и *Mathematical Reviews*, реферативный журнал по математике (о нем см. в отчете «Статистики цитирования»).

ИГРА В ЦЫФИРЬ,
или как теперь оценивают труд ученого
(сборник статей о библиометрике)

Издательство Московского центра
непрерывного математического образования
119002, Москва, Большой Власьевский пер., 11. Тел. (499) 241-74-83

Подписано в печать 11.04.2011 г. Формат 60×90¹/₁₆.
Бумага офсетная. Печать офсетная. Печ. л. 4,5.

Отпечатано с готовых диапозитивов в ООО «Принт Сервис Групп».
Москва, 2-й Лихачёвский пер., д. 7.

Книги издательства МЦНМО можно приобрести в магазине
«Математическая книга», Большой Власьевский пер., д. 11.
Тел. (499) 241-72-85. E-mail: biblio@mcsme.ru
