

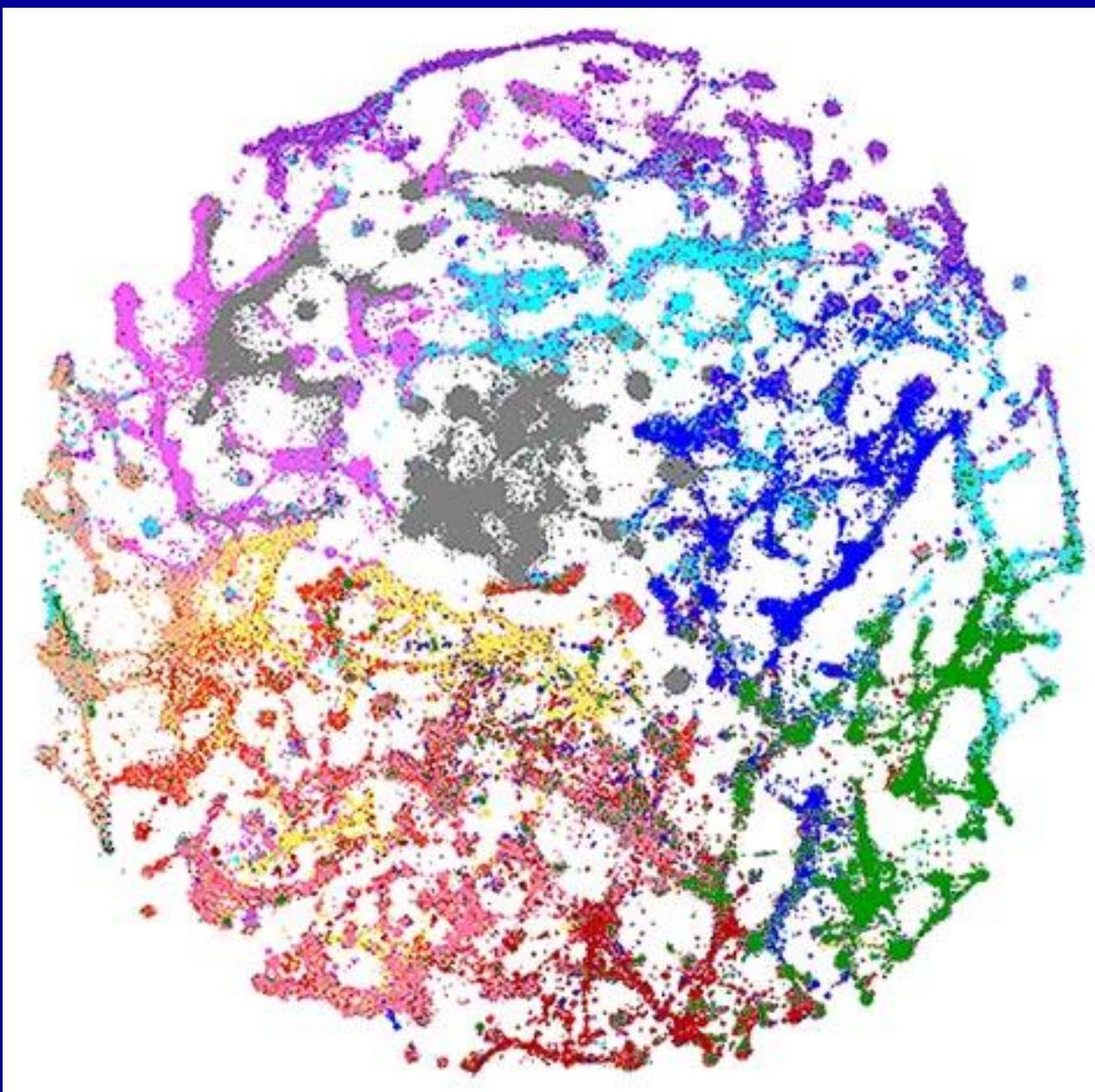
# Как сделать научные исследования более достоверными и полезными

Джон П.А. Иоаннидис, доктор медицинских наук  
(MD), доктор наук (DSc)

Кафедра профилактики заболеваний К. Ф. Ренборга,  
Профессор медицины, исследований и политики в области здоровья,  
науки о биомедицинских данных и статистики Стэнфордского  
университета,  
Содиректор Стэнфордского инновационного центра мета-исследований  
(METRICS)

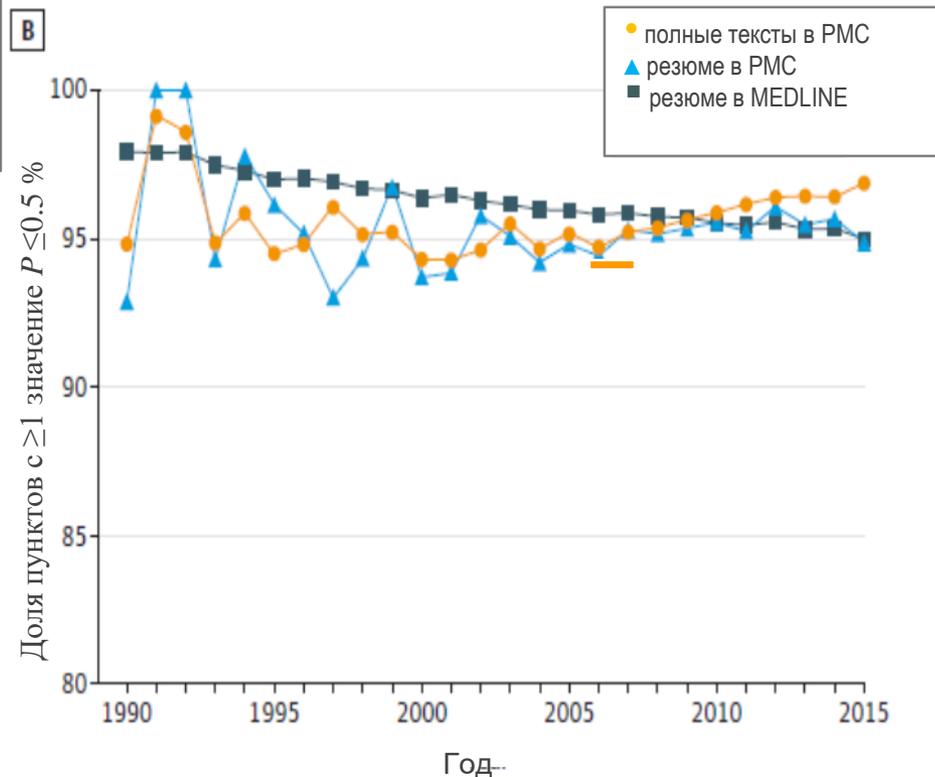
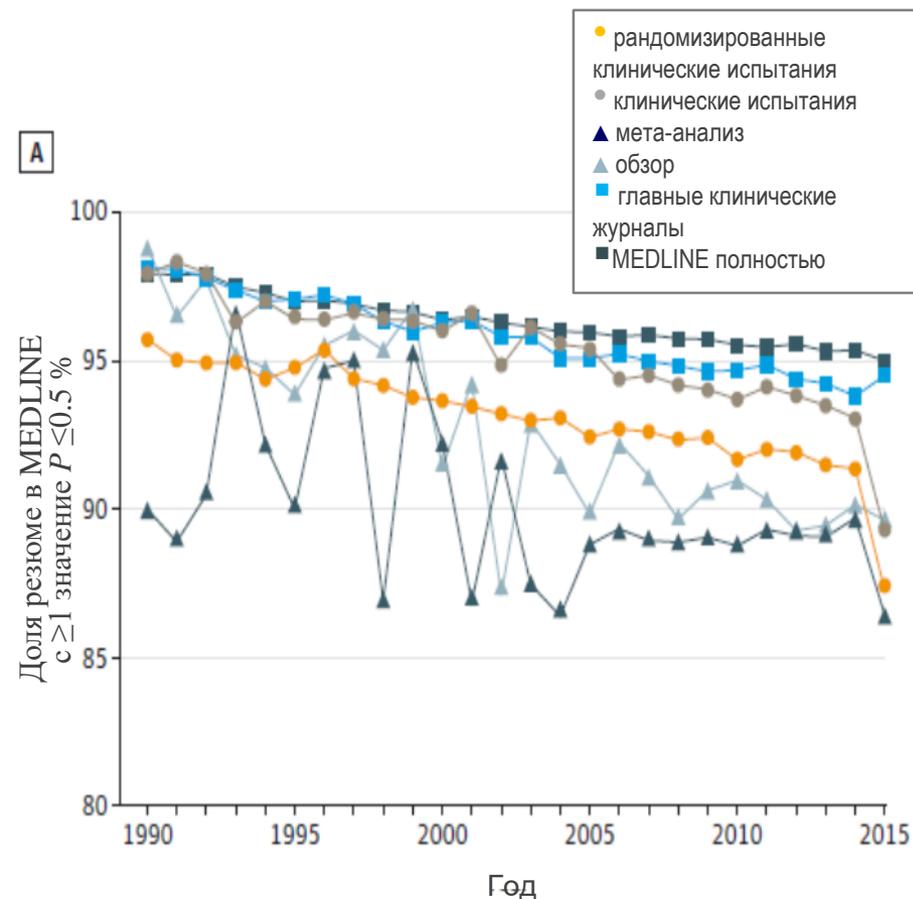
# Наука: с высоты птичьего полета

- 15 153 100 разных ученых публиковали статьи в крупных научных журналах в 1996-2011 гг. (Scopus)
- 25 805 462 разных статей опубликовано за тот же период и индексируется в Scopus
- По оценкам, 160 миллионов научных документов в Google Scholar
- Каждая эмпирическая статья может содержать от нескольких до многих миллионов результатов

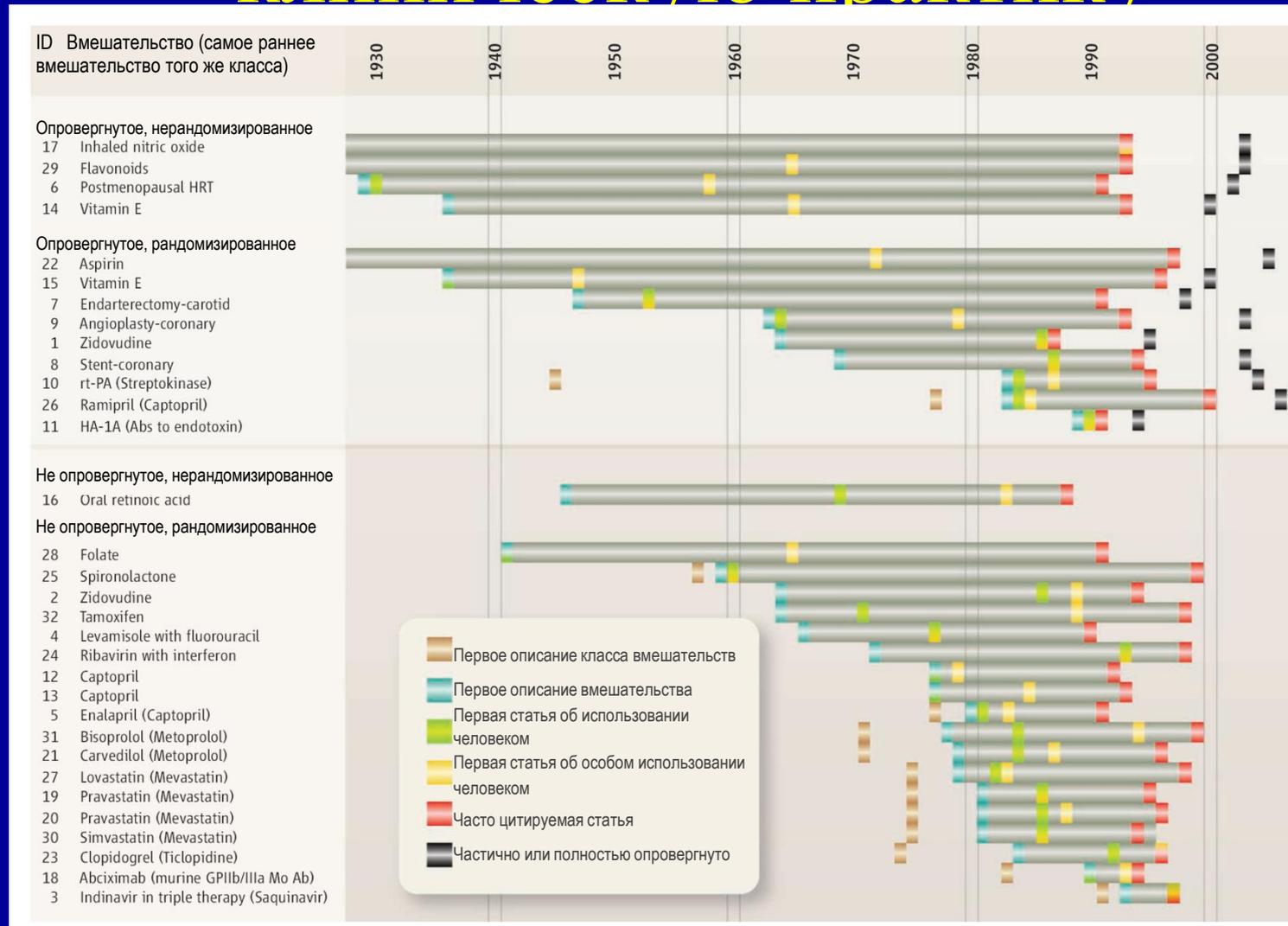


Карта современной науки: 20 миллионов статей, 2 миллиона патентов, 200000 кластеров, сохраняющихся 2-16 лет каждый

# Научное открытие стало досадным недоразумением: в 96% биомедицинской литературы заявляется о получении важных результатов



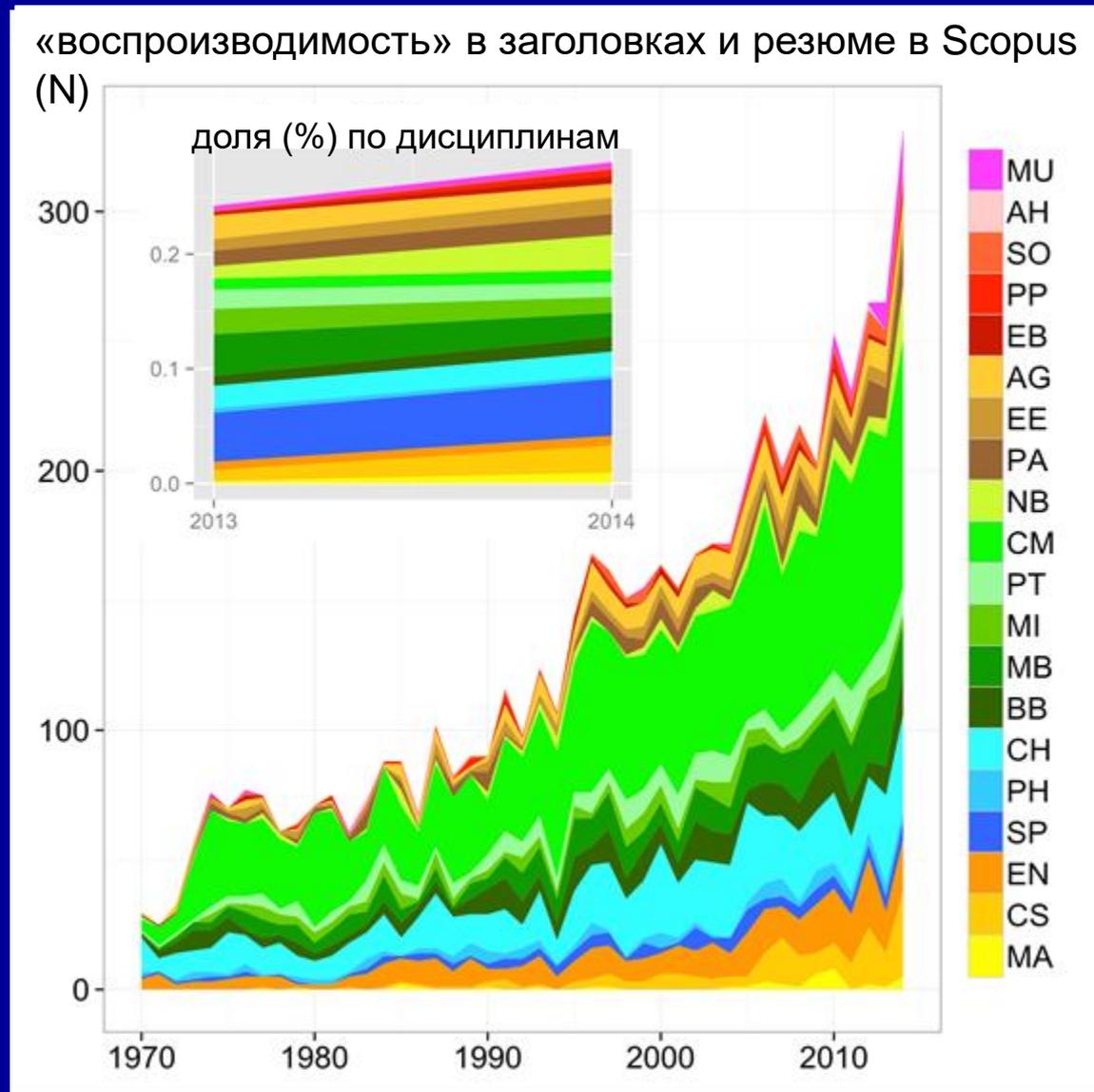
# Замедленный темп внедрения в клиническую практику



# Стандартный рецепт исследовательской практики

- Самостоятельный, изолированный исследователь
- Исследования на малой выборке
- Придирчивый отбор одной/лучшей гипотезы
- Post-hoc
- Достаточно, чтобы  $P < 0.05$
- Нет учета
- Нет воспроизведения
- Нет совместного использования данных

Поиск: воспроизводимость научных исследований; воспроизводимость результатов;  
воспроизводимость исследований; воспроизводимые исследования; воспроизводимый  
результат.



# Различные типы воспроизводимости

- **Воспроизводимость методов:** возможность понять или повторить экспериментальные и вычислительные процедуры настолько точно, насколько возможно.
- **Воспроизводимость результатов:** возможность получить подкрепляющие результаты в ходе нового исследования при применении таких же экспериментальных методов.
- **Воспроизводимость заключений:** такие же веские заявления о полученных знаниях при воспроизведении исследования.

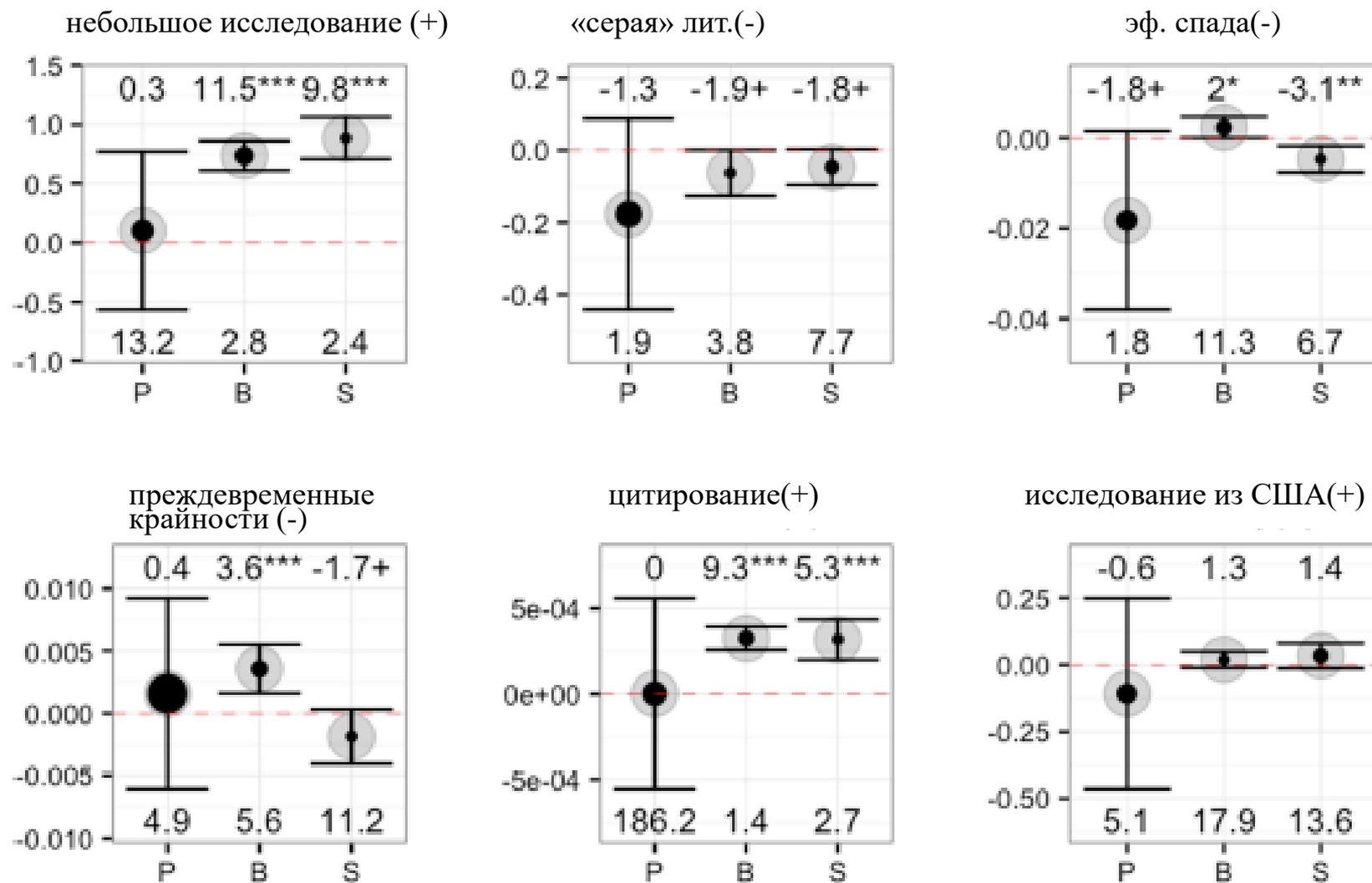
# Различия между отраслями, влияющие то, что значит «воспроизводимое исследование»

- Степень детерминизма
- Измерение коэффициента ошибок
- Сложность разработок/измерительных инструментов
- Близость соответствия между гипотезой и экспериментальной разработкой/данными.
- Статистические/аналитические метод проверки гипотез
- Типичная гетерогенность результатов экспериментов
- Культура проведения повторных испытаний, прозрачности и накопления знаний
- Статистические критерии проверки истинности заявлений
- Цели применения результатов и последствия неверных выводов.

# Мета-оценка предвзятости в науке

Daniele Fanelli <sup>1</sup>, Rodrigo Costas <sup>2</sup>, John PA Ioannidis <sup>1,3</sup>

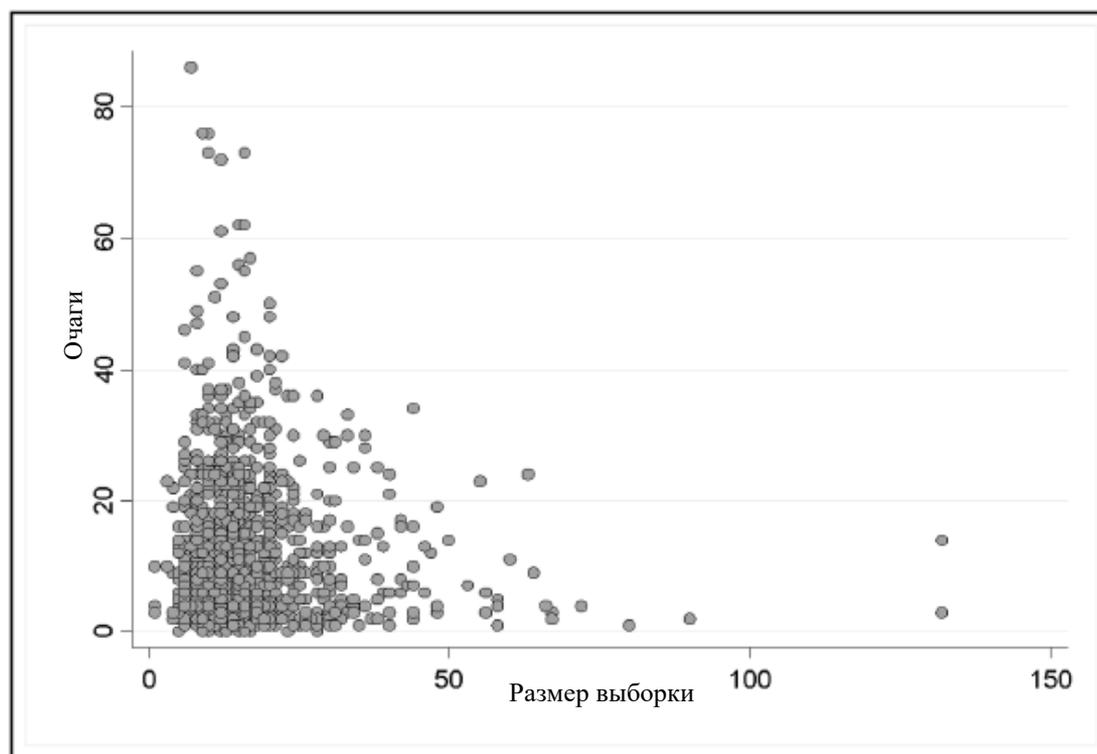
Проверяемая гипотеза	Переменные, измеренные для проверки гипотезы	Предположенная связь с размером эффекта	Результат		
Постулируемые шаблоны предвзятости	эффект небольшого исследования	стандартная ошибка в исследовании	+	S	
	эффект «серой» литературы	«серая литература» (любого типа) vs. статья в журнале	-	S	
	эффект спада	порядок лет в МА	-	P	
	преждевременные крайности	порядок лет в МА, в обратном порядке по абсолютной величине эффекта	-	N	
	необъективность цитирования	всего цитирований на исследование	+	S	
	США-эффект	Исследование автора из США vs. любой другой страны	+	P	
Постулируемые факторы риска для возникновения предвзятости	предвзятость в отрасли	Исследования авторов, связанных с частной отраслью vs. нет	+	P	
	публикация под давлением	политика страны	денежный стимул	+	N
		продуктивность автора	карьерный стимул	+	N
			влияние автора	институциональный стимул	+
	взаимный контроль	профессиональный уровень		(первая/последняя) всего авторских публикаций, публикаций/год	+
			(первое/последнее) всего цитирований, цитирований в среднем, нормализованных цитирований в среднем, среднее влияние в журналах, %	+	N
	индивидуальные факторы риска	пол	топ-10 журналов	-	S
			размер группы	-	S
		честность исследования	страны/автор, среднее расстояние между адресами	+	S
		годы деятельности (первый/последний) автора	-	S	
	(первое/последнее) женское имя автора	-	N		
	(первый/последний) автор с $\geq 1$ отказом)	+	P		



**Рис. 2.** Шаблоны предвзятостей, разделенные по дисциплинарным областям. Каждая панель пре-

# Исследования фМРТ: только очень небольшие исследования обнаруживают множество очагов

Необъективность в исследованиях фМРТ



**Рисунок 3. Отношение между числом очагов и размером выборки исследования.** График рассеяния очагов и размер выборки по каждому набору данных (N=1778) по всем мета-анализам  
[doi:10.1371/journal.pone.0070104.g003](https://doi.org/10.1371/journal.pone.0070104.g003)

# Недостаточная мощность: почему маленький размер выборки подрывает надежность в нейронауке

*Katherine S. Button<sup>1,2</sup>, John P. A. Ioannidis<sup>3</sup>, Claire Mokrysz<sup>1</sup>, Brian A. Nosek<sup>4</sup>, Jonathan Flint<sup>5</sup>, Emma S. J. Robinson<sup>6</sup> and Marcus R. Munafò<sup>1</sup>*

Abstract | A study with low statistical power has a reduced chance of detecting a true effect, but it is less well appreciated that low power also reduces the likelihood that a statistically significant result reflects a true effect. Here, we show that the average statistical power of studies in the neurosciences is very low. The consequences of this include overestimates of effect size and low reproducibility of results. There are also ethical dimensions to this problem, as unreliable research is inefficient and wasteful. Improving reproducibility in neuroscience is a key priority and requires attention to well-established but often ignored methodological principles.

А. Размеры эффектов, журналы, подобласти и мощность

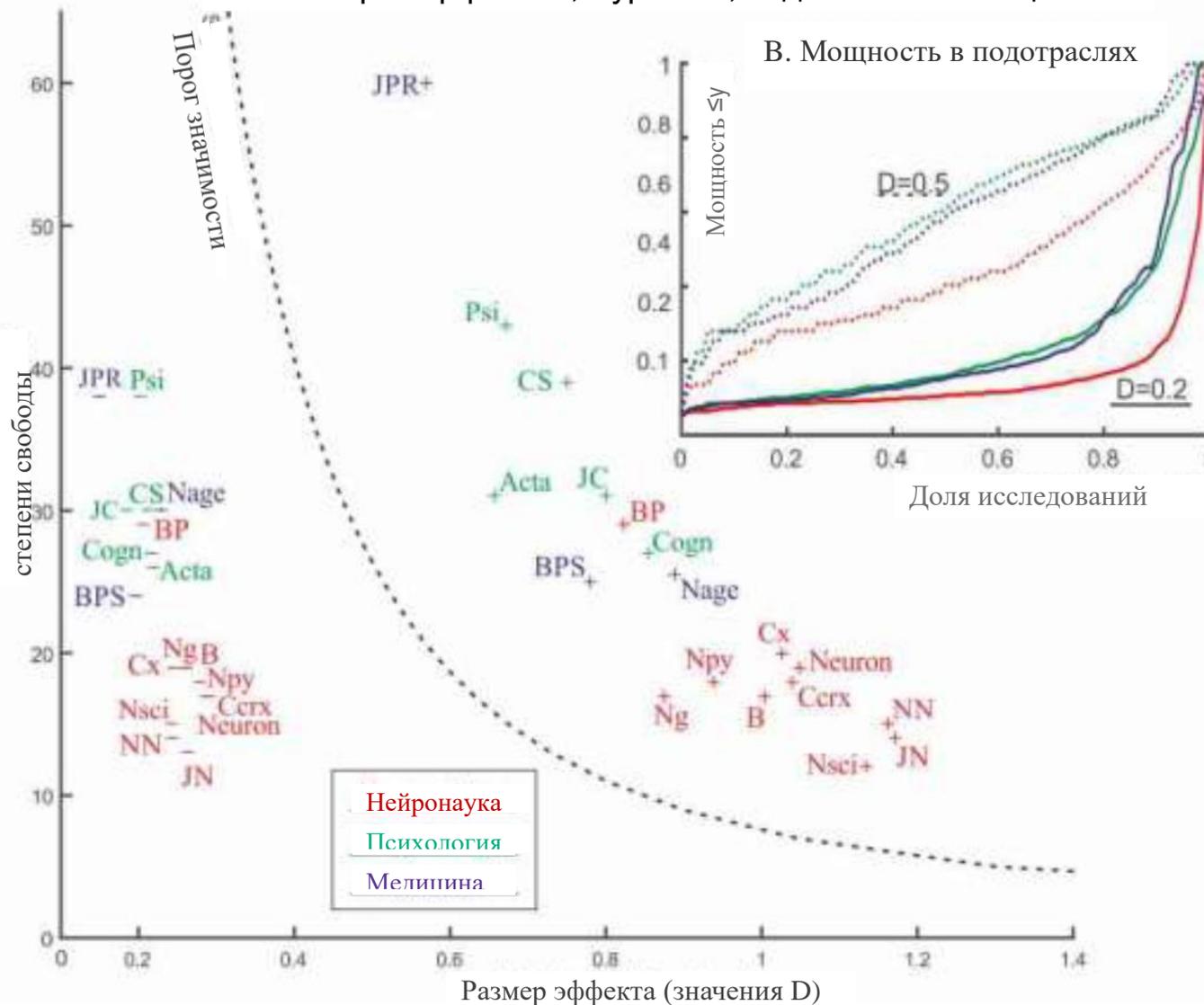
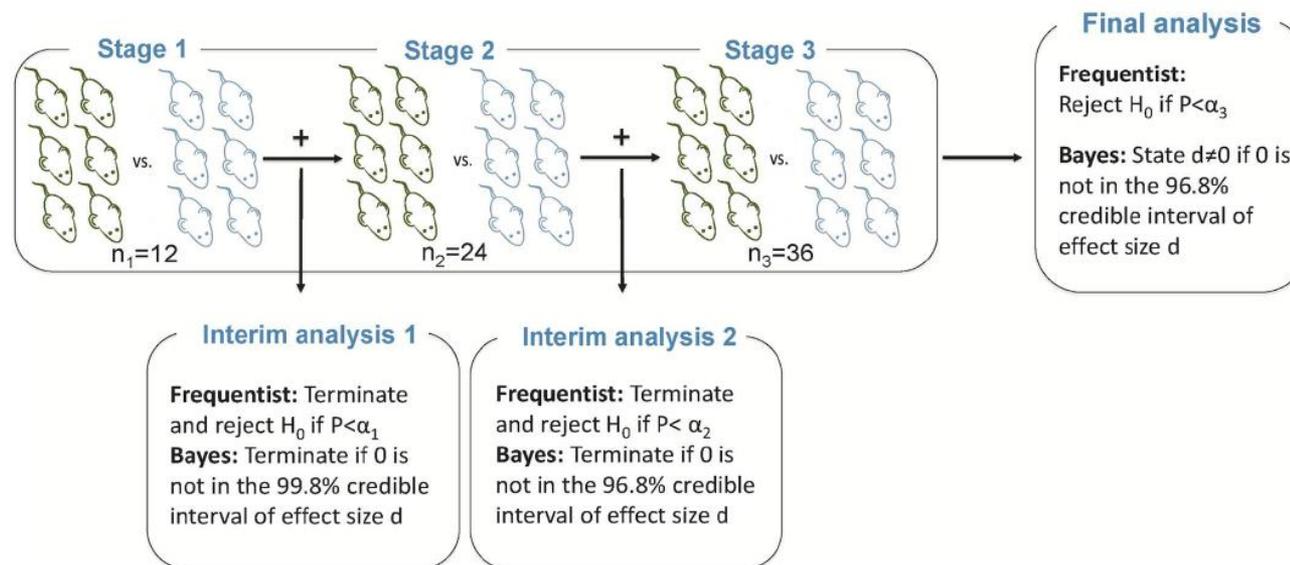


Figure 2. Power in journals and subfields. (A) Median effect sizes and degrees of

PERSPECTIVE

# Повышая эффективность доклинических испытаний посредством последовательных разработок групп

Konrad Neumann<sup>1</sup>, Ulrike Grittner<sup>1,2</sup>, Sophie K. Piper<sup>1,2,3</sup>, Andre Rex<sup>2,4</sup>, Oscar Florez-Vargas<sup>5</sup>, George Karystianis<sup>6</sup>, Alice Schneider<sup>1,2</sup>, Ian Wellwood<sup>2,7</sup>, Bob Siegerink<sup>2,8</sup>, John P. A. Ioannidis<sup>9</sup>, Jonathan Kimmelman<sup>10</sup>, Ulrich Dirnagl<sup>2,3,4,8,11,12</sup>



**Fig 1. Study design and sequential analysis approach allowing two interim analyses.** Stage 1: 33% of samples acquired, stage 2: 66% of samples acquired, and stage 3: 100% of samples acquired.  $H_0$ : null hypothesis,  $P$ :  $p$ -value, Credible interval: specific Bayesian interval of certainty about an estimate,  $d$ : effect size Cohen's  $d$ ,  $\alpha_i$ : significance levels for each stage derived from [11]  $\alpha_1 = 0.0006$ ,  $\alpha_2 = 0.0151$ ,  $\alpha_3 = 0.0471$ . Additionally, we used a Bayes factor approach (Table 1) and Pocock boundaries for the frequentist approach (S1 Table). All sequential approaches used were calibrated by using simulations to get a type I error of about 5%.

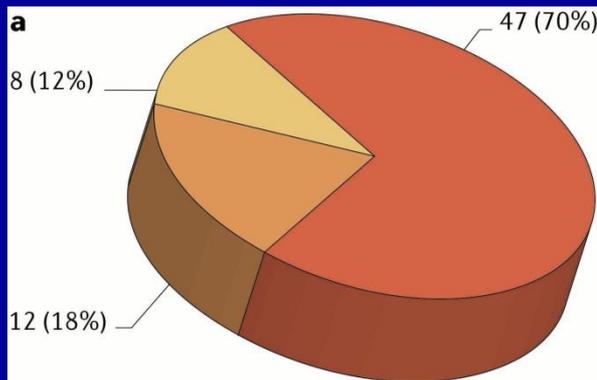
Эмпирические исследования в областях, где распространены практики воспроизводимости, показывают, что большинство статистических эффектов, изначально заявленных как значимые, оказываются ложноположительными или существенно преувеличенными

# Гены-кандидаты, воспроизведенные посредством полногеномного поиска ассоциаций (GWAS):

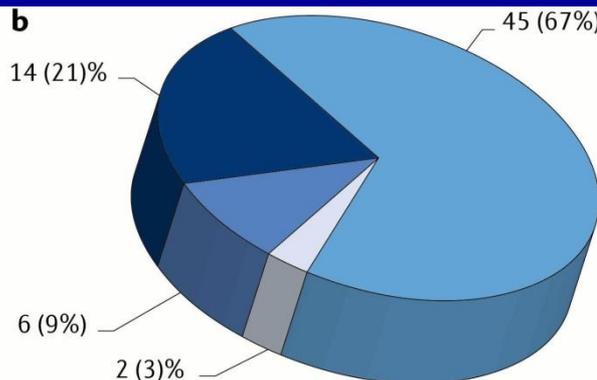
Таблица. Широкомасштабные усилия по массовому воспроизведению заявленных ассоциаций генов-кандидатов

Первый автор	Заболевание/фенотип	Протестированные локусы генов	Размер выборки (проектный)	Реплицированные локусы генов
Bosker (16)	Сильное депрессивное расстройство	57	3540 («Случай-контроль»)	1
Caropaso (17)	Курение (7 фенотипов)	359	4611 (Когортные)	1
Morgan (18)	Острый коронарный синдром	70	1461 («Случай-контроль»)	0
Richards (19)	Остеопороз (2 фенотипа)	150	19 195 (Когортные)	9
Samani (20)	Заболевание коронарных артерий	55	4864; 2519 («Случай-контроль»)	1
Scuteri (21)	Ожирение (3 фенотипа)	74	6148 (Когортные)	0
Soeber (22)	Кровяное давление	149	1644; 8023 (Когортные)	0
Wu (23)	Детская астма	237	1476 (Триады)	1

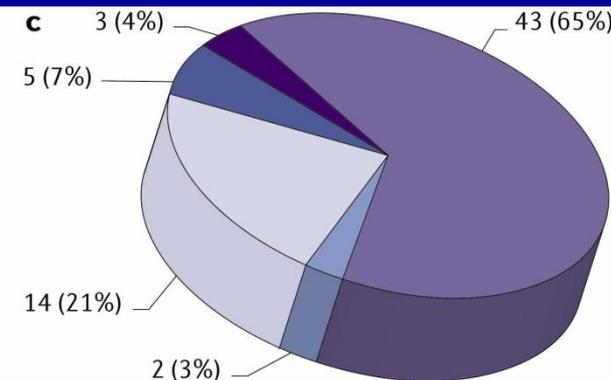
# Неудачное воспроизведение в доклинических исследованиях



- Онкология
- Женское здоровье
- Сердечно-сосудистые



- Модель адаптирована под внутренние нужды
- Данные из литературы перенесены на другие показания
- Не применимо
- Модель воспроизведена 1:1



- Непоследовательности
- Не применимо
- Данные из литературы совпадают с собственными
- Основной набор данных был воспроизводим
- Некоторые результаты были воспроизводимы

**d**

	Модель воспроизведена 1:1	Модель адаптирована под внутренние нужды (клеточные линии, анализ)	Данные из литературы перенесены на другие показания	Не применимо
Собственные данные в одном ряду с опубликованными результатами	1 (7%)	12 (86%)	0	1 (7%)
Непоследовательности, ведущие к прекращению проекта	11 (26%)	26 (60%)	2 (5%)	4 (9%)

Воспроизведено: только 6 из 53  
знаковых исследований в  
целевых проектах Amgen по  
онкологическим препаратам

- «Неудачу в «войне с раком» объясняли многими факторами, ... Но недавно появился новый виновник: слишком много базовых научных открытий... неверно».

Begley et al. Nature 2012

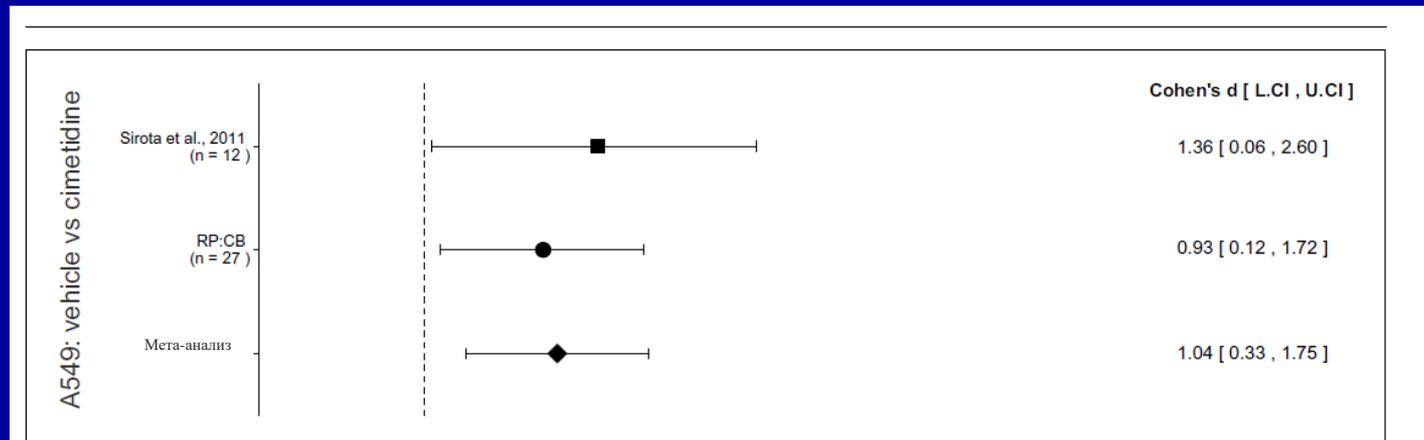
---

Войны воспроизводимости:  
Успешная, безуспешная, неинтерпретируемая, точная,  
концептуальная, триангулированная, оспариваемая  
воспроизводимость

John P.A. Ioannidis<sup>1,2,3,4\*</sup>

# Повторное исследование: обнаружение и доклиническая оценка показаний к применению препарата с использованием компендиума публичных данных по экспрессии генов

Irawati Kandela, Fraser Aird, Reproducibility Project: Cancer Biology\*



# Повторное исследование: комбинированное назначение проникающего в опухоль пептида повышает эффективность антираковых препаратов

Christine Mantis, Irawati Kandela, Fraser Aird,  
Reproducibility Project: Cancer Biology\*

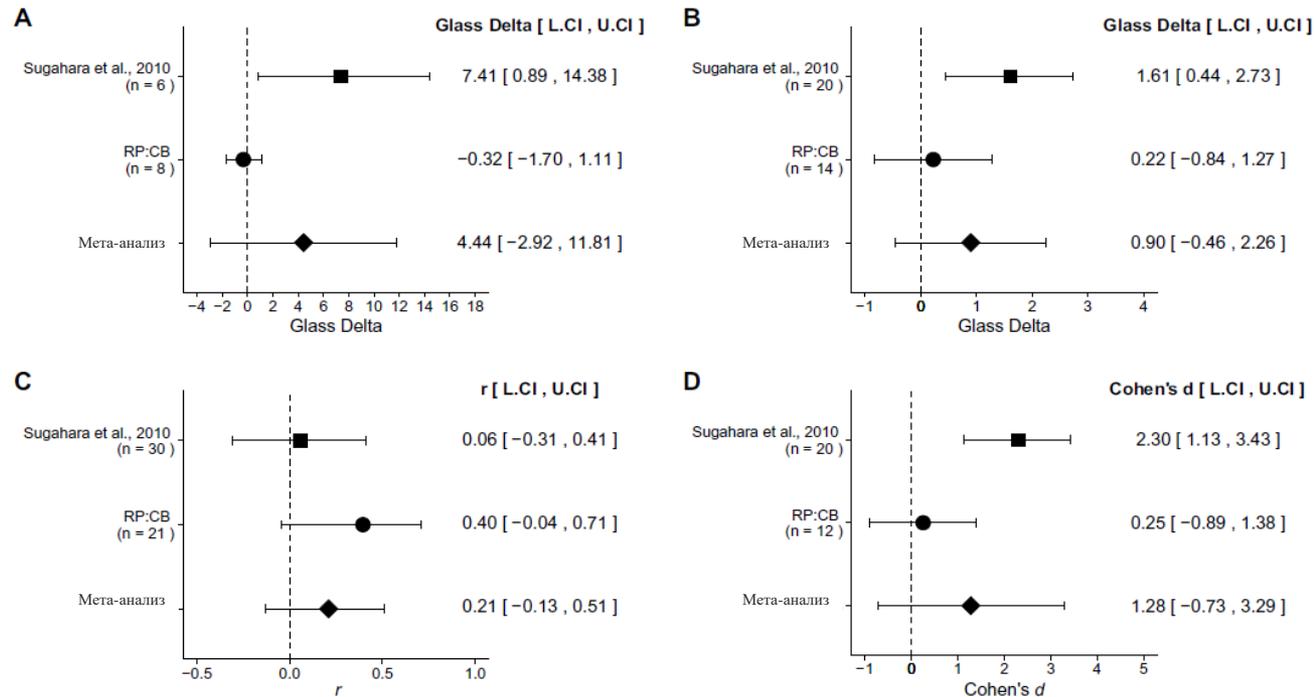


Таблица 1. Примеры некоторых отмеченных проблем с воспроизводимостью в доклинических исследованиях

Автор	Сфера	Отмеченные проблемы
Ioannidis et al (2009) <sup>22</sup>	Данные микрочипов	16/18 исследований в принципе невозможно воспроизвести на сырых данных
Baggerly et al (2009) <sup>23</sup>	Данные микрочипов	Многочисленные; недостаточные данные/плохая документация
Sena et al (2010) <sup>24</sup>	Исследования инсульта на животных	Явная предвзятость в публикации: только 2% исследований отрицательны
Prinz (2011) <sup>1</sup>	Общая биология	75-80% из 67 исследований не были повторены
Begley & Ellis (2012) <sup>2</sup>	Онкология	90% из 53 исследований не были повторены
Nekrutenko & Taylor(2012) <sup>25</sup>	Доступ к данным СНП	в 26/50 нет доступа к изначальным наборам данных/программному обеспечению
Perrin (2014) <sup>26</sup>	Мыши, на живом организме	0/100 представленных методов лечения не показали результата в исследованиях БАС
Tsilidis et al (2013) <sup>27</sup>	Неврологические исследования	Слишком много значимых результатов, явное выборочное представление отчетов
Lazic & Essioux (2013) <sup>28</sup>	Модель ВПК на мышах	Только в 3/34 использованы корректные экспериментальные измерения
Haibe-Kains et al (2013) <sup>29</sup>	Геномика/анализ клеточных линий	Прямое сравнение 15 препаратов и 471 клеточных линий из 2 групп, мало/нет сходных данных
Witwer (2013) <sup>30</sup>	Данные микрочипов	93/127 статей не соответствовали стандарту МИЭМ
Elliott et al (2006) <sup>31</sup>	Коммерческие антитела	Коммерческие антитоды обнаруживают неверные антигены
Prassas et al (2013) <sup>32</sup>	Коммерческий набор ELISA	Набор ELISA обнаружил неверный антиген
Stodden et al (2013) <sup>33</sup>	Журналы	Вычислительная биология: 105/170 журналов не соблюдают рекомендации Национальной Академии
Baker et al (2014) <sup>34</sup>	Журналы	Полное несоответствие согласованным стандартам исследований на животных
Vaux (2012) <sup>35</sup>	Журналы	Несоблюдение собственного статистического руководства

БАС означаем боковой амиотрофический склероз; МИЭМ – минимум информации об эксперименте на микрочипе; СНП – секвенирование нового поколения; и ВПК – вальпроевая кислота (модель аутизма)

**Таблица 2. Дополнительные основные научные отрасли, вызывающие беспокойство в отношении воспроизводимости**

Дисциплина	Поднятые вопросы	Автор
Нейронаука	Низкая статистическая мощность; маленький размер выборки	Button et al (2013) <sup>36</sup>
Фармакология	Недостаточное обучение и статистическая мощность, маскировка, гипотезы необходимые фармакокинетические исследования, рандомизация, эффект дозы, контроль, перспективная схема, подтверждение, независимое повторение, отбор недопустимых для человека доз	Henderson et al(2013) <sup>37</sup> ; Kenakin et al (2014) <sup>38</sup> ; McGonigie et al (2014) <sup>39</sup> ; Winquist et al (2014) <sup>40</sup> ; Marino (2014) <sup>41</sup>
Геномика/биоинформатика	Невоспроизводимость резонансных исследований	Sugden et al (2013) <sup>42</sup>
Биология стволовых клеток	Недостаток надежных, качественных данных	Plant & Parker (2013) <sup>43</sup>
Онкология, тесты в пробирке	Использование концентраций, недостижимых клинически	Smith & Houghton(2013) <sup>44</sup>
Химия, обнаружение ведущего вещества	Артефакты; ложноположительные и отрицательные	Davis & Erlanson (2013) <sup>45</sup>
Вычислительная биология	10 типичных ошибок	Sandve et al (2013) <sup>46</sup>
Патология/Биомаркеры	Качество биообразцов	Simeon-Dubach et al (2012) <sup>47</sup>
Организационная психология	Подавление негативных исследований	Kepes & McDonald (2103) <sup>48</sup>
Наблюдательное исследование	0/52 гипотез подтверждено в рандомизированных испытаниях	Young & Karr (2011) <sup>41</sup>

# В некоторых странах – еще хуже (PNAS, 2013)

Исследования в США могут переоценивать размер эффекта в более мягких исследованиях

Daniele Fanelli<sup>a,1</sup> and John P. A. Ioannidis<sup>b,c,d,e</sup>

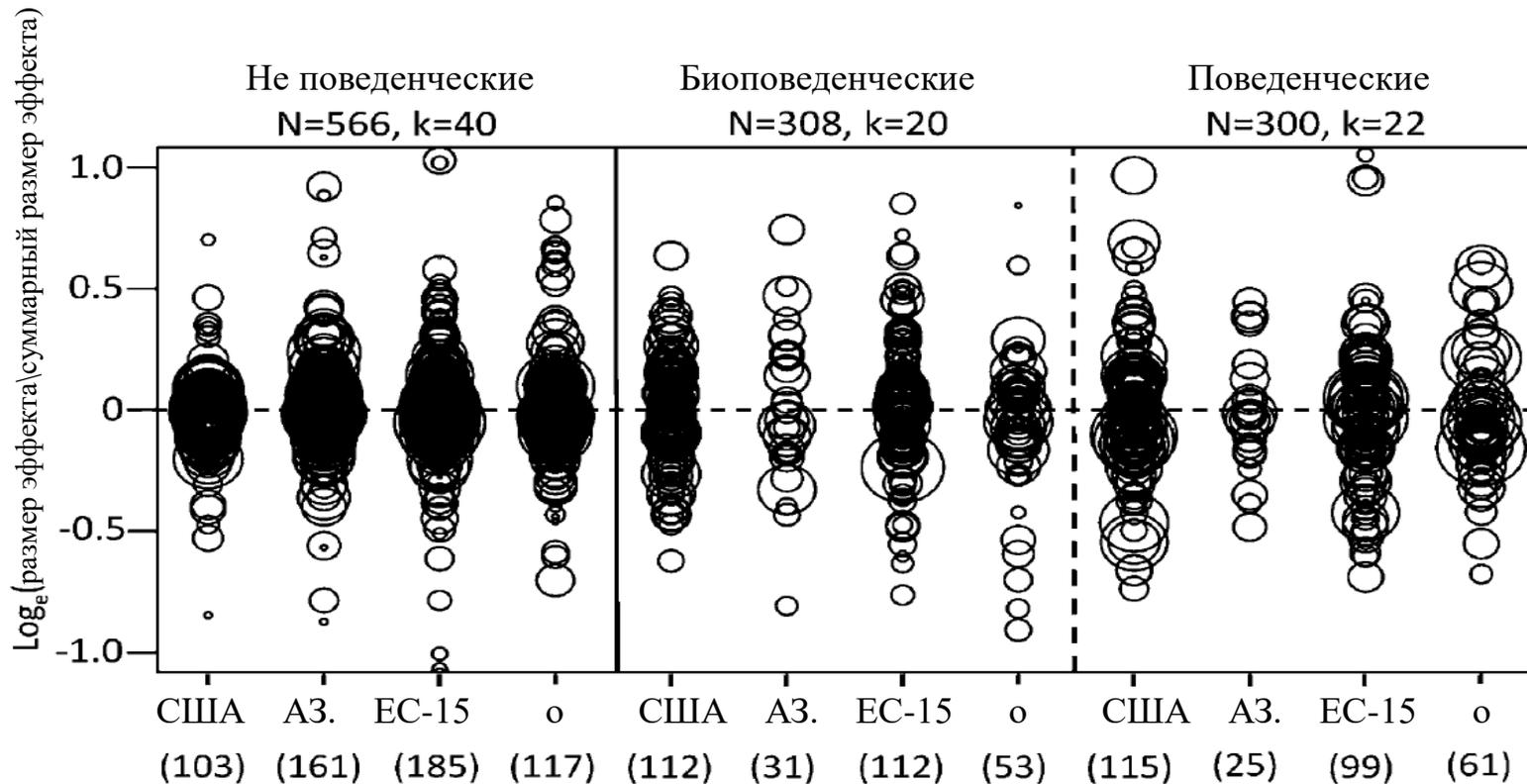


Рис. 1. Величина размера эффекта первичных исследований относительно суммарного размера эффекта при ретроспективном мета-анализе, разделенные по географическому признаку

# Большие данные, Большой шум, Большие ошибки

MEDICINE

## Большие данные сталкиваются с общественным здравоохранением

Human well-being could benefit from large-scale data if large-scale noise is minimized

By **Muin J. Khoury**<sup>1</sup> and  
**John P. A. Ioannidis**<sup>2</sup>

**N**1854, as cholera swept through London, John Snow, the father of modern epidemiology, painstakingly recorded the locations of affected homes. After long, laborious work, he implicated the Broad Street water pump as the source of the outbreak, even without knowing that a *Vibrio* organism caused cholera. Today, Snow might have crunched Global Positioning System information and disease prevalence data, solving the problem within hours (1). That is the potential impact of “Big Data” on the public’s health. But the promise of Big Data is also accompanied by claims that “the scientific method itself is becoming obsolete” (2), as next generation computers, such as IBM’s Watson (3), sift through the digital world to provide predictive models based on massive information. Separating the true signal from the gigantic amount of noise is neither easy nor straightforward, but it is a challenge that must be tackled if information is ever to be translated into societal well being.

The term “Big Data” refers to volumes of large, complex, linkable information (4). Beyond genomics and other “omic” fields, Big Data includes medical, environmental, financial, geographic, and social media information. Most of this digital information was unavailable a decade ago. This swell of data will continue to grow, stoked by sources that are currently unimaginable. Big Data stands to improve health by providing insights into



**From validity to utility.** Big Data can improve tracking and response to infectious disease outbreaks, discovery of early warning signals of disease, and development of diagnostic tests and therapeutics.

For non-genomic associations, false alarms due to confounding variables or other biases are possible even with very large-scale studies, extensive replication, and very strong signals (9). Big Data’s strength is in finding associations, not in showing whether these associations have meaning. Finding a signal is only the first step.

Even John Snow needed to start with a plausible hypothesis to know where to look, i.e., choose what data to examine. If all he had was massive amounts of data, he might well have ended up with a correlation as spurious as the honey bee-marijuana connection. Crucially, Snow “did the experiment.” He removed the handle from the water pump and dramatically reduced the spread of cholera, thus moving from correlation to causation and effective intervention.

How can we improve the potential for Big Data to improve health and prevent disease? One priority is that a stronger epidemiological foundation is needed. Big Data analysis is currently largely based on convenient samples of people or information available on the Internet. When associations are probed between perfectly measured data (e.g., a genome sequence) and poorly measured data (e.g., administrative claims health data), research accuracy is dictated by the weakest link. Big Data are observational in nature and are fraught with many biases such as selection, confounding variables, and lack of generalizability. Big Data analysis may be embedded in epidemiologically well-characterized and representative populations. This epide-



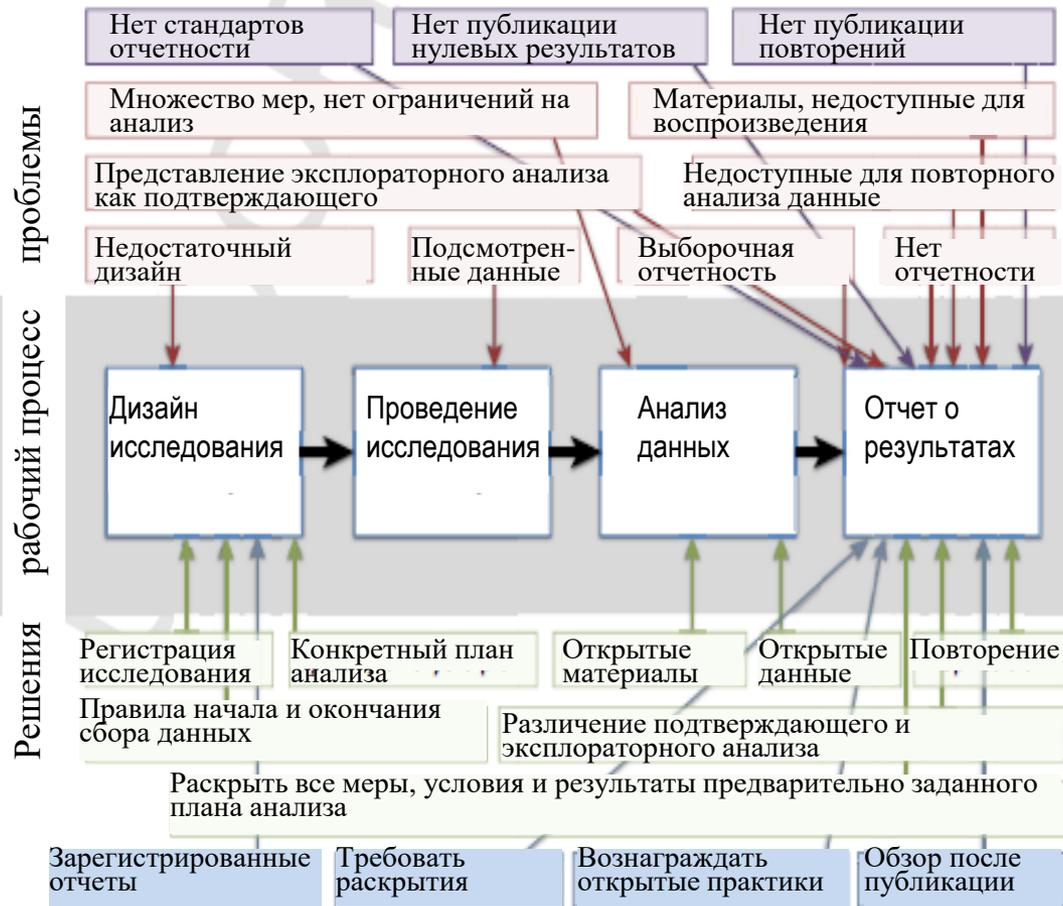
VIEWPOINT

## **Невидимые исследования**

Биомедицинские инновации происходят за пределами рецензируемой литературы?

# Потенциальные решения для улучшения исследовательских практик

- Некоторые решения уже сработали в отдельных отраслях, и, возможно, стоит рассмотреть их применение в других
- Другие решения более умозрительны
- Необходимы эмпирические данные об их эффективности
- Решения, кажущиеся эффективными, могут также нести сопутствующий ущерб
- Не навреди

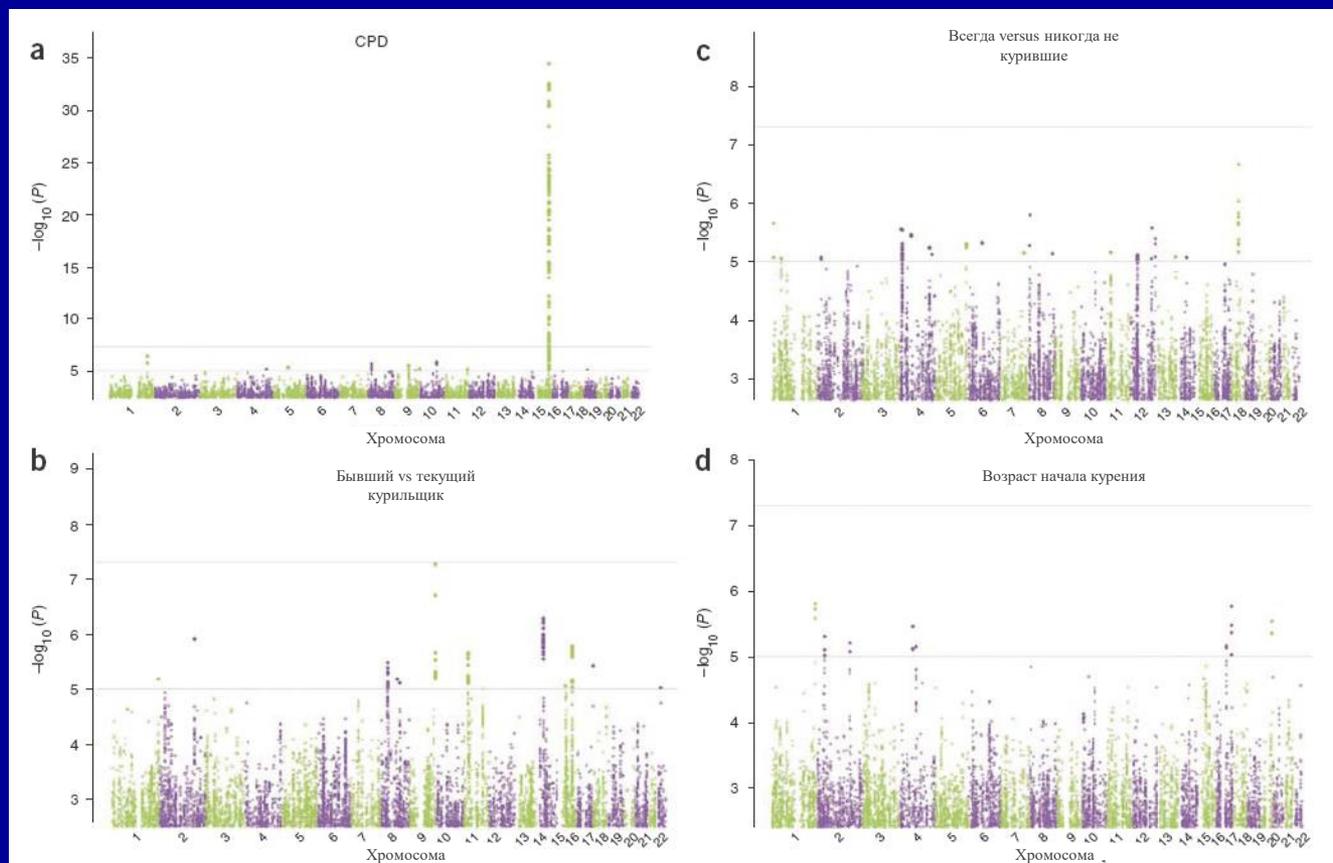


TRENDS in Cognitive Sciences

## **Вставка 1. Некоторые исследовательские практики, которые могут помочь в увеличении доли действительно достоверных результатов**

- Широкомасштабные совместные исследования
- Усвоение культуры проведения повторных испытаний
- Учет (исследований, протоколов, кодов анализа, наборов данных, сырых данных и результатов)
- Совместное использование (данных, протоколов, материалов, программного обеспечения и других инструментов)
- Методики воспроизводимости
- Политика сдерживания в отношении противоречивых спонсоров и авторов
- Более подходящие статистические методы
- Стандартизация определений и процедур анализа
- Более точные пороговые значения для претендентов на звание открытия или "успеха"
- Улучшения в стандартах разработки исследования
- Улучшение процедур оценки коллегами, отчетности и распространения исследования
- Улучшение обучения научных сотрудников в отношении методов и статистической грамотности

# Широкомасштабное сотрудничество и принятие культуры проведения повторных испытаний



# Воспроизведение – кем?

- Теми же исследователями
- Разными исследователями одной школы
- Разными исследователями, представляющими соперничающие теории/гипотезы
- Скомбинировать вышеперечисленных
- Открыть широкой публике

# Уровни регистрации

- Уровень 0: без учета
- Уровень 1: регистрация набора данных
- Уровень 2: регистрация протокола
- Уровень 3: регистрация плана проведения анализа
- Уровень 4: регистрация плана проведения анализа и сырых данных
- Уровень 5: открытая прямая трансляция



# Зарегистрированный отчет: систематическая идентификация геномных маркеров чувствительности к препарату в раковых клетках

John P Vanden Heuvel<sup>1,2</sup>, Jessica Bullenkamp<sup>3</sup>,  
Reproducibility Project: Cancer Biology\*

<sup>1</sup>Indigo Biosciences, State College, United States; <sup>2</sup>Veterinary and Biomedical Sciences, Penn State University, University Park, PA, United States; <sup>3</sup>King's College London, London, United Kingdom

## REPRODUCIBILITY PROJECT CANCER BIOLOGY

**Abstract** The Reproducibility Project: Cancer Biology seeks to address growing concerns about the reproducibility in scientific research by conducting replications of selected experiments from a number of high-profile papers in the field of cancer biology. The papers, which were published between 2010 and 2012, were selected on the basis of citations and Altmetric scores (*Errington et al., 2014*). This Registered Report describes the proposed replication plan of key experiments from "Systematic identification of genomic markers of drug sensitivity in cancer cells" by Garnett and colleagues, published in *Nature* in 2012 (*Garnett et al., 2012*). The experiments to be replicated are those reported in Figures 4C, 4E, 4F, and Supplemental Figures 16 and 20. Garnett and colleagues performed a high throughput screen assessing the effect of 130 drugs on 639 cancer-derived cell lines in order to identify novel interactions for possible therapeutic approaches. They then tested this approach by exploring in more detail a novel interaction they identified in which Ewing's sarcoma cell lines showed an increased sensitivity to PARP inhibitors (Figure 4C). Mesenchymal progenitor cells (MPCs) transformed with the signature *EWS-FLI1* translocation, the hallmark of Ewing's sarcoma family tumors, exhibited increased sensitivity to the PARP inhibitor olaparib as compared to MPCs transformed with a different translocation (Figure 4E). Knockdown mediated by siRNA of *EWS-FLI1* abrogated this sensitivity to olaparib (Figure 4F). The Reproducibility Project: Cancer Biology is a collaboration between the Center for Open Science and Science Exchange, and the results of the replications will be published by *eLife*.

DOI: 10.7554/eLife.13620.001

\*For correspondence: [tim@cos.io](mailto:tim@cos.io)

Group author details:  
Reproducibility Project: Cancer  
Biology [See page 18](#)

Competing interest: [See page 18](#)



# Совместное использование данных— кем, когда и как?

Doshi, Goodman,  
Ioannidis, TiPS  
2013

## Таблица 1. Обсуждаемые вопросы по оптимальному порядку совместного использования данных в клинических испытаниях

### Организация, делящаяся данными и/или устанавливающая правила

- Регулирующие органы (Управление по санитарному надзору за качеством пищевых продуктов и медикаментов (FDA), Европейское агентство лекарственных средств (EMA))
- Спонсор(ы) каждого испытания
- Исследователи, проводящие каждое испытание
- Главная организация, представляющая спонсоров (напр., PhRMA– EFPIA)
- Другая/новая организация (будет создана, возможно, с участием некоторых/всех вышеперечисленных)

### Доступность данных

- Все собранные сырые данные как они есть
- Обработанные версии данных (напр., очищенные и/или деидентифицированные)  
В этом случае, кто будет проводить обработку и с какими ресурсами?
- Ограниченный доступ (т.е. отсутствие совместного использования, но предоставление уполномоченным пользователям доступа на запрос, но не на загрузку данных)
- Ограниченные версии (напр., классифицированные по конкретным проектным запросам)

### Имеющие право на запрос данных

- Кто угодно, например, открытый общий доступ
- Только запрашивающие с особыми полномочиями  
В этом случае, какими?

### Критерии одобрения

- Никаких критериев, например, неограниченный общественный доступ
- Минимальные критерии по контракту
- Обзор предложений  
В этом случае, кем, кто будет назначать комиссию рецензентов и какими для них будут критерии приемлемости (напр., конфликт интересов, содержательная экспертиза)?

### Срок доступности данных для совместного использования

- Непосредственно по завершении исследования
- Публикация основного анализа
- С некоторым временным промежутком (напр., 6 месяцев или 1 год), чтобы предоставить первичной команде руководство по осуществлению дополнительных анализов
- Как-то связано с лицензионным циклом (для лицензионных продуктов)
- Особые издания с доступом к данным прошлых испытаний
- Доступность архивированной информации
- Предварительные правовые, контрактные и разрешающие ограничения

### Принуждение или стимулы

- Законодательные обязательства по совместному использованию данных
- Стимулы, связанные с совместным использованием данных (или антистимулы за его отсутствие)
- Исследователям:
  - От журналов (напр., раскрытие данных – предпосылка для публикации)
  - От спонсоров (напр., финансирование исследователей не относящимися к компании спонсорами в зависимости от прошлых практик совместного использования)
  - Компаниям (напр., лицензирование или патенты в связи с совместным использованием данных)

### Дальнейшее совместное использование данных

- Неограниченное (напр., открытый общественный доступ)
- Доступно одобренным (упомянутым выше) (по контракту)

# Повторяемость

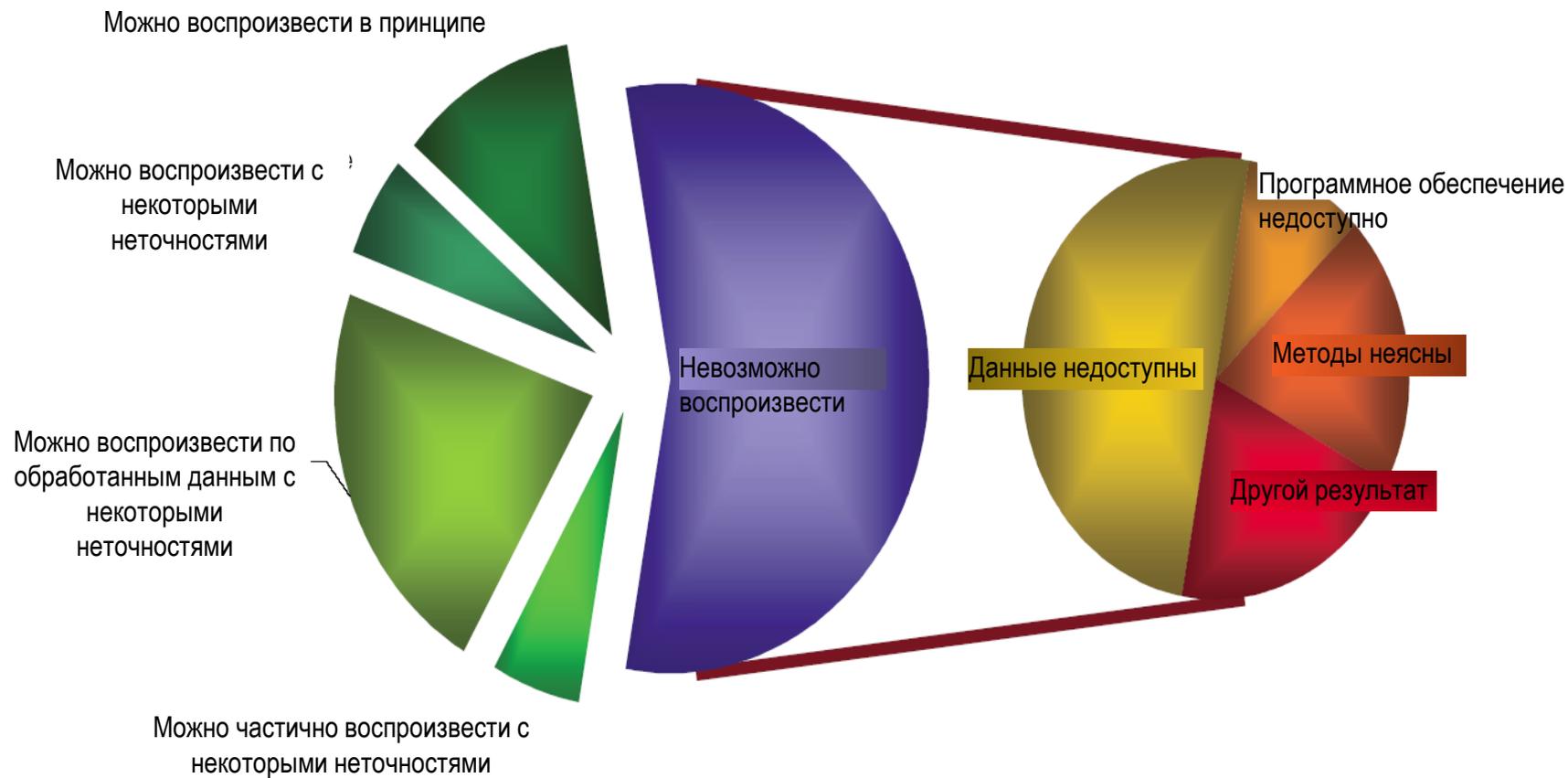
ANALYSIS

nature  
genetics

---

## Повторяемость опубликованных анализов экспрессии генов с помощью микрочипов

John P A Ioannidis<sup>1-3</sup>, David B Allison<sup>4</sup>, Catherine A Ball<sup>5</sup>, Issa Coulibaly<sup>4</sup>, Xiangqin Cui<sup>4</sup>, Aedín C Culhane<sup>6,7</sup>, Mario Falchi<sup>8,9</sup>, Cesare Furlanello<sup>10</sup>, Laurence Game<sup>11</sup>, Giuseppe Jurman<sup>10</sup>, Jon Mangion<sup>11</sup>, Tapan Mehta<sup>4</sup>, Michael Nitzberg<sup>5</sup>, Grier P Page<sup>4,12</sup>, Enrico Petretto<sup>11,13</sup> & Vera van Noort<sup>14</sup>



**Рисунок 1** Резюме усилий по воспроизведению опубликованных анализов.

REPRODUCIBILITY

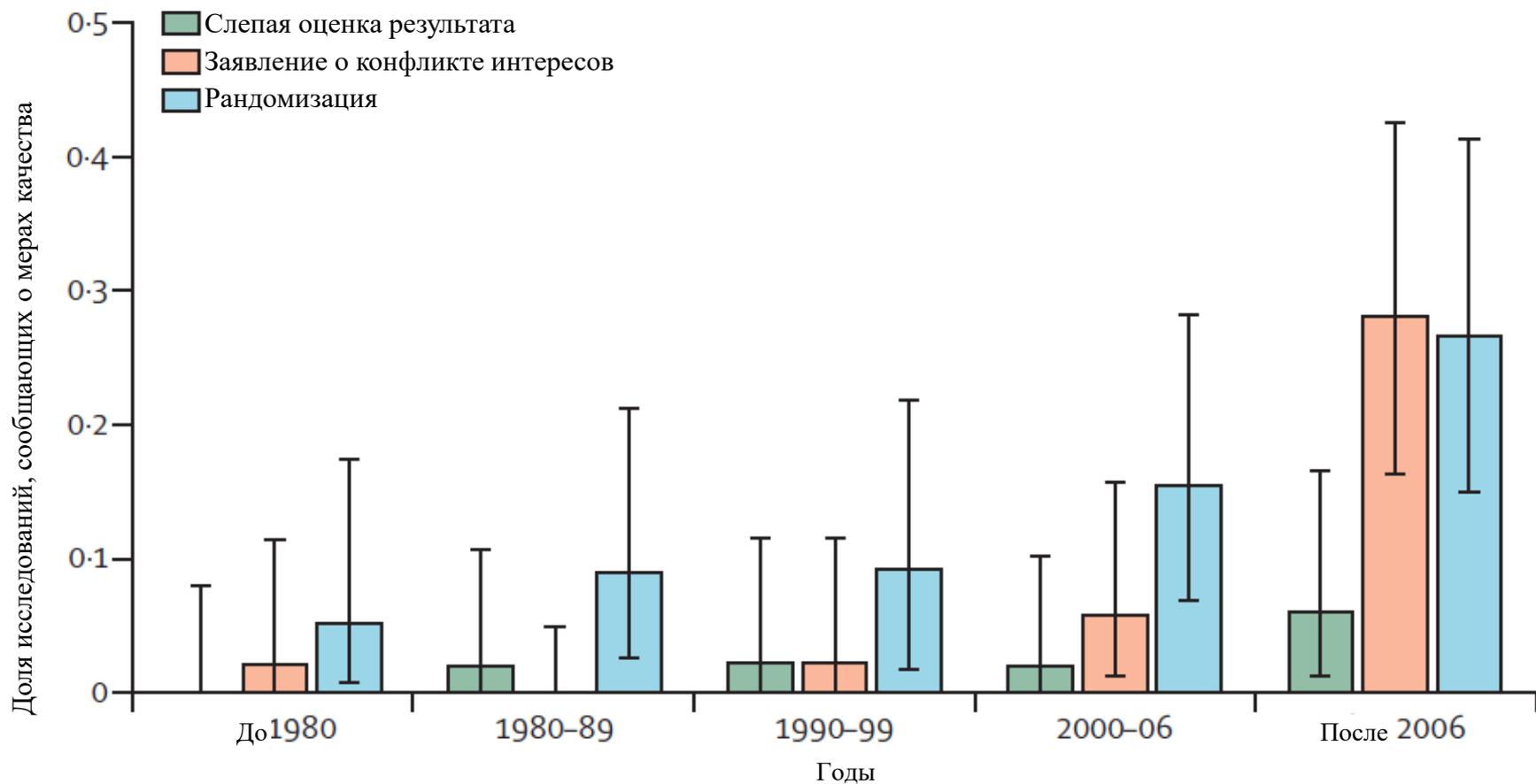
# Повышение воспроизводимости вычислительными методами

Data, code and workflows should be available and cited.

*By Victoria Stodden, Marcia McNutt, David H. Bailey, Ewa Deelman, Yolanda Gil, Brooks Hanson, Michael A. Heroux, John P.A. Ioannidis, Michela Taufer*

# Улучшенные статистики и методы

- Прозрачные (зарегистрированные?) планы статистического анализа
- Обучение научных работников статистике и повышение их уровня грамотности/умения работать с числами
- Улучшенная разработка исследований
- Стандартные черты: напр., рандомизация и маскировка исследователей в экспериментах над животными
- Какова роль перечней разработки/проведения?



**Рисунок:** Тренды в отношении трех методологических индикаторов качества в отчетах об исследованиях на живых организмах. Мы случайным образом отобрали 2000 записей из PubMed (опубликованных в 1960-2012 гг.) на основе ID PubMed (детали и набор данных исследования см. в приложении). 254 отчета описывали эксперименты на живых организмах, вне организма или в пробирке, проводимые на животных, не являющихся человеком. Два исследователя независимо определяли, имела ли место слепая оценка результата, заявление о конфликте интересов или рандомизация. Отчет о соотношениях, включающий эту информацию

Таблица 4. Вопросы, которые можно решить с помощью политики достойной институциональной практики базовых исследований

Ориентация на	Предложение
Студентов/постдокторантов	<p>Базовое обучение экспериментальным методам и дизайну; отбору данных; анализу данных; маскировке; включению контрольных; интерпретации статистики; оценке реагентов; воспроизведению и повторению экспериментов</p> <p>Старший коллега из независимого отдела в роли наставника</p>
Исследователь	<p>Требование, чтобы субъективные конечные пункты оценивал «слепой» исследователь</p> <p>Обязательные курсы повышения квалификации по экспериментальному дизайну; отбору данных; включению контрольных анализу данных; оценке реагентов; вопросам развивающихся технологий</p> <p>Требование соблюдать указания и рекомендации Федерального и Научного сообщества</p>
Институт	<p>Указания по борьбе с мошенничеством</p> <p>Независимый комитет для обзора соответствия</p> <p>Требование, чтобы доступ к сырым данным открывался по запросу</p> <p>Указания по ведению лабораторных записей</p> <p>Случайные проверки лабораторных записей</p> <p>Прозрачная процедура повышений, при которой качество ценится выше показных, невозпроизводимых исследований, наставничество и обучение – вознаграждаются</p>

# Улучшая стандарты учета исследований: EQUATOR



COMMUNITY PAGE

# Мета-исследования: оценка и усовершенствование исследовательских методов и практик

**John P. A. Ioannidis\*, Daniele Fanelli, Debbie Drake Dunne, Steven N. Goodman**

Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California, United States of America

\* [jiannid@stanford.edu](mailto:jiannid@stanford.edu)

## Таблица 1. Главные темы мета-исследований

### Область мета-исследований

### Особый интерес (не исчерпывающий список)

**Методы:** «проведение исследований» – разработка исследований, методы, статистики, синтез исследований, сотрудничество и этика

Предвзятости и сомнительные практики проведения исследований, методы уменьшения таких предвзятостей, тема-анализ, синтез исследований, объединение фактических данных, синтез пересекающихся разработок, совместная работа научных коллективов и консорциумов, честность и этика исследований

**Отчет:** «передача исследований» – стандарты отчетности, регистрации исследований, раскрытия конфликта интересов, информация для пациентов, общественности и лиц, определяющих политику

Предвзятости и сомнительные практики отчетности, объяснения, распространения и популяризации исследований, раскрытие и управление конфликтами интересов, регистрация исследований и другие меры предотвращения появления предвзятости и методы отслеживания и уменьшения таких проблем

**Воспроизводимость:** «проверка исследований» – совместное использование данных и методов, повторяемость, сходимость, воспроизводимость и само-исправление

Препятствия к совместному использованию данных и методов, повтор исследований, повторяемость и воспроизводимость опубликованных исследований, методы их улучшения, эффективность исправления и само-исправления в литературе и методы улучшения

**Оценка:** «оценка исследований» – рецензирование до публикации и после нее, критерии финансирования исследований и другие способы оценки научного качества

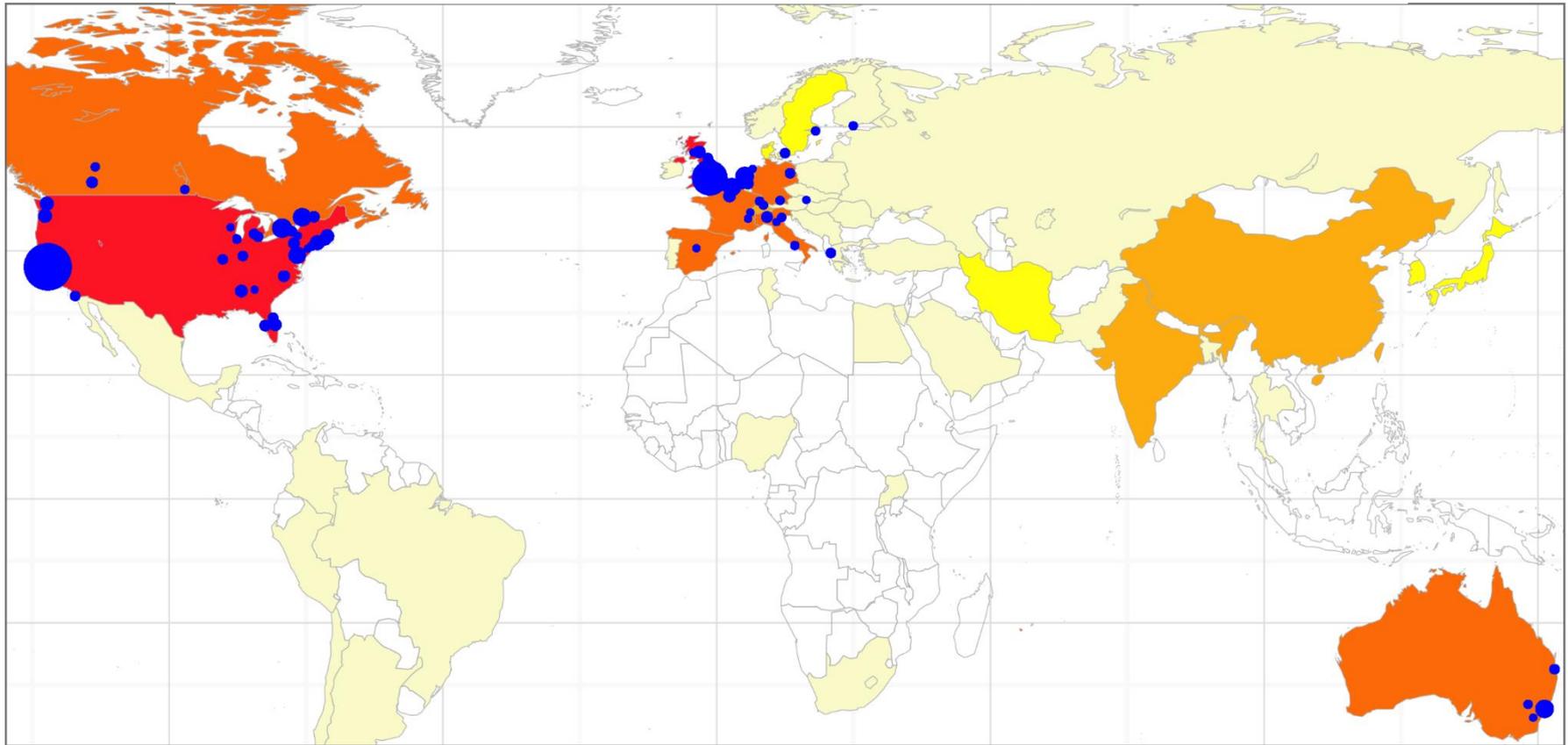
Точность, эффективность, расходы и выгоды от старых и новых подходов к рецензированию и другие методы оценки науки и методы их улучшения

**Стимулы:** «награда исследований»: критерии получения повышений, наград и взысканий при оценке исследований для отдельных лиц, групп и институтов

Точность, эффективность, расходы и выгоды от старых и новых подходов к ранжированию и оценке производительности, качества, ценности и исследований, отдельных лиц, групп и институтов

[doi:10.1371/journal.pbio.1002264.t001](https://doi.org/10.1371/journal.pbio.1002264.t001)

современная мета-исследовательская литература (общ. N=815)



**Таблица 2. Не исчерпывающий список инициатив, касающихся различных тем мета-исследований**

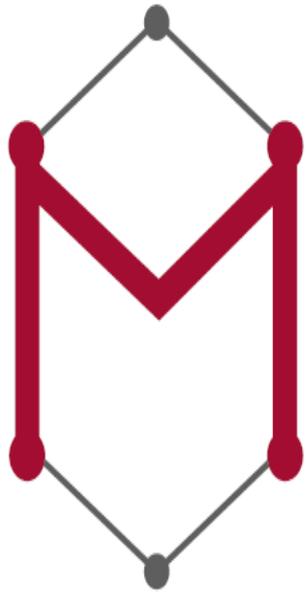
Инициатива	Область работы (веб-сайт)
<b><u>МЕТОДЫ</u></b>	
Cochrane Collaboration	Систематические обзоры области здравоохранения- ( <a href="http://cochrane.org">cochrane.org</a> )
Campbell Collaboration	Систематические обзоры социальных наук ( <a href="http://campbellcollaboration.org">campbellcollaboration.org</a> )
James Lind Library	Эволюция честного тестирования методов лечения ( <a href="http://jameslindlibrary.org">jameslindlibrary.org</a> )
Society for Clinical Trials	Клинические испытания ( <a href="http://sctweb.org">sctweb.org</a> )
SRSM	Методы синтеза исследований ( <a href="http://srsm.org">srsm.org</a> )
BioSharing	Стандарты для биологических, естественных наук ( <a href="http://biosharing.org">biosharing.org</a> )
Human Proteome Project	Центр сотрудничества по протеомам ( <a href="http://thehpp.org">thehpp.org</a> )
NCPRE	Этика исследований ( <a href="http://ethicscenter.csl.illinois.edu">ethicscenter.csl.illinois.edu</a> )
<b><u>ОТЧЕТНОСТЬ</u></b>	
<a href="http://ClinicalTrials.gov">ClinicalTrials.gov</a>	Учет клинических испытаний ( <a href="http://clinicaltrials.gov">clinicaltrials.gov</a> )
EQUATOR network	Стандарты отчетов об исследованиях ( <a href="http://equator-network.org">equator-network.org</a> )
Sense About Science	Информирование общественности об исследованиях ( <a href="http://senseaboutscience.org">senseaboutscience.org</a> )
Health News Reviews	Экспертный обзор новостей науки ( <a href="http://healthnewsreview.org">healthnewsreview.org</a> )
<b><u>ВОСПРОИЗВОДИМОСТЬ</u></b>	
Center for Open Science	Открытая наука, психология и не только; ( <a href="http://centerforopenscience.org">centerforopenscience.org</a> )
BITSS	Прозрачность в социальных науках ( <a href="http://bitss.org">bitss.org</a> )
BPS	Лучшие практики в социальных науках ( <a href="http://bps.stanford.edu">bps.stanford.edu</a> )
Political Science Replication	Воспроизводимость в политических науках ( <a href="http://politicalsciencereplication.com">politicalsciencereplication.com</a> )
YODA	Раскрытие данных клинических исследований ( <a href="http://yoda.yale.edu">yoda.yale.edu</a> )
Neurovault	Хранилище карт ПЭТ и МРТ ( <a href="http://neurovault.org">neurovault.org</a> )
OpenfMRI	Хранилище фМРТ-данных ( <a href="http://openfmri.org">openfmri.org</a> )
Хранилища НИИ, примеры:	
dbGAP	Сырые данные по генотипам и фенотипам; ( <a href="http://ncbi.nlm.nih.gov/gap">ncbi.nlm.nih.gov/gap</a> )
GEO	Хранилище функциональной геномики, ( <a href="http://ncbi.nlm.nih.gov/geo">ncbi.nlm.nih.gov/geo</a> )
Science Exchange	Проверки воспроизводимости ( <a href="http://validation.scienceexchange.com">validation.scienceexchange.com</a> )

## ОЦЕНКА

Peer Review Congress	Фактические данные по рецензированию ( <a href="http://peerreviewcongress.org">peerreviewcongress.org</a> )
Center for Scientific Integrity	Отслеживание отказов от научных статей ( <a href="http://retractionwatch.com/the-center-for-scientific-integrity">retractionwatch.com/the-center-for-scientific-integrity</a> )
PubMed Commons	Комментарии после публикации ( <a href="http://ncbi.nlm.nih.gov/pubmedcommons">ncbi.nlm.nih.gov/pubmedcommons</a> )
ArXiv	Хранилище статей до печати ( <a href="http://arxiv.org">arxiv.org</a> )
ICMJE	Стандарты журнальных публикаций ( <a href="http://icmje.org">icmje.org</a> )
COPE	Этика журнальных публикаций ( <a href="http://publicationethics.org">publicationethics.org</a> )
PubPeer	Комментарии исследования от коллег ( <a href="http://pubpeer.com">pubpeer.com</a> )
PEERE	Новые модели рецензирования ( <a href="http://www.peere.org">www.peere.org</a> )

## СТИМУЛЫ

REWARD	Сокращение негодных и вознаграждение старательных исследований ( <a href="http://researchwaste.net">researchwaste.net</a> )
AAAS	Научная политика ( <a href="http://aaas.org">aaas.org</a> )
ICSU	Международная научная политика ( <a href="http://icsu.org">icsu.org</a> )



# METRICS

META•RESEARCH INNOVATION  
CENTER AT STANFORD

[metaresearch.stanford.edu](http://metaresearch.stanford.edu)

# Перестроение системы вознаграждений

Таблица. Индекс ПКВРВ (PQRST) для оценки и вознаграждения исследований

Ввод в действие		
Пункт индекса ПКВРВ	Пример	Источник данных
П (Продуктивность)	Число публикаций среди % наиболее цитируемых для данной научной области и года	ISI Essential Science Indicators (автоматизированный)
	Доля предложений по финансированию, приведших к ≥1 опубликованному отчету об основных результатах	Записи финансирующих органов и автоматические записи признанных грантов (напр., PubMed)
К (качество научной работы)	Доля зарегистрированных протоколов, опубликованных через 2 года после завершения исследований:	Реестры исследований, такие, как ClinicalTrials.gov для испытаний
	Доля публикаций, удовлетворяющих ≥1 стандарту качества	
В (воспроизводимость научной работы)	Доля воспроизводимых публикаций	Сейчас нет автоматизированной базы данных с широким охватом, но может быть легко создана, особенно если учитывать только основные, самые цитируемые статьи в каждой области
Р (раскрытие данных или других источников)	Доля публикаций, раскрывающих данные, материалы и/или протоколы (любые из актуальных пунктов)	Сейчас нет автоматизированной базы данных с широким охватом, но можно легко создать, напр., встроить в PubMed при создании записи и обновлять ее, если потом раскрывается больше
В (эффект внедрения исследования в практику)	Доля публикаций, приведших к преодолению «внедренческого» рубежа, напр., получив многообещающие результаты при испытаниях на людях для вмешательств, протестированных на животных или клеточных культурах, или лицензирование мер для клинических испытаний	Сейчас нет автоматизированной базы данных с широким охватом, должна была бы курироваться оценщиком (напр., финансирующим органом), и, возможно, пришлось бы ограничиться самыми цитируемыми статьями, если получалось бы громоздко

# Манифест о воспроизводимой науке

Marcus R. Munafò<sup>1,2\*</sup>, Brian A. Nosek<sup>3,4</sup>, Dorothy V. M. Bishop<sup>5</sup>, Katherine S. Button<sup>6</sup>,  
Christopher D. Chambers<sup>7</sup>, Nathalie Percie du Sert<sup>8</sup>, Uri Simonsohn<sup>9</sup>, Eric-Jan Wagenmakers<sup>10</sup>,  
Jennifer J. Ware<sup>11</sup> and John P. A. Ioannidis<sup>12,13,14</sup>



**Рисунок 1** | Угрозы, направленные на воспроизводимую науку. Демонстрируется идеализированная версия гипотетико-дедуктивной модели научного метода. Разл

**Таблица 1 | Манифест о воспроизводимой науке.**

Тема	Предложение	Примеры инициатив/возможных решений (степень внедрения на данный момент)	Заинтересованные стороны
Методы	Защита от когнитивных искажений	Все из нижеперечисленных инициатив (от * до *****) Ослепление (**)	Ж, С
	Улучшение методологического обучения	Строгое обучение статистике и методам исследования будущих исследователей (*) Интенсивное продолжение обучения статистике и методам исследования (*)	И, С
	Независимая методологическая поддержка	Причастность методологов к исследованию (**) Независимый надзор (*)	С
	Сотрудничество и командная наука	Многоплощадочные исследования/сбор распространенных данных (*) Консорциумы командной науки (*)	И, С
Обмен и распространение	Продвижение предварительной регистрации исследований	Зарегистрированная отчетность (*) Структура открытой науки (*)	Ж, С
	Улучшение качества отчетности	Использование отчетных перечней (**) Протокольные перечни (*)	Ж
	Защита от конфликтов интересов Исключение/предотвращение финансовых и нефинансовых конфликтов интересов	Раскрытие конфликтов интересов (***)	Ж
Воспроизводимость	Поощрение прозрачности и открытой науки	Открытые данные, материалы, программное обеспечение и т.д. (от * до **) Зарегистрированная отчетность (***) при клинических испытаниях, * при других исследованиях)	Ж, С, Р
Оценка	Диверсифицированное рецензирование	Препринты (* в биомедицинских/поведенческих науках, **** в физических науках) Экспертная оценка до и после публикации, например, Publons, PubMed Commons (*)	Ж
Стимулы	Награждение открытых и воспроизводимых практик	Жетоны (*) Зарегистрированная отчетность (*) Принципы продвижения прозрачности и открытости (*) Финансирование воспроизводимых исследований(*) Практика открытой науки при найме на работу и при повышении (*)	Ж, И, С

Предполагаемая текущая степень применения: \*, <5%; \*\*, 5–30%; \*\*\*, 30–60%; \*\*\*\*, >60%. Аббревиатуры ключевых заинтересованных сторон: Ж - журналы/издатели; С - спонсоры; И - институты; Р - регуляторы.

# Понимать и сопоставлять интересы заинтересованных сторон

**Таблица 1.** Некоторые крупные заинтересованные стороны в науке, а также степень их заинтересованности в исследованиях и их результатах с разных углов; представлены типичные шаблоны (бывают и исключения)

	Степень заинтересованности в результатах исследования			
	Публикуемые	Финансируемые	Внедряемые в практику	Прибыльные
Ученые	+++	+++	+	
Индустрия – продажи и маркетинг				+++
Индустрия – НИОКР			+++	+++
Частные инвесторы, в том числе хедж-фонды			++	+++
Государственные спонсоры – открытые (например, NIH (Национальный институт здоровья), NSF (Национальный научный фонд))	++		+	
Государственные спонсоры – закрытые (например, военные)			+++	
Некоммерческие спонсоры/филантропы	++		+++	
Редакторы журналов	+++			+
Коммерческие издатели	+			+++
Профессиональные и научные сообщества	+			
Университеты	+	+++		+
Некоммерческие исследовательские институты	+++	+++	+	+
Вспомогательный ненаучный персонал		+++		
Больницы и другие профессиональные организации, предоставляющие услуги, которые связаны с наукой			+	+++
Другие финансовые организации, на которые влияют эти услуги (например, страховые)				+++
Правительства и государственные/федеральные власти				++
Потребители продуктов и услуг			+++	

# Заключительные комментарии

- Воспроизводимость имеет разные значения в разных научных дисциплинах, но в конечном счете все они направлены на повышение достоверности результатов научных исследований
- Существует множество возможных мер, которые могут повысить эффективность исследовательских практик и воспроизводимости научной литературы
- Эмпирические мета-исследования могли бы помочь оценить не только распространенность проблемы воспроизводимости, но и эффективность и возможный вред от вмешательств, нацеленных на повышение воспроизводимости исследований.

# Особая благодарность.

- Daniele Fanelli, Stanford University
- Steve Goodman, Stanford University
- Shanil Ebrahim, Stanford University
- Debbie Dunne, Stanford University
- Despina Contopoulos-Ioannidis, Stanford University
- Robert Tibshirani, Stanford University
- Chirag Patel, Stanford University and Harvard University
- David Chavalarias, ISC, Paris
- Lamberto Manzoli, University of Ferrara
- Maria Elen Flacco, University of Chieti
- Paolo Villari, University of Rome La Sapienza
- Fainia Kavvoura, Oxford University
- Kostas Siontis, Mayo Clinic
- George Siontis, University of Bern
- Vangelis Evangelou, University of Ioannina
- Muin Khoury, CDC and NCI
- Orestis Panagiotou, National Cancer Institute
- Kevin Boyack, Map of Science, SciTech
- Glenn Begley, Amgen, TetraLogic
- Joseph Lau, Brown University
- Tiago Pereira, U Sao Paolo
- Malcolm MacLeod, University of Edinburgh
- Kostas Tsilidis, University of Ioannina
- David Allison, University of Alabama at Birmingham
- Brian Nosek, Center for Open Science
- Belinda Burford, University of Melbourne