



Точка зрения

Мифологизация статистики в биомедицинских исследованиях

Славенко М.Г.^{1,2} • Мироненко О.Н.¹ • Кузнецов М.С.^{1,3} • Бакин Е.А.¹

Славенко Матвей Григорьевич – преподаватель¹; биостатистик²; ORCID: <https://orcid.org/0000-0001-6610-9824>

✉ 197342, г. Санкт-Петербург, ул. Кантемировская, 2а, Российская Федерация.
E-mail: matvei.slavenko@gmail.com

Мироненко Ольга Николаевна – канд. экон. наук, преподаватель¹; ORCID: <https://orcid.org/0000-0001-8952-8386>. E-mail: freelomir@yandex.ru

Кузнецов Максим Сергеевич – преподаватель¹; ст. лаборант кафедры пропедевтики внутренних болезней имени профессора С.С. Зимницкого³; ORCID: <https://orcid.org/0000-0002-4891-9169>. E-mail: bqmaks@gmail.com

Бакин Евгений Александрович – канд. техн. наук, преподаватель¹; ORCID: <https://orcid.org/0000-0002-5694-4348>. E-mail: eugene.bakin@gmail.com

В последние годы большое внимание уделяется проблеме воспроизводимости научных исследований. В качестве одной из причин кризиса воспроизводимости выделяют некорректное применение исследователями методов статистического анализа. С опорой на концепцию «статистических мифов» (устойчивое, воспроизводящееся искажение статистических понятий) утверждается, что мифы носят неслучайный характер, имеют свои функции и свидетельствуют о глубоком методологическом разрыве между предметной областью и практикой статистического анализа. Связь между научным и статистическим выводами показана в виде схемы биомедицинского исследования, состоящего из семи этапов: концептуализация и постановка исследовательского вопроса; операционализация; выбор эстиманда; создание и применение статистической модели; интерпретация оценки с учетом неопределенности; клиническая интерпретация; ответ на исследовательский вопрос. Схема иллюстрирована примерами трех реальных исследований из различных областей медицины. На основании выполняемых функций авторы классифицируют мифы на эпистемологические, фундаментальные, инструментальные и терминологические. Приведены

характерные признаки и примеры каждого типа, проанализировано несколько популярных мифов. Предложенный подход позволяет рассматривать статистические мифы не только как педагогическую проблему, но и как индикатор более глубоких нарушений логики исследования. С методической точки зрения исследователю предлагается не запомнить очередной список «правильных» и «неправильных» статистических действий, а последовательно двигаться от научного вопроса к способу его статистической формализации.

Ключевые слова: биостатистика, биомедицинские исследования, методология, миф, воспроизводимость, статистический вывод, научный вывод, терминология

Для цитирования: Славенко МГ, Мироненко ОН, Кузнецов МС, Бакин ЕА. Мифологизация статистики в биомедицинских исследованиях. Альманах клинической медицины. 2026;54(2):104–112. doi: 10.18786/2072-0505-2026-54-010.

Поступила 25.06.2026; доработана 10.08.2026; принята к публикации 14.08.2026; опубликована онлайн 25.08.2026

В 2016 г. в журнале Nature была опубликована статья М. Baker “1,500 scientists lift the lid on reproducibility”, ставшая одним из символов так называемого кризиса воспроизводимости [1]. В ней были описаны результаты опроса более 1500 ученых, большинство из которых признали, что им не удавалось воспроизвести результаты как чужих, так и собственных исследований. Работа подчеркнула системный характер проблемы, вызванной давлением на ученых с целью повышения публикационной активности, селективностью публикуемых материалов, а также необходимостью тратить существенное время на бюрократические издержки. В первую очередь эти факторы приводили к слабому обоснованию дизайна исследования, неадекватной статистической обработке, а также отсутствию полноценного наставничества для молодых ученых. Данная работа подтвердила реальность опасений, формулируемых на теоретическом уровне задолго до ее появления [2].

Как следствие, в научном сообществе активизировалось обсуждение действий, направленных на повышение воспроизводимости научных результатов [3]. Большая часть предложений касалась организационных мер, направленных на 1) уменьшение влияния субъективных факторов через пререгистрацию исследования с подробным описанием дизайна, конечных точек и методов анализа; 2) устранение конфликтов интересов посредством привлечения независимых экспертов к оценке качества исследований; 3) повышение точности исследований за счет поощрения сотрудничества разных команд, работающих над одной тематикой.

Особо выделим меры по устранению непреднамеренных ошибок, вызванных непониманием сути применяемых методов. M.R. Munafo и соавт. предложили обратить внимание на повышение качества образования за счет создания открытых и наглядных курсов для обучения статистике в рамках концепции непрерывного образования [3].



Эта идея поддерживается данными других исследований, показавшими, что для студентов статистика представляет собой одну из наиболее сложных дисциплин, будучи самой частой причиной отчислений из университетов и колледжей из-за академической неуспеваемости [4, 5]. Ведущими факторами называют слабую математическую подготовку слушателей и высокую степень абстрактности статистических концепций [6]. Особенно критичными они становятся для студентов нетехнических специальностей, в частности биологического и медицинского направлений. Как правило, статистика читается в специалитете по устаревшим методикам и усваивается плохо в силу кажущегося отсутствия прикладного смысла, а значит, и мотивации заниматься сложной абстрактной дисциплиной [7]. Впоследствии аспиранты и молодые ученые биомедицинских специальностей сталкиваются с необходимостью самостоятельно проводить статистический анализ данных, во многом в силу отсутствия штатных статистиков в учреждениях биомедицинского профиля [8]. В сочетании с необходимостью регулярных публикаций это создает благоприятную среду для появления различных учебных курсов, самоучителей и пособий, авторы которых обещают обучить статистике быстро и без серьезных трудозатрат. Обещание подкрепляется возможностями современных программных пакетов, позволяющих путем несложных действий сформировать правдоподобный статистический отчет на основе имеющихся данных¹ [9].

В действительности же подобные методики способствуют алгоритмизации навыков статистической обработки вместо осознанности, возникновению ложного чувства доступности и простоты статистического инструментария [10]. В попытке упростить восприятие статистических концепций нарушается строгость терминологии, размывается смысл понятий, возникают серьезные проблемы с интерпретацией полученных результатов. Как отметил О. Wegwarth, подобное положение вещей приводит к ошибкам не только при проведении научных исследований, но и при принятии клинических решений при работе с пациентами [11]. Обучившиеся по подобным методикам специалисты в дальнейшем и сами распространяют некорректные концепции среди коллег и своих учеников.

Как результат, рождается *статистический миф* – устойчивое упрощенное или искаженное представление о статистических понятиях, методах или принципах, которое воспроизводится в образовательной, исследовательской или профессиональной среде и некритично воспринимается как

истина, несмотря на его несоответствие формальной теории вероятностей и принципам статистического вывода. Статистические мифы позволяют говорить о статистике без необходимости понимать ее.

В мировой литературе проблема мифологизации статистики имеет давнюю историю обсуждения. Отдельно отметим инициативу А. Althouse, который популяризировал понятие «статистический миф» и вдохновил авторов настоящей статьи ознакомить русскоязычное научное сообщество с наиболее распространенными ошибочными концепциями². Этой работой мы начинаем серию статей, посвященных обсуждению таких концепций, их корней и того, к каким последствиям они могут привести исследователя. Ниже рассмотрим общий источник появления мифов, мифологизацию статистической терминологии и наиболее частые примеры статистических мифов. Материалы, положенные в основу настоящей публикации, были частично представлены в серии открытых лекций, проводимых авторами на базе ООО «Институт биоинформатики и биостатистики» (Институт биоинформатики)³.

Место статистики в биомедицинском исследовании

В прикладных биомедицинских исследованиях статистический анализ нередко используется как набор стандартизованных процедур, слабо связанных с логикой исследовательского вопроса. В результате статистика превращается из инструмента научного поиска в ритуал легитимации результата. Такое ритуализированное восприятие приводит к тому, что вопрос «что именно мы хотим узнать?» незаметно подменяется вопросом «какой тест следует применить?» [12]. Кроме того, результаты статистического анализа воспринимаются как «математически доказанные истины», которые говорят сами за себя. Именно поэтому принципиально важно понимать функцию статистического вывода: статистика не производит научный вопрос и не заменяет его. Она формализует связь между собранными данными и *заранее* заданным объектом анализа, а также производит некоторые факты на языке математики, которые требуют дальнейшей интерпретации в контексте прикладной области. Иными словами, научный вывод (англ. *scientific inference*) не равен выводу статистическому (англ. *statistical inference*) [13, 14].

Переход между плоскостями исследовательского вопроса и статистического вывода осуществляется в ходе определенной последовательности этапов: концептуализация исследовательского вопроса, операционализация переменных, выбор эстиманда,



Место статистического анализа в процессе научного поиска

построение статистической модели, оценивание и интерпретация с учетом неопределенности (рисунок). Термины «концептуализация» и «операционализация» широко применяются в социальных науках [15], и мы берем на себя смелость утверждать, что они полезны и в контексте биомедицинских исследований.

На первом этапе формулируется исследовательский вопрос. Здесь необходимо определить, какими прикладными (предметными) понятиями (конструктами) мы оперируем в исследовании, а также интересует ли исследователя каузальный (причинно-следственный) эффект вмешательства, ассоциация между признаками, индивидуальный прогноз или гетерогенность наблюдаемых признаков. Пока этот уровень не прояснен, статистические решения остаются техническими действиями без ясного эпистемического* статуса.

На втором этапе происходит операционализация: теоретические конструкты переводятся в наблюдаемые переменные, а воздействие, исход, ковариаты, измерения и пропущенные данные получают четкие дефиниции.

После этого должен быть определен эстиманд, то есть точный объект статистического оценивания. Современный регулятивный документ ICH E9(R1) рассматривает эстиманд как связующее звено между целью исследования, дизайном, данными, методом оценивания и интерпретацией результата⁴. Эстиманд определяет, какой именно эффект, для какой популяции, на каком временном горизонте и при каких условиях должен быть оценен. Только после выбора эстиманда становится оправданным создание математической модели.

Необходимо различать предсказательные модели и модели статистического вывода. Главная цель предсказательных моделей – прогноз будущих или ненаблюдаемых исходов, моделей вывода – оценивание параметра, ассоциации или каузального эффекта и описание неопределенности этого вывода. В прикладной литературе эти способы мышления часто смешиваются, из-за чего подходы, полезные

для решения одного типа задач, начинают использоваться для решения задач другого типа, что может привести к получению некорректного ответа на поставленный исследовательский вопрос [16]. Например, коэффициент детерминации R^2 , являющийся метрикой предсказательного качества модели линейной регрессии, часто применяется для оценки моделей, используемых для оценки эффекта, где он представляет ограниченный интерес.

Безусловно, при выборе модели важно понимать ее допущения, отвечающие за корректность результатов выбранных процедур: состоятельность оценок и доверительных интервалов, соблюдение ожидаемых вероятностей ошибки 1-го и 2-го рода и пр. Однако во многих прикладных текстах внимание сосредоточено на ритуальной проверке допущений (например, «нормальности данных»), хотя главная проблема может заключаться в неясном эстиманде, некорректной операционализации или несоответствии модели и задачи.

Завершается анализ применением выбранной модели (моделей) и интерпретацией результатов. Интерпретация также осуществляется в несколько этапов, согласованных с соответствующими этапами, предшествующими построению модели. Сначала проводится статистическая интерпретация с учетом неопределенности: на основе точечных оценок, доверительных интервалов, p -значений и прочих инструментов делается вывод о величине эффекта и неопределенности, связанной с его оценкой. На следующем этапе проводится интерпретация с учетом специфики доменной области: делается вывод о клинической значимости эффекта, прикладной ценности полученных результатов, согласованности с имеющимися практиками и др. Завершается процесс резюмированием степени успешности проведенного исследования и оценкой необходимости изменения текущей научной парадигмы.

Обратите внимание: полученные результаты обладают вероятностной природой, которая подчас игнорируется. Как подчеркивается в заметках A. Gelman, исследователи часто склонны отрицать

* Эпистемический – относящийся к знаниям, оценке их надежности и процессам познания в целом.



не только неопределенность, но и саму вариацию данных, однако именно вариация и гетерогенность нередко составляют содержательную часть науки^{5,6}. В таблице 1 приведены примеры, иллюстрирующие работу указанной схемы.

Таким образом, статистика – это не автономный этап «математической обработки данных».

Ее функция состоит в том, чтобы формально связать исследовательский вопрос, популяцию интереса, способ порождения данных, эстиманд, модель и меру неопределенности. Там, где эта связь теряется, статистика превращается в язык ритуалов, и именно в этом пространстве возникают и воспроизводятся статистические мифы.

Таблица 1. Примеры формализации решения научных проблем*

Этап	Пример 1	Пример 2	Пример 3
Научный вопрос	Является ли телемедицинское ведение детей и подростков с сахарным диабетом 1-го типа не уступающей альтернативой очным визитам?	Следует ли учитывать уровень экспрессии PD-L1 при выборе терапии 1-й линии у пациентов с распространенным НМРЛ без драйверных мутаций?	Как корректно учесть типовые значения концентрации гомоцистеина?
Концептуализация	Рассматривается ведение сахарного диабета 1-го типа у детей и подростков – аутоиммунного эндокринного заболевания, при котором дефицит инсулина обусловлен разрушением β -клеток поджелудочной железы. Клинически значимым считается не только контроль гликемии, но и нагрузка, связанная с получением медицинской помощи: затраты времени, усилий и ресурсов пациента и семьи. Телемедицинское ведение понимается как альтернативная организационная модель наблюдения, которая должна сохранять качество контроля заболевания и уменьшать эту нагрузку	Рассматривается распространенный НМРЛ без драйверных мутаций, при котором в 1-й линии терапии возможен выбор между стандартной химиотерапией и химиотерапией с добавлением иммунотерапии анти-PD-1-препаратом тислилизумаб. Экспрессия PD-L1 концептуализируется как потенциальный предиктивный биомаркер – характеристика опухоли, с которой может быть связано изменение эффекта иммунохимиотерапии относительно химиотерапии	Рассматривается концентрация гомоцистеина как прогностический фактор, ассоциированный с риском развития сердечно-сосудистых событий. Интерес представляет не эффект лечения, а распределение маркера среди людей без сердечно-сосудистых заболеваний. Диапазон типовых значений понимается как значение, выше которого результат становится необычно высоким для здорового человека того же пола и возраста, обследованного тем же методом и в сопоставимых условиях
Операционализация	Гликемический контроль операционализируется через концентрацию HbA1c в крови. Дополнительно могут использоваться показатели непрерывного мониторинга глюкозы: доля времени, проведенная в целевом и иных гликемических диапазонах. Нагрузка, связанная с визитом, операционализируется как время, затраченное на плановое посещение диabetолога каждые 3 месяца: от выхода из дома до возвращения домой	Сравнимые лечебные стратегии: иммунохимиотерапия и химиотерапия в 1-й линии. Исход: ВБП. Модификатор эффекта: исходный уровень экспрессии PD-L1 в разбивке на 3 уровня: низкий (< 1%), средний (1–49%) и высокий ($\geq 50\%$); задача состоит в оценке эффекта терапии в зависимости от этого уровня	Переменная интереса: концентрация гомоцистеина в плазме (ммоль/л), измеренная откалиброванным аналитическим прибором. Ковариаты для стратификации или моделирования референсной границы: пол и возраст
Выбор эстиманда	Основной эстиманд: разность средних значений HbA1c между телемедицинским ведением и очными визитами в целевой популяции детей и подростков с сахарным диабетом 1-го типа: $\Delta = \mu_{\text{телемедицина}} - \mu_{\text{очные визиты}}$. Телемедицина считается не уступающей альтернативой, если средний уровень HbA1c в группе телемедицинского вмешательства по сравнению с контролем выше не более чем на 5 ммоль/моль. Дополнительные целевые величины: разность средних долей времени в гликемических диапазонах и разность среднего времени на визит. Если интерес представляет вариабельность, она задается отдельной целевой величиной, например дисперсией HbA1c или IQR времени в целевом диапазоне	Основной эстиманд гетерогенности эффекта: зависимость HR при иммунохимиотерапии по сравнению с химиотерапией от исходного уровня экспрессии PD-L1 в целевой популяции пациентов с распространенным НМРЛ без драйверных мутаций, получающих терапию 1-й линии. Клинически существенным будет считаться изменение интенсивности событий ВБП на 20% ($HR \leq 0,8$ или $HR \geq 1,2$)	Оцениваемая величина: условный 0,95-квантиль (95-й процентиль) распределения концентрации гомоцистеина в зависимости от пола и возраста
Статистическая модель / процедура	Оценка разности средних с 95% ДИ на основе t-процедуры Уэлча. Для вывода о не меньшей эффективности используется сравнение верхней границы 95% ДИ с заранее заданной границей 5 ммоль/моль	Регрессия Кокса для ВБП с предикторами: стратегия терапии 1-й линии, уровень экспрессии PD-L1 (2 бинарные переменные для среднего и высокого уровней экспрессии, с низким	Обобщенная аддитивная модель параметров положения, масштаба (GAMLSS) и формы условного распределения концентрации гомоцистеина в зависимости



		экспрессией в качестве референсной категории) и их взаимодействие. Для вывода о практической значимости данного биомаркера при выборе линии терапии используется сравнение границ 95% ДИ для HR с заранее заданной границей $\leq 0,8$ или $\geq 1,2$	от пола (бинарный фактор) и возраста (сплайновая функция)
Результат	95% ДИ для разности средних значений HbA1c (телемедицина – очные визиты) составил (-10,2; 1,9) ммоль/моль. 95% ДИ для разности среднего времени на визит составил (-237; -157) минут, $p < 0,001$	Значение p для проверки гипотезы об одновременном равенстве единице HR для эффектов взаимодействия между терапией, с одной стороны, и средним и высоким уровнями экспрессии PD-L1, с другой, составило 0,0324. HR (95% ДИ) для иммунохимиотерапии по сравнению с химиотерапией при низкой экспрессии PD-L1 составило 0,733 (0,456–1,179), при средней экспрессии – 1,095 (0,526–2,277), при высокой – 0,336 (0,185–0,611)	Построена функция для вычисления верхнего референсного предела при подстановке соответствующих значений пола и возраста, например: • в 40 лет у мужчин верхний референсный предел равен 16,6 мкмоль/л; • в 60 лет у женщин верхний референсный предел равен 14,8 мкмоль/л
Интерпретация оценки эстиманда	Полученные данные совместимы как со снижением среднего уровня HbA1c в группе телемедицины на 10,2 ммоль/моль по сравнению с очными визитами, так и с его повышением на 1,9 ммоль/моль. Для времени на визит данные совместимы со снижением среднего времени на 157–237 минут	Полученные данные плохо совместимы с нулевой гипотезой об отсутствии зависимости эффекта иммунохимиотерапии от уровня экспрессии PD-L1. При этом для пациентов с высокой экспрессией PD-L1 ($\geq 50\%$) данные хорошо согласуются с клинически существенным положительным эффектом иммунохимиотерапии; для пациентов с низкой экспрессией PD-L1 ($< 1\%$) – как с клинически существенным положительным, так и с клинически несущественным эффектом, для пациентов со средней экспрессией PD-L1 (1–49%) – от клинически существенного положительного до клинически существенного отрицательного эффекта	Если рассматриваются здоровые люди заданного пола и возраста, обследованные тем же методом и при тех же условиях, примерно 95% их значений концентрации гомоцистеина будут не выше соответствующего порога
Интерпретация в контексте доменной области	В оптимистичном сценарии телемедицинское ведение связано со снижением HbA1c в среднем на 10,2 ммоль/моль, что соответствует существенному улучшению качества ведения диабета. В пессимистичном сценарии средний уровень HbA1c выше на 1,9 ммоль/моль, что в целом соответствует пренебрежимо малому ухудшению качества ведения диабета. Кроме того, телемедицинский формат существенно сокращает время, затрачиваемое пациентами на консультации, по сравнению с классическими очными визитами: экономия составляет примерно 3 часа	Чем выше уровень экспрессии PD-L1, тем больше ожидаемое преимущество иммунохимиотерапии по ВБП относительно химиотерапии; однако при низкой и средней экспрессии PD-L1 данных недостаточно для более точных выводов относительно эффекта иммунохимиотерапии	Если значение концентрации гомоцистеина у пациента превышает его половозрастной верхний референсный предел, это может указывать на риск развития сердечно-сосудистых событий и могут потребоваться дополнительные диагностические процедуры, рекомендации по изменению образа жизни и др.
Ответ на исследовательский вопрос	Поскольку HbA1c отражает среднюю гликемию за последние 3–4 месяца и связан с отдаленными осложнениями сахарного диабета, выбранная граница не меньшей эффективности имеет клиническую интерпретацию. Правая граница 95% ДИ для Δ находится ниже границы 5 ммоль/моль; следовательно, телемедицинское ведение может быть признано не уступающей альтернативой полностью очным консультациям. Одновременно телемедицина существенно экономит время пациентов, что в перспективе может повысить приверженность пациента лечению	Уровень экспрессии PD-L1 может учитываться при принятии решения о выборе между иммунохимиотерапией и химиотерапией в 1-й линии, в том числе с учетом профиля безопасности обеих стратегий. При высоком уровне экспрессии PD-L1 ($\geq 50\%$) можно говорить о существенном ожидаемом преимуществе иммунохимиотерапии по ВБП по сравнению с химиотерапией. Для пациентов с низким и средним уровнем экспрессии PD-L1 необходимы дополнительные исследования, позволяющие получить более точные оценки	Полученная функция позволяет определять верхний референсный предел концентрации гомоцистеина в зависимости от пола и возраста. Эта граница описывает поведение маркера в здоровой популяции и помогает отделять необычно высокие значения от ожидаемого половозрастного диапазона

HbA1c – гликированный гемоглобин, HR (hazard ratio) – отношение рисков, IQR (interquartile range) – межквартильный размах, PD-L1 (programmed death-ligand 1) – лиганд рецептора программируемой клеточной смерти 1, ВБП – выживаемость без прогрессирования, ДИ – доверительный интервал, НМРЛ – немелкоклеточный рак легкого

* Пример 1 основан на публикации [17], пример 2 – [18], пример 3 – [19]



Таблица 2. Классификация статистических мифов

Тип мифа	Признак	Функция	Примеры
Эпистемологический	Сведёние исследования к проведению статистического анализа	Компенсировать недостаток понимания доменной области и отсутствие четко сформулированной цели исследования	Если статистический анализ выполнен правильно (в соответствии с канонами математической статистики), то исследование уже является научно корректным. Данные говорят сами за себя, их нужно просто правильно проанализировать; понимать способ их происхождения для этого необязательно
Фундаментальный	Упрощение картины мира, приводящее к искажению природы статистического вывода	Создание иллюзии строгости и однозначности там, где на деле присутствует неопределенность	Отождествление свойств выборки и свойств генеральной совокупности. Отождествление статистической и практической (клинической) значимости. Интерпретация отсутствия статистической значимости как отсутствия различий / эффекта. Необоснованная категоризация непрерывных признаков
Инструментальный	Рутинизация и алгоритмизация выбора статистических методов. Упрощение использования технического аппарата. Наделение методов функциями, которыми они не обладают	Снижение когнитивной нагрузки при проведении статистического анализа. Создание ощущения простоты в освоении навыков статистического анализа. Придание уверенности в корректности полученных результатов при следовании опубликованному алгоритму. Придание результатам видимости «объективности», мнимое снижение «человеческого фактора» и «субъективности суждения»	Блок-схемы для выбора «правильного» статистического теста на основе свойств данных (например, «нормальности»). Автоматический отбор признаков для множественной регрессии. Перенос методов построения, анализа и интерпретации предиктивных моделей в область статистического вывода. Расчет мощности исследования постфактум, после проведения статистического анализа. «Несравнимые» рандомизированные исследования. Использование теста Манна – Уитни для сравнения медиан. Вера в то, что рандомизация делает ненужной поправку на сопутствующие ковариаты
Терминологический	Упрощение определений статистических терминов до некорректных. Привнесение понятий, не имеющих четких определений в теории вероятностей и статистике	Упрощение коммуникации между исследователями, создание иллюзии одинакового понимания осуществляемых процедур	Некорректные определения р-значения и доверительного интервала. Часто употребляемые понятия: тенденция к статистической значимости; достоверность различий; выброс

Типы статистических мифов

В соответствии с выполняемыми функциями⁷ [20] статистические мифы можно условно классифицировать на типы (табл. 2). Подчеркнем: классификация носит условный характер, а предложенные типы не являются взаимоисключающими; часть мифов может относиться одновременно к нескольким типам.

Некоторые наиболее частые примеры статистических мифов

Чрезмерная сосредоточенность на тестировании гипотез

Одно из наиболее устойчивых проявлений статистического мифа – восприятие проверки нулевой гипотезы как главного, а иногда и единственного инструмента статистического вывода. При таком подходе исследовательский вопрос подменяется вопросом о том, удалось ли получить р-значение ниже заранее принятого порога. Между тем статистический вывод не исчерпывается тестированием:

во многих ситуациях центральное значение имеют оценивание величины эффекта, описание неопределенности, анализ чувствительности и сопоставление альтернативных моделей. Сведёние анализа к тестированию гипотез создает ложную бинарную картину мира, в которой сложные количественные выводы заменяются дихотомией «значимо / незначимо», «позитивными» и «негативными» результатами исследования [21].

Проверки на нормальность

Предварительная проверка «нормальности данных» остается одним из самых распространенных статистических ритуалов в биомедицинских публикациях. Между тем проблема обычно формулируется неверно: для многих стандартных процедур существенны не «нормальные данные» как таковые, а свойства оценок, остатков, средних или статистик теста в конкретной задаче. В работе Т. Lumley и соавт. показано, что t-тест и линейная регрессия зачастую работоспособны и при



резко ненормальных распределениях исхода [22], тогда как J. Rochon и соавт. продемонстрировали, что двухэтапная схема «сначала тест на нормальность, потом выбор между t-тестом и непараметрическим тестом» сама по себе может ухудшать статистические свойства анализа [23]. Отдельно отметим, что выбор между параметрическим и непараметрическим тестом на основе проверки на нормальность нарушает описанную выше последовательность проведения научного поиска (см. рисунок), так как анализ данных начинает диктовать выбор эстиманда.

Жесткие границы для размера эффекта и р-значений

Стремление устанавливать жесткие универсальные границы (например, трактовать корреляцию 0,29 как «слабую», а 0,31 – уже как «умеренную» или считать $p = 0,049$ качественно отличным от $p = 0,051$) отражает более общее желание замкнуть рассуждение механическим «объективным» правилом. Современные методологические публикации, напротив, подчеркивают необходимость интерпретировать величину эффекта и меру неопределенности в связи с предметной областью, дизайном исследования и целью анализа⁸ [24, 25]. Использование жестких порогов упрощает процесс статистического анализа, но методологически зачастую создает ложное ощущение объективности там, где требуется исследовательское суждение.

Применение термина «достоверность» к результатам статистического анализа

Одним из характерных мифологических терминов русскоязычного пространства является определение «достоверно» в применении к результатам статистического анализа. Такая практика создает ложное впечатление, будто статистически значимый результат является не только

следствием применения определенной модели и выбранного уровня значимости, но и в некотором смысле «истинным» или «надежно доказанным». По мнению Н.А. Зорина, подмена термина «статистическая значимость» термином «достоверность» в российских медицинских публикациях способна искажать восприятие исследования в целом, поскольку переводит вероятностный вывод в детерминистский [26].

Заключение

Статистические мифы – не случайный набор частных ошибок или неудачных формулировок, а закономерное следствие разрыва между исследовательским вопросом и статистическим анализом. Когда статистика начинает восприниматься как автономная процедура «обработки данных», а не как формальный язык, связывающий вопрос исследования, операционализацию, эстиманд, модель и неопределенность, создаются условия для мифологизации статистических понятий, методов и интерпретаций.

Предложенный нами подход позволяет рассматривать статистические мифы не только как педагогическую проблему, но и как индикатор более глубоких нарушений логики исследования. В зависимости от того, на каком этапе происходит этот разрыв, формируются различные типы мифов: эпистемологические, фундаментальные, инструментальные и терминологические. Такое рассмотрение позволяет перейти от простого перечисления ошибочных представлений к анализу их происхождения и функций в исследовательской практике. Практический вывод из этого состоит не в необходимости запомнить очередной список «правильных» и «неправильных» статистических действий, а в требовании последовательно двигаться от научного вопроса к способу его статистической формализации. ☺

¹ Fox J. Guest post: Is statistical software harmful? [Internet]. Available from: <https://dynamicecology.wordpress.com/2014/06/04/guest-post-is-statistical-software-harmful> (accessed August 10, 2026).

² Althouse A. Reference collection to push back against “Common Statistical Myths” [Internet]. Available from: <https://discourse.datamethods.org/t/reference-collection-to-push-back-against-common-statistical-myths/1787> (accessed August 10, 2026).

³ Институт биоинформатики. Онлайн-лекторий «Разрушители статистических мифов» [Интернет]. Доступно по: https://bioinf.me/stat_myths (дата обращения 10.08.2026).

⁴ ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. 17 February 2020 [Internet]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-and-sensitivity-analysis-clinical-trials-guideline-statistical-principles-clinical-trials-step-5_en.pdf (accessed August 10, 2026).

⁵ Gelman A. The inclination to deny all variation [Internet]. Available from: <https://statmodeling.stat.columbia.edu/2014/12/10/inclination-deny-variation/> (accessed August 10, 2026).

⁶ Gelman A. When there's a lot of variation, it can be a mistake to make statements about 'typical' attitudes [Internet]. Available from: <https://statmodeling.stat.columbia.edu/2014/10/08/var/> (accessed August 10, 2026).

⁷ Buxton RG, Bolle KW, Smith JZ. Myth. In: Encyclopedia Britannica. 2026 [Internet]. Available from: <https://www.britannica.com/topic/myth> (accessed August 10, 2026).

⁸ Jané MB, Xiao Q, Yeung S, Azevedo F, Ben-Shachar MS, Caldwell AR, Cousineau D, Dunleavy DJ, Elsherif M, Harlow TJ, Johnson B, Moreau D, Riethuis P, Röseler L, Steele J, Vieira FF, Zloteanu M, Feldman G. Guide to effect sizes and confidence intervals. 2024 [Internet]. Available from: <https://matthewbjane.quarto.pub/guide-to-effect-sizes-and-confidence-intervals#cite-this-guide> (accessed August 10, 2026).



Дополнительная информация

Финансирование

Работа проведена без привлечения дополнительного финансирования со стороны третьих лиц.

Конфликт интересов

Авторы заявляют об отсутствии явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Участие авторов

Все авторы внесли равный вклад в разработку концепции исследования, подготовку рукописи и ее редактирование. Все авторы прочли и одобрили финальную версию статьи перед публикацией, согласны нести ответственность за все аспекты работы и гарантируют, что ими надлежащим образом были рассмотрены и решены вопросы, связанные с точностью и добросовестностью всех частей работы.

Использование искусственного интеллекта при написании статьи

Авторами было использовано программное обеспечение ChatGPT (OpenAI) как вспомогательный источник для подбора релевантной литературы. Авторы изучили и перепроверили все выдачи искусственного интеллекта и полностью принимают на себя ответственность за содержание статьи.

Благодарности

Авторы благодарят всех участников лектория «Разрушители статистических мифов» и в особенности Наталью Зорину, Алексея Глазкова, Артемия Охотина, Богдана Сотникова, Алису Селезневу, Елену Чуклину и Елену Убогоеву за плодотворные дискуссии, критику и ценные замечания, высказанные при обсуждении данных материалов.

Список литературы / References

- Baker M. 1,500 scientists lift the lid on reproducibility. *Nature*. 2016;533(7604):452–454. doi: 10.1038/533452a.
- Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124. doi: 10.1371/journal.pmed.0020124.
- Munafò MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, du Sert NP, Simonsohn U, Wagenmakers EJ, Ware JJ, Ioannidis JPA. A Manifesto for reproducible science. *Nat Hum Behav*. 2017;1(1):0021. doi: 10.1038/s41562-016-0021.
- Fitzmaurice O, Leavy A, Hannigan A. Why is statistics perceived as difficult and can practice during training change perceptions? Insights from a prospective mathematics teacher. *Teach Math Appl*. 2014;33(4):230–248. doi: 10.1093/teamat/hru010.
- Babalola BT, Awe OO, Adarabioyo MI. Challenges of statistics education that leads to high dropout rate among undergraduate statistics students in developing countries. In: Awe OO, Vance EA, Olubusoye OE, editors. *Promoting statistical practice and collaboration in developing countries*. Boca Raton: Chapman and Hall/CRC, 2022. P. 75–83. doi: 10.1201/9781003261148-7.
- Silva PN, Sarnecka BW. What do your students struggle with? A survey of statistics instructors. *J Stat Data Sci Educ*. 2025;33(1):1–12. doi: 10.1080/26939169.2025.2455560.
- Янцер ЛВ. Методика обучения математической статистике в медицинских вузах. *KANT*. 2017;(3):75–78. Yantser LV. [Methodology for teaching mathematical statistics in medical institutions of higher education]. *KANT*. 2017;(3):75–78. Russian.
- Awe OO. Building biostatistics capacity in developing countries. In: Awe O, Love K, Vance E, editors. *Promoting statistical practice and collaboration in developing countries*. Boca Raton: Chapman and Hall/CRC, 2022. P. 433–442. doi: 10.1201/9781003261148-43.
- Ercan I, Yazıcı B, Yang Y, Özkaya G, Cangur S, Ediz B, Kan I. Misusage of statistics in medical research. *Eur J Gen Med*. 2007;4(3):128–134. doi: 10.29333/ejgm/82507.
- Gigerenzer G. Mindless statistics. *J Socio Econ*. 2004;33(5):587–606. doi: 10.1016/j.soec.2004.09.033.
- Wegwarth O. Statistical illiteracy in residents: What they do not learn today will hurt their patients tomorrow. *J Grad Med Educ*. 2013;5(2):340–341. doi: 10.4300/JGME-D-13-00084.1.
- McElreath R. *Statistical rethinking: A bayesian course with examples in R and Stan*. 2nd edition. Boca Raton: Chapman and Hall/CRC, 2020. 612 p. doi: 10.1201/9780429029608.
- Marmolejo-Ramos F, Cousineau D. Perspectives on the use of null hypothesis statistical testing. Part I: The mighty frames of scientific and statistical inference. *Educ Psychol Meas*. 2017;77(3):471–474. doi: 10.1177/0013164416667986.
- Hubbard R, Haig BD, Parsa RA. The limited role of formal statistical inference in scientific inference. *Am Stat*. 2019;73(Suppl 1):91–98. doi: 10.1080/00031305.2018.1464947.
- Conceptualization, operationalization, and measurement. In: Lewis-Beck MS, Bryman A, Liao TF, editors. *The SAGE encyclopedia of social science research methods*. Thousand Oaks (CA): Sage Publications, Inc., 2004. P. 162–165. doi: 10.4135/9781412950589.n150.
- van Diepen M, Ramspek CL, Jager KJ, Zoccali C, Dekker FW. Prediction versus aetiology: Common pitfalls and how to avoid them. *Nephrol Dial Transplant*. 2017;32(Suppl 2):ii1–ii5. doi: 10.1093/ndt/gfw459.
- Plachy L, Neuman V, Velichova K, Slavenko MG, Santova A, Anne Amaratunga S, Obermannova B, Kolouskova S, Pruhova S, Sumnik Z, Petruzelkova L. Telemedicine maintains good glucose control in children with type 1 diabetes but is not time saving for healthcare professionals: KITES randomized study. *Diabetes Res Clin Pract*. 2024;209:111602. doi: 10.1016/j.diabres.2024.111602.
- Lu S, Wang J, Yu Y, Yu X, Hu Y, Ai X, Ma Z, Li X, Zhuang W, Liu Y, Li W, Cui J, Wang D, Liao W, Zhou J, Wang Z, Sun Y, Qiu X, Gao J, Bao Y, Liang L, Wang M. Tislelizumab plus chemotherapy as first-line treatment for locally advanced or metastatic nonsquamous NSCLC (RATIONALE 304): A randomized phase 3 trial. *J Thorac Oncol*. 2021;16(9):1512–1522. doi: 10.1016/j.jtho.2021.05.005.
- Ma C, Li L, Wang X, Hou L, Xia L, Yin Y, Cheng X, Qiu L. Establishment of reference interval and aging model of homocysteine using real-world data. *Front Cardiovasc Med*. 2022;9:846685. doi: 10.3389/fcvm.2022.846685.
- Campbell J. *Pathways to bliss: Mythology and personal transformation*. Novato (CA): New World Library, 2004. 194 p.
- Wasserstein RL, Schirm AL, Lazar NA. Moving to a world beyond “p < 0.05”. *Am Stat*. 2019;73(Suppl 1):1–19. doi: 10.1080/00031305.2019.1583913.
- Lumley T, Diehr P, Emerson S, Chen L. The importance of the normality assumption in large public health data sets. *Annu Rev Public Health*. 2002;23:151–169. doi: 10.1146/annurev.publhealth.23.100901.140546.
- Rochon J, Gondan M, Kieser M. To test or not to test: Preliminary assessment of normality when comparing two independent samples. *BMC Med Res Methodol*. 2012;12:81. doi: 10.1186/1471-2288-12-81.
- Lakens D, Adolphi FG, Albers CJ, Anvari F, Apps MAJ, Argamon SE, Baguley T, Becker RB, Benning SD, Bradford DE, Buchanan EM, Caldwell AR, Van Calster B, Carlsson R, Sau-Chen S-C, Chung B, Colling LJ, Collins GS, Crook Z,



Cross ES, Daniels S, Danielsson H, DeBruine L, Dunleavy DJ, Earp BD, Feist MI, Ferrelle JD, Field JG, Fox NW, Friesen A, Gomes C, Gonzalez-Marquez M, Grange JA, Grieve AP, Guggenberger R, Grist J, van Harmelen A-L, Hasselman F, Hochard KD, Hoffarth MR, Holmes NP, Ingre M, Isager PM, Isotalus HK, Johansson C, Juszczak K, Kenny DA, Khalil AA, Konat B, Lao J, Larsen EG, Lodder GMA, Lukavský J, Madan CR, Manheim D, Martin SR, Martin AE, Mayo DG, Mc-

Carthy RJ, McConway K, McFarland C, Nio AQX, Nilsson G, De Oliveira CL, De Xivry J-J-O, Parsons S, Pfuhr G, Quinn KA, Sakon JJ, Saribay SA, Schneider IK, Selvaraju M, Sjoerds Z, Smith SG, Smits T, Spies JR, Sreekumar V, Steltenpohl CN, Stenhouse N, Świątkowski W, Vadillo MA, Van Assen MALM, Williams MN, Williams SE, Williams DR, Yarkoni T, Ziano I, Zwaan RA. Justify your alpha. *Nat Hum Behav.* 2018;2:168–171. doi: 10.1038/s41562-018-0311-x.

25. Maier M, Lakens D. Justify your alpha: A primer on two practical approaches. *Adv Methods Pract Psychol Sci.* 2022;5(2):25152459221080396. doi: 10.1177/25152459221080396.
26. Зорин НА. «Достоверность» или «статистическая значимость» – 12 лет спустя. *Педиатрическая фармакология.* 2011;8(5):13–19.
Zorin NA. [“Validity” or the “significance” – 12 years later]. *Pediatric Pharmacology.* 2011;8(5):13–19. Russian.

Mythologization of statistics in biomedical research

M.G. Slavenko^{1,2} • O.N. Mironenko¹ • M.S. Kuznetsov^{1,3} • E.A. Bakin¹

Reproducibility of scientific research has become a major concern in recent years. Inappropriate application of statistical analysis methods by researchers is considered one of the causes of the reproducibility crisis. Drawing on the concept of “statistical myths” – persistent, recurring distortions of statistical concepts – it is argued that these myths are systematic rather than accidental, serve specific functions, and reflect a deep methodological gap between the research domain and the practice of statistical analysis. The relationship between scientific and statistical inference is shown in the form of a biomedical research framework consisting of seven stages: conceptualization and formulation of the research question; operationalization; selection of the estimand; statistical model building and application; interpretation of the estimate accounting for uncertainty; clinical interpretation; and answering the research question. The framework is illustrated using examples from three real-world studies across various medical fields. Based on their functions, the authors classify statistical myths into four categories: epistemological, fundamental, instrumental, and terminological. The defining characteristics and

examples of each type are provided, alongside an analysis of several widespread myths. The proposed approach allows statistical myths to be viewed not merely as a pedagogical issue, but as indicators of deeper disruptions in research logic. From a methodological perspective, rather than memorizing yet another list of “correct” and “incorrect” statistical practices, researchers are encouraged to proceed systematically from the research question at hand to its statistical formalization.

Key words: biostatistics, biomedical research, methodology, myth, reproducibility, statistical inference, scientific inference, terminology

For citation: Slavenko MG, Mironenko ON, Kuznetsov MS, Bakin EA. Mythologization of statistics in biomedical research. *Almanac of Clinical Medicine.* 2026;54(2):104–112. doi: 10.18786/2072-0505-2026-54-010.

Received June 25, 2026; revised August 10, 2026; accepted for publication August 14, 2026; published online August 25, 2026

Conflict of interests

The authors declare no conflict of interests regarding the publication of this article.

Authors' contribution

All authors contributed equally to the development of the study concept, manuscript preparation and editing. All the authors have read and approved the final version of the manuscript, agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Use of artificial intelligence in manuscript preparation

The authors used ChatGPT (OpenAI) software as an auxiliary source for identifying relevant literature. The authors have reviewed and verified all outputs generated by the artificial intelligence and accept full responsibility for the content of the manuscript.

Acknowledgments

The authors thank all participants of the “Statistical MythBusters” lecture series, and in particular Natalia Zorina, Alexey Glazkov, Artemy Okhotin, Bogdan Sotnikov, Alisa Selezneva, Elena Chuklina and Elena Ubogoeva for fruitful discussions, criticism and valuable comments made during working on the contents of the article.

Matvei Slavenko – MSc (in Probability and Statistics), Lecturer¹; Biostatistician²; ORCID: <https://orcid.org/0000-0001-6610-9824>
✉ 2a Kantemirovskaya ul., Saint Petersburg, 197342, Russian Federation.
E-mail: matvei.slavenko@gmail.com

Olga N. Mironenko – PhD (in Economics), Lecturer¹; ORCID: <https://orcid.org/0000-0001-8952-8386>.
E-mail: freelomir@yandex.ru

Maksim S. Kuznetsov – MD, Lecturer¹; Senior Laboratory Technician, Department of Propaedeutics of Internal Diseases named after Professor S.S. Zimnitsky³; ORCID: <https://orcid.org/0000-0002-4891-9169>.
E-mail: bqmaks@gmail.com

Evgeny A. Bakin – PhD (in Engineering), Lecturer¹; ORCID: <https://orcid.org/0000-0002-5694-4348>.
E-mail: eugene.bakin@gmail.com

¹ Bioinformatics Institute; 2a Kantemirovskaya ul., Saint Petersburg, 197342, Russian Federation

² Motol and Homolka University Hospital; 84/1 V Úvalu, Prague 5, 150 00, Czech Republic

³ Kazan State Medical University; 49 Butlerova ul., Kazan, 420012, Russian Federation