

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/340581578>

Санкт-Петербург 2020

Chapter · April 2020

CITATIONS

0

READS

421

1 author:



[Nikita Khromov-Borisov](#)

First Pavlov State Medical University Of St Petersburg

102 PUBLICATIONS 286 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



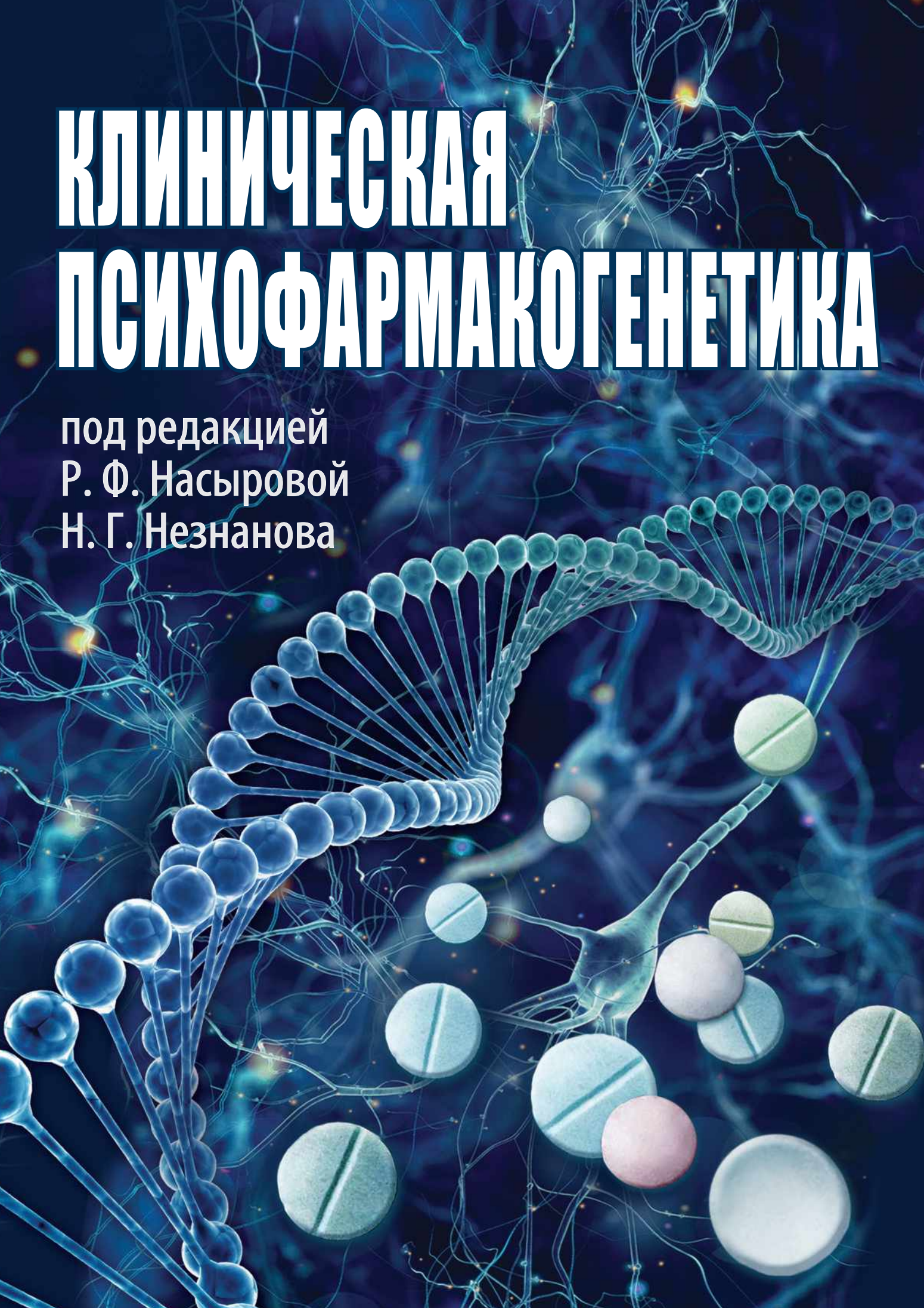
Biostatistics [View project](#)



Molecular Biology [View project](#)

КЛИНИЧЕСКАЯ ПСИХОФАРМАКОГЕНЕТИКА

под редакцией
Р. Ф. Насыровой
Н. Г. Незнанова



КЛИНИЧЕСКАЯ ПСИХОФАРМАКОГЕНЕТИКА

под редакцией
Р. Ф. Насыровой
Н. Г. Незнанова



Санкт-Петербург
2020

УДК 615.03:615.21
ББК 52.81:56.14

Рецензенты

Заслуженный деятель науки РФ, академик РАН *В. В. Новицкий*,
член-корреспондент РАН *Д. Ф. Хритинин*

Клиническая психофармакогенетика / Под ред. Р. Ф. Насыровой, Н. Г. Незнанова. — СПб: Издательство ДЕАН, 2020. — 408 с.

ISBN 978-5-6043573-7-8

Коллективная монография содержит современные данные по фармакогенетике психотропных препаратов. Систематизированы результаты исследований лекарственных средств, применяемых при лечении психических расстройств и болезней зависимости. Книга дает возможность ознакомиться с базовыми понятиями фармакогенетики, получить представление о современном состоянии данной науки в области психиатрии, наркологии и неврологии. Материал охватывает исторические, фармакоэкономические и практические аспекты психофармакогенетики.

Для врачей психиатров, наркологов, неврологов, клинических фармакологов, генетиков, аспирантов, а также студентов биологических и медицинских вузов.

УДК 615.03:615.21
ББК 52.81:56.14

© Регина Насырова, 2020
© Николай Незнанов, 2020
© Издательство ДЕАН, оформление, 2020

ISBN 978-5-6043573-7-8

Все права защищены. Никакая часть данной книги не может быть воспроизведена в какой бы то ни было форме без письменного разрешения владельцев авторских прав.

Кризис воспроизводимости результатов научных исследований, его истоки

Мировое научное сообщество взбудоражено многочисленными научными публикациями, неоспоримо свидетельствующими о плохой воспроизводимости результатов биомедицинских (и не только) исследований. Кризис воспроизводимости становится системной проблемой современной биомедицины. Вырисовывается причинно-следственная последовательность: неопределенность — непредсказуемость — невозпроизводимость.

В генетических и геномных исследованиях низкая воспроизводимость результатов во многом связана с их принципиальной непредсказуемостью, обусловленной неопределенностью взаимоотношений внутри триады генотип — среда — фенотип, в которой большую роль играют неконтролируемые нелинейные (неаддитивные) взаимодействия ген — среда и ген — ген, прояв-

ляющиеся в таких фундаментальных явлениях, как экспрессивность, пенетрантность, норма реакции, плейотропия.

Многие установленные локусы предрасположенности к болезням имеют тенденцию быть плейотропными.

В становление одного и того же признака могут быть вовлечены несколько генов, и в то же время один и тот же ген может влиять на становление многих признаков (рис. 1).

Один и тот же генотип может усиливать проявление одного признака, и одновременно он же способен ослаблять проявление другого, не менее жизненно важного признака. Например, один и тот же генотип может оказаться предрасполагающим к развитию данного заболевания, но одновременно он же может оказаться протективным в отношении другого заболевания (плейотропия). На рис. 2 представлена схема участия множества генов в развитии болезни Крона [1].

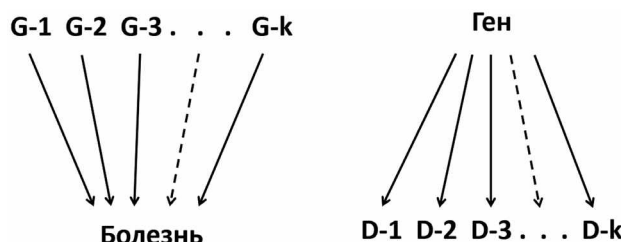


Рис. 1. Схема многообразных взаимоотношений между генами (G) и признаками (D)

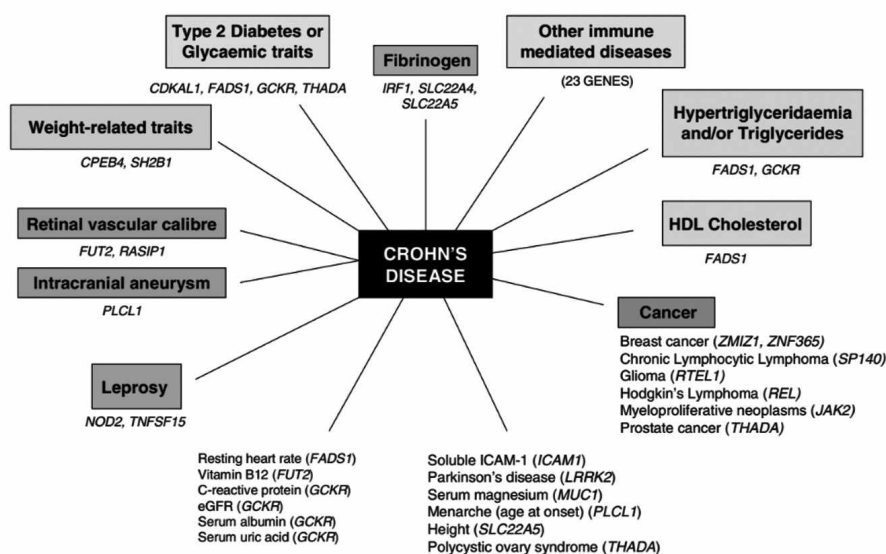


Рис. 2. Иллюстрация «избыточного» плейотропного действия генов, связанных с болезнью Крона [1]

Одним из основных источников неопределенности и низкой предсказательной способности в генетической предрасположенности следует признать плейотропию — часто один и тот же генотип оказывается предрасполагающим к различным заболеваниям и синдромам.

Очевидно, что плейотропия может быть *разнонаправленной*. Это, пожалуй, наиболее правдоподобное и едва ли не единственное объяснение высокой популяционной частоты многих, вроде бы несомненно предрасполагающих аллелей. Одна из аллелей какого-либо гена может оказаться предрасполагающей к одной болезни, однако другая аллель может оказаться предрасполагающей к другой болезни или оказывать противоположное действие в разных группах риска.

Например, ангиотензинпревращающий фермент (АПФ) не только является монитором артериального давления, но и участвует в процессе оплодотворения, в образовании иммунных клеток, в развитии атеросклероза. Его высокая экспрессия в макрофагах иммунных клеток предотвращает образование злокачественных опухолей. Поэтому использование ингибиторов АПФ может вызвать рак и болезнь Альцгеймера. Вследствие этого различные его аллели гена *ACE* проявляют широкое, зачастую противоположное плейотропное действие [2]. В частности, в гене *ACE* аллель *D* (делеция) считается предрасполагающей к сердечно-сосудистым заболеваниям, а аллель *I* (инсерция) оказывается предрасполагающей к болезни Альцгеймера. Делеция *GSTT1*0* считается предрасполагающей к различным формам рака у курящих, но защитной для некурящих. Кровопотери при травмах, родах, менструациях у гетерозигот по аллели *Leiden V* значительно ниже, чем у женщин без этой аллели.

Известная аллель гена С-С-рецептора хемокина 5 *CCR5*Δ32* может быть полезной для ее носителя при некоторых инфекциях (например, ВИЧ-1, возможно, при оспе), но вредна при других (например, клещевой энцефалит, вирус лихорадки Западного Нила). В гомозиготном состоянии *CCR5*Δ32* приводит к невозможности присоединения вируса ВИЧ к Т-клетке, но сильно увеличивает восприимчивость организма к лихорадке Западного Нила, проявляет побочные эффекты при целиакии и аутоиммунных заболеваниях щитовидной железы у пациентов с диабетом 1-го типа. Потеря *CCR5* оказывается благоприятной для рассеянного склероза, спонтанного клиренса вируса гепатита С и хронического и агрессивного периодонтита. Это объясняется тем, что *CCR5* явно участвует в иммунной системе человека, трудно оценить его потенциальные побочные

эффекты. Например, вариант *CCR5*Δ32* оказывает значительное влияние на ревматоидный артрит (РА), болезнь Стилла (БС), ишемическую болезнь сердца (ИБС), ИБС без реваскуляризации, стеноз позвоночника и бронхит. Примечательно, что среди этих шести заболеваний эффекты делеции *CCR5*Δ32* на аутоиммунные (РА и БС) и на остальные перечисленные заболевания имеют противоположные направления. Делеция *CCR5*Δ32* оказывается высоко плейотропной и потенциально способна увеличивать восприимчивость к другим заболеваниям, более распространенным, нежели ВИЧ. Создается впечатление, что делеция *CCR5*Δ32* приносит больше вреда, чем пользы.

Впечатляющими явились результаты недавней масштабной систематизации информации о генотипировании и регистре смерти 409 693 человек британского происхождения для исследования физических эффектов аллели *CCR5*Δ32*. По полученным оценкам генотип $\Delta32/\Delta32$ оказывает вредный эффект: показатель смертности от всех причин у лиц, гомозиготных по аллели $\Delta32$, составил 20% (рис. 3).

Особенно остро проблема разнонаправленной плейотропии встает в связи с попытками редактировать геном человека. «Исправление» аллели или генотипа, «вредных» («патогенных») в отношении одной болезни, может привести к непредсказуемым негативным последствиям в отношении других болезней [3].

В британском биобанке были собраны сведения о 409 693 людях британского происхождения в возрасте от 41 до 76 лет на предмет влияния *CCR5*Δ32* на уровень смертности. Различий между людьми, гетерозиготными по *CCR5*Δ32*, и дикого типа, обнаружено не было (обе эти группы окрашены светло-серым цветом). Напротив, люди, гомозиготные по *CCR5*Δ32* (темно-серый цвет), имели уровень смертности на 20% выше [4].

Известны гены — рекордсмены по плейотропии.

Ген VDR — ген-рецептор витамина D. Активность VDR простирается далеко за пределы метаболизма кальция и паратиреоидного гормона (ПТГ). Он участвует в транскрипции 900 генов, некоторые из которых являются ключевыми для здоровья, такие как MTSS1 (белок 1, супрессирующий метастазирование), а также ключевые компоненты врожденного иммунитета (антимикробный пептид кателицидин, бета-дефензины, TLR2 — толл-подобный рецептор и др.). Роль VDR во врожденном иммунитете уникальна для человека. Никакое другое модельное животное, в частности, мышь, не выработало эволюционно такую функцию для этого рецептора.

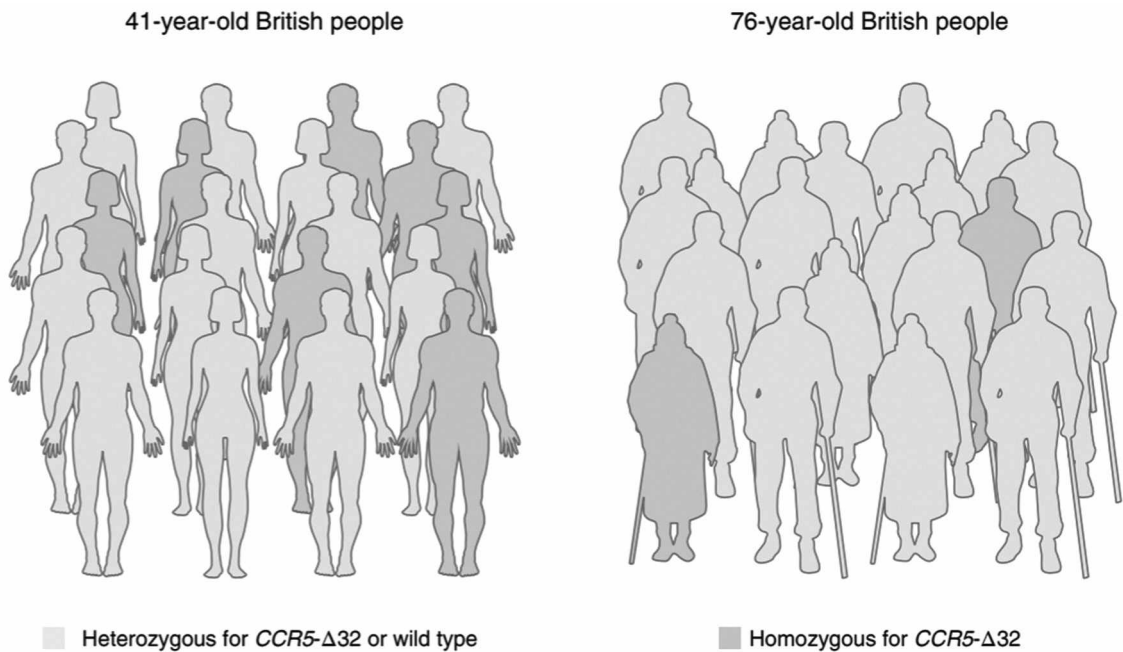


Рис. 3. Повышенная смертность среди людей, гомозиготных по *CCR5*Δ32*

Ген рецептора эстрогена (*ESR*). Рецептор эстрогена прямо или опосредованно отвечает за экспрессию 6000 генов, или 26% генома человека (рис. 4) [5].

Этими коллизиями, очевидно, обусловлено то, что многие аллели и генотипы, кажущиеся «болезнетворными», не элиминируются естественным отбором из популяций. Впору говорить о принципиально непреодолимом принципе неопределенности в биомедицине [6].

В результате все чаще появляются консенсусные заявления «авторитетных» исследователей, подобные тому, с которым выступили 24 специалиста по спортивной генетике: «Генетические тесты не играют никакой роли при выявлении талантов или при разработке индивидуальных тренировочных режимов для достижения максимальных спортивных достижений. Независимыми исследованиями выявлены проблемы, связанные с контролем качества лабораторных процедур и с противоречи-

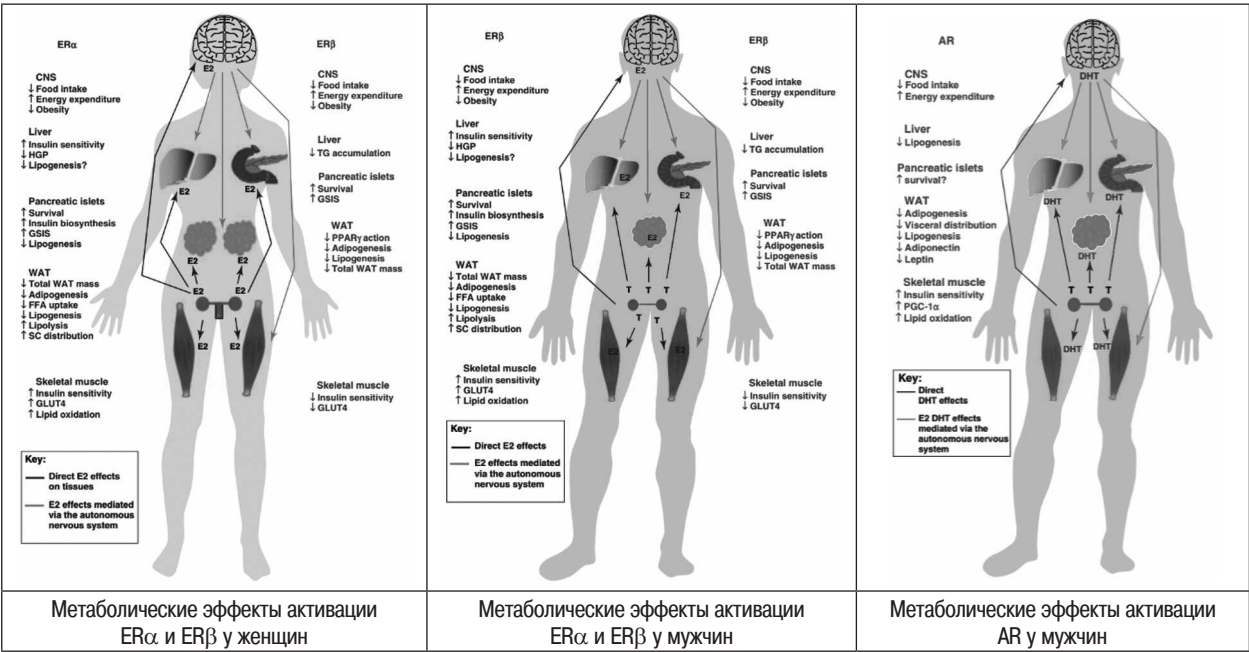


Рис. 4. Множественные метаболические эффекты активации рецепторов эстрогена у мужчин и женщин и андрогена у мужчин [5]

Авторы еще одного недавнего обзора предупреждают: «Многие фармакогенетические алгоритмы дозирования разработаны для целевого МНО 2–3 и поэтому их полезность для оценки терапевтических доз варфарина с другими целевыми диапазонами МНО является неопределенной. Дозирование варфарина под фармакогенетическим контролем не меняет требований к регулярному мониторингу МНО. Есть пациенты, для которых генетическое тестирование, вероятно, будет мало или вообще не принесет пользы, в том числе те, кто уже проходил длительное лечение стабильными дозами варфарина, и те, кто не может добиться стабильной дозировки из-за различной приверженности. Наибольшая потенциальная выгода наступает в начале курса терапии (до начала терапии или в первые дни терапии). Вполне вероятно, что пациенты, получающие терапию в течение многих недель или месяцев при тщательном мониторинге МНО, не получают большой пользы от последующего тестирования фармакогенетики варфарина» [17]. На эту неопределенность накладывается еще тот факт, что несмотря на международные усилия по калибровке МНО, все еще сохраняются статистически значимые различия между различными наборами, что ставит под сомнение пригодность МНО при долгосрочной антикоагулянтной терапии.

Недавно исследователи оценили степень согласованности между четырьмя коммерческими инструментами поддержки принятия решений (ИППР) на основе фармакогенетики. Были исследованы пять амбулаторных пациентов с серьезным депрессивным расстройством и по крайней мере с двумя предыдущими неудачами с антидепрессантами. Сравнения были сделаны по семи фармакокинетическим (*CYP1A2*, *CYP2B6*, *CYP2C19*, *CYP2C9*, *CYP2D6*, *CYP3A4* и *UGT2B15*) и семи фармакодинамическим (*BDNF*, *COMT*, *HLA-A*, *HTR2A*, *HTR2C*, *OPRM1* и *SLC6A4*) генам, которые были включены не менее чем в две из четырех панелей тестирования ИППР. Среди этих перекрывающихся наборов генов согласие генотипа (33–100%) и прогнозируемого фенотипа (20–100%) существенно различалось. Согласованность с рекомендациями по лечению была наибольшей для стабилизаторов настроения (84%), затем следовали антидепрессанты (56%), анксиолитики/снотворные (56%) и антипсихотики (55%). Примерно одна четверть (26%) всех рекомендаций по лечению были совместимы с двумя или более ИППР как «действенные», но 19% этих рекомендаций для одного и того же лекарства были противоречивыми (например, дозирование). Авторы нашли незначительное соответствие между рекомендациями по ЛС, предоставляемым коммерческими

ИППР на основе фармакокинетики, предположительно в результате различий в тестируемых генах/вариантах, стратегиях генотипирования и алгоритмах, используемых для прогнозирования взаимодействий между ЛС и генами. Наблюдаемый уровень разногласий в рекомендациях по лечению среди ИППР указывает на то, что эти инструменты нельзя считать эквивалентными, и заставляет предположить, что они могут различаться по своей клинической применимости. Тем не менее авторы полагают, что дальнейшая разработка профессиональных руководств по использованию ИППР в клинической практике является оправданной, но потребует проверки в клинических испытаниях [18].

Пути преодоления кризиса воспроизводимости

Давно известно, что большинство результатов исследования генетических предрасположенностей не воспроизводятся. См., например, Editorial (1999) [19]. Тогда же были ужесточены требования к работам, представляемым к публикации:

- 1) объемы выборок должны быть большими;
- 2) *p*-значения должны быть малыми;
- 3) должны быть представлены результаты независимых повторностей;
- 4) ассоциации должны быть изучены как в семьях, так и в популяциях;
- 5) значения отношения шансов (*OR* — odds ratios) и (или) дополнительного (добавочного, привносимого) риска (*AR* — attributable risk) должны быть высокими;
- 6) для выявленных ассоциаций должны быть предложены разумные биологические и физиологические объяснения.

Одной из основных причин кризиса воспроизводимости является применение неадекватных методов статистического анализа данных. Поэтому следует тщательно и обоснованно выбирать статистические методы. В частности, становится понятным, что нельзя ограничиваться сообщением исключительно так называемых *P*-значений.

В биомедицинских исследованиях (как у нас, так и за рубежом) широко распространены как минимум три чреватых негативными последствиями заблуждения, которые недопустимы при современном состоянии биомедицинской и биостатистической науки.

1. Излишне часто авторы ограничиваются лишь сообщениями о статистической значимости наблюдаемых эффектов, ориентируясь при этом почти исключительно на получаемые *p*-значения.

2. В качестве порогового значения для уровня значимости α в подавляющем большинстве случаев авторы ограничиваются исключительно са-

мым снисходительным значением $\alpha = 0,05$ (ибо «все так делают»).

3. Часто неосознанно статистическая значимость наблюдаемых эффектов отождествляется с их практической (клинической) важностью и ценностью. Другими словами, исподволь p -значение интерпретируется как вероятность проверяемой нулевой гипотезы H_0 об отсутствии эффекта.

Синдром статистической снисходительности

Распространенным заблуждением является интерпретация p -значения как вероятности нулевой гипотезы H_0 . На самом деле p -значения ничего не говорят непосредственно ни о вероятности проверяемой статистической нулевой гипотезы H_0 , ни о размере оцениваемого эффекта, ни о вероятности его воспроизведения. Поэтому давно осознано и признано, что констатации статистической значимости наблюдаемых эффектов (основанной в основном на преодолении критических значений уровня α) совершенно недостаточно для объявления этих эффектов клинически (практически) важными и ценными. Давно рекомендуется подход «статистика с доверием», т. е. обязательно для измеряемых эффектов следует представлять результаты интервальных оценок (доверительных интервалов — ДИ). Более того, настоятельно рекомендуется оценивать достигнутую статистическую мощность (Pow) исследования и объемы выборок, необходимые для достижения заданной мощности при заранее выбранном уровне значимости.

Можно сказать, что современная экспериментальная биомедицина страдает «синдромом статистической снисходительности и доверчивости». В частности, становится понятным, что нельзя ограничиваться сообщением исключительно p -значений. На наших глазах в современной биостатистике разрабатывается подход, который можно назвать «гармонизация статистических доказательств и предсказаний». В нем методы традиционной (ортодоксальной, частотной) статистики сочетаются с методами бейзовской¹ статистики. В результате современная биостатистика становится способной не только получать статистические доказательства, но и прогнозировать (предсказывать, моделировать) воспроизводимость результатов повторных буду-

щих исследований (см., например, Хромов-Борисов, 2016 [20]).

На заре зарождения биостатистики (биометрии) исследователи были более требовательны к надежности и убедительности результатов анализа статистических данных и выводов из него. Популярно говоря, они руководствовались «правилом трех сигм». Согласно ему вероятность значений нормально распределенной случайной величины за пределами плюс-минус три стандартных отклонения ($\pm 3\sigma$) равна 0,0027, или округленно 0,003, и такие значения интерпретировались как *практически невозможные*. Однако не следует забывать, что в более общем случае для любого распределения (с конечным средним значением и дисперсией), согласно неравенству Чебышёва, вероятность отклонения от ожидаемого за пределы $\pm 3\sigma$ может достигать 0,11 и даже за пределами плюс-минус 5 сигм ($\pm 5\sigma$) могут находиться до 0,04 значений (см. например, Королев, 2008 [21]).

В подавляющем большинстве руководств по статистике для уровня значимости α указываются три сакраментальных значения: 0,05, 0,01 и 0,001. Однако многие исследователи (и у нас, и за рубежом) довольствуются преодолением уровня 0,05 и объявляют наблюдаемый эффект достоверным. Предпочитая преодоление порога 0,05, исследователи низвели свою нетребовательность до «правила двух сигм». Такое явление можно назвать «синдромом статистической снисходительности».

Калибровка p -значений

Интуитивно наиболее понятной является бейзовская интерпретация p -значения, согласно которой половина p -значения ($p/2$) есть доля повторных экспериментов, в которых будет наблюдаться направление эффекта, противоположное наблюдаемому в первом эксперименте (и не более того). Или $(1 - p/2)$ есть всего лишь предсказательная вероятность получить эффект с тем же знаком (в том же направлении) при повторении эксперимента с бесконечно большим набором данных. Можно понять, что p -значение говорит нам лишь о знаке эффекта, о его направлении, но не о размере эффекта.

Понять p -значение можно также следующим образом. Если дополнение к нему $(1 - p)$ использовать в качестве доверительной вероятности, то полученный $100(1 - p)\%$ -ный ДИ будет одним из своих концов касаться безразличного значения.

Обычно p -значение интерпретируют как свидетельство против нулевой гипотезы. Считается, что чем оно меньше, тем больше основание сомневаться в нулевой гипотезе, и отсюда

¹ В отечественной литературе нет единого мнения относительно транскрипции английской фамилии Bayes. Чаще всего его пишут Байес, однако в отечественных математических энциклопедиях и в словарях английских фамилий указано Бейес. У нас общепринята фонетическая транскрипция. Bayes произносится как Бейз [beiz], поэтому следует настоятельно рекомендовать именно эту транскрипцию, она и звучит благозвучнее.

возникает соблазн интерпретировать p -значение как вероятность нулевой гипотезы. На самом деле соотношение между p -значением и вероятностью нулевой гипотезы далеко не столь однозначно.

В самом общем виде соотношения между p -значениями и вероятностью нулевой гипотезы можно продемонстрировать в виде так называемой калибровки p -значений, практически важные значения которой представлены в табл. 1.

Таблица 1
Калибровка p -значений и вербальная шкала для них

p_{obs} — наблюдаемое p -значение	Нижняя граница для $P_L(H_0 p_{\text{obs}})$ при p_{obs}	Верхняя граница для $P^U(H_1 p_{\text{obs}})$ при p_{obs}	Вербальная шкала статистической значимости
0,05	> 0,29	< 0,71	Ничтожная
0,01	> 0,11	< 0,89	
0,005	> 0,067	< 0,933	Низкая
0,001	> 0,018	< 0,982	Средняя
0,0001	> 0,0025	< 0,9975	Высокая
0,00001	> 0,00031	< 0,99969	

Примечание. $P_L(H_0|p_{\text{obs}})$ — нижняя граница для апостериорной вероятности нулевой гипотезы H_0 при наблюдаемом значении p_{obs} . Соответственно, $P^U(H_1|p_{\text{obs}})$ — верхняя граница для апостериорной вероятности альтернативной гипотезы H_1 при наблюдаемом значении p_{obs} . Предполагается, что *a priori* обе гипотезы равновероятны. Жирным шрифтом выделены значения вероятностей, которые общепринято считать достаточно малыми (или достаточно большими), чтобы сомневаться в отсутствии эффекта. См., например, GraphPadSoftware [22] со смещением на одну строку согласно современным рекомендациям [23, 24]. Для калибровки других p -значений можно использовать интерактивную веб-страницу [25, 26].

Можно видеть, что при $p_{\text{obs}} = 0,05$ апостериорная вероятность нулевой гипотезы превышает 29%. Поэтому, получив $p_{\text{obs}} = 0,05$, теперь вряд ли кто осмелится признать его весомым аргументом против гипотезы H_0 . Только при p_{obs} , исчисляемом тысячными долями, апостериорная вероятность $P(H_0|p_{\text{obs}})$ становится приемлемо малой для отклонения гипотезы H_0 (выделены жирным шрифтом в табл. 1). Так, при $p_{\text{obs}} = 0,001$ можно ожидать, что апостериорная вероятность $P(H_0|p_{\text{obs}})$ может достигнуть вполне убедительного малого значения 1,8%, но не меньше.

Еще в 1982 году в учебнике «Биометрия», удостоенном премии Ленгосуниверситета, на основе знаний того времени было предложено считать результат проверки нулевой гипотезы неопределенным, когда p -значение попадает в диапазон $0,01 < p < 0,05$ («серая зона») [27]. Сейчас этот совет подхвачен и обоснован в контексте бейзовского подхода в статистике: «Следует в обязательном порядке вычислять фактор Бейза при попадании p -значений в «серую зону» (0,01–0,05)» [23].

Гармонизация статистических доказательств и предсказаний

Половина p -значения ($p/2$) есть доля повторных экспериментов, в которых будет наблюдаться направление эффекта, противоположное

наблюдаемому в первом эксперименте. Когда эффект отсутствует, то при многократных повторениях эксперимента отклонения от безразличного значения будут происходить равновероятно в противоположных направлениях. В этом случае $p/2 = 0,5$.

Чем меньше p -значение, т. е. чем реже результаты будут отклоняться от безразличного значения в одном из двух направлений, тем убедительнее наличие противоположного эффекта.

В настоящее время имеется возможность предсказывать вероятности воспроизведения наблюдаемых p -значений (табл. 2).

Таблица 2
Воспроизводимость и предсказательность p -значений

p -значение:	p_{exp} — ожидаемое (желаемое)				
	0,05	0,01	0,005	0,001	10^{-4}
p_{obs} — наблюдаемое	P_{rep} — вероятность воспроизведения p_{obs}				
0,05	0,50	0,33	0,27	0,17	0,086
0,01	0,67	0,50	0,44	0,31	0,18
0,005	0,73	0,57	0,50	0,37	0,22
0,001	0,83	0,69	0,63	0,50	0,34
10^{-4}	0,91	0,82	0,78	0,66	0,50

Жирным шрифтом выделены значения вероятностей, которые можно считать убедительно высокими. Закрашены равновероятные значения.

Как видим, вероятность того, что при повторении эксперимента воспроизведется наблюдаемое p -значение, равна 50%. Чтобы достигнуть высокой вероятности для желаемого p -значения, исходное наблюдаемое p -значение должно быть существенно (фактически на порядок) меньше. Например, можно ожидать, что при наблюдаемом $p_{\text{obs}} = 0,001$ с вероятностью 83% будет получено $p_{\text{exp}} = 0,05$. Новейшие публикации по этим проблемам см. [23, 28–33]. Эти авторы разработали полезные программы для вычисления таких вероятностей: [24, 34–38].

Размер эффекта

Понятие «размер эффекта» является принципиально важным в современной статистике. Для анализа мерных данных наиболее практичным является стандартизированный размер эффекта по Коэну d_c . Он измеряет разность средних в единицах сводного стандартного отклонения и удобен для содержательной интерпретации наблюдаемых различий. Для стандартизированного эффекта d_c по Коэну предложены вербальные шкалы, которые представлены в табл. 3а и 3б.

Таблица 3а

Вербальные шкалы для стандартизированного эффекта d_c по Коэну

Эффект (различие)	d_c	d_c	d_c
Ничтожный	0–0,1	0–0,2	0–0,2
Очень малый	0,1–0,2		
Малый	0,2–0,5	0,2–0,6	
Умеренный	0,5–0,8	0,6–1,2	0,2–0,5
Большой	0,8–1,2	1,2–2,0	0,5–0,8
Очень большой	1,2–2,0	2,0–4,0	> 0,8
Несомненный	> 2,0	> 4,0	
Ссылки	[39]	[40]	[41]

Таблица 3б

Сводная пользовательски-дружественная, практически удобная вербальная шкала для интерпретации стандартизированного размера эффекта по Коэну d_c

d_c	Интерпретация эффекта
0–0,5	Пренебрежимо малый, ничтожный, не заслуживающий внимания, практически бесполезный
0,5–1,0	Малый, слабый
1,0–1,5	Средний, умеренный
1,5–2,0	Большой, сильный
2,0–3,0	Очень большой, очень сильный
> 3,0	Исключительно, чрезвычайно, невероятно большой, сильный

Как уже было сказано, без интервальной оценки любая интерпретация размера эффекта бессмысленна. Указанным диапазонам той или иной градации должны удовлетворять полученные значения его нижней доверительной границы.

Понятие «размер эффекта» настолько важно для анализа научных данных, что за рубежом его уже обсуждают и пропагандируют научные журналисты и блогеры (Reichel C., 2019) [42].

Статистическая мощность исследования

В биомедицине уже давно рекомендуется проводить анализ статистической мощности исследования (см. классический труд: Cohen, 1998 [41]). Мощность исследования (Pow) есть доля повторных экспериментов, в которых будут наблюдаться статистически значимые эффекты (данного размера и при одном и том же фиксированном уровне α). Мощность можно интерпретировать также как чувствительность статистического метода (критерия) к распознаванию реального эффекта.

Далее в таблицах 4, 5 и 6 представлены вербальные шкалы для показателей, которые широко используются в качестве мер эффектов для счетных данных (качественных признаков), мер связи, мер корреляции и мер согласия.

Таблица 4

Вербальные шкалы для коэффициентов сопряженности (φ Пирсона и V Крамера)

Сила связи	φ или V	φ или V	φ или V
Ничтожная	0–0,1	0–0,2	0–0,1
Слабая	0,1–0,2	0,2–0,3	0,1–0,3
Умеренная	0,2–0,4	0,3–0,4	0,3–0,5
Относительно сильная	0,4–0,6	0,4–0,7	> 0,5
Сильная	0,6–0,8		
Очень сильная	0,8–1,0	> 0,7	
Ссылки	[43]	[44]	[45]

Таблица 5

Вербальные шкалы для мер согласованности (каппа Коэна κ) и корреляции (коэффициенты корреляции Пирсона r или Спирмена r_s)

Согласованность или корреляция	κ , r или r_s	κ , r или r_s	κ , r или r_s
Низкая, слабая	0,0–0,2	0,0–0,4	0,0–0,4
Посредственная	0,2–0,4		
Умеренная	0,4–0,6	0,4–0,6	0,4–0,75
Хорошая	0,6–0,8	0,6–0,75	
Отличная	0,8–1,0	0,75–1,0	0,75–1,0
Ссылки	[46, 47]		

Понятно, что для отрицательных корреляций градация остается такой же. Только в случае с коэффициентом согласованности κ Козна, когда $\kappa = -1$, это означает полную рассогласованность между мнениями экспертов: если один говорит «белое», другой говорит «черное».

Таблица 6

Вербальные шкалы для отношения шансов за/против (*OR*)

Сила связи	<i>OR</i>	<i>OR</i>	<i>OR</i>	<i>OR</i>	<i>OR</i>
Ничтожная	1–1,5	1–1,5	1–1,5	1–2,2	1–5
Слабая	1,5–3,5	1,5–3,5	1,5–2,5		
Умеренная	3,5–9	2,5–9	2,5–4,3	2,2–5,4	> 5
Сильная	9–32	> 9	> 4,3	> 5,4	
Очень сильная	32–360				
Сверхсильная	> 360				
Источник	[40]	[45]	[48]	[49]	[50]

Можно видеть, что при словесной интерпретации показателей статистической связи у комментаторов-интерпретаторов нет единогласия (консенсуса). Тем не менее все согласны, что значения при $OR < 1,5$ не заслуживают внимания, а значения в диапазоне $1,5 < OR < 3,5$ большинство интерпретируют как слабую связь. Опять-таки указанным границам должны удовлетворять значения нижней границы ДИ для *OR*.

«Отношение шансов» (*OR*) не имеет вероятностной интерпретации

Шансы не являются вероятностями. Это частое заблуждение обусловлено в первую очередь тем, что два английских слова *chance* и *odds* на русский язык переводятся одним словом «шансы». Но *chance* выступает синонимом слова «вероятность» (величина, к которой сводится частота осуществления случайного события *A*), а *odds* есть отношение двух вероятностей: вероятности осуществления случайного события *A* к вероятности его неосуществления: $Odds(A) = P(A) / [1 - P(A)]$. Чтобы избежать недоразумения и недопонимания, наверное, как минимум, следует использовать уточняющее выражение «шансы за/против» события *A* либо ввести в обиход кальку «одды», или «проконы» (от *pro* et *contra*) или (более по-русски) «запротивы».

Однако есть возможность избежать использования двусмысленного и приводящего к недоразумениям слова «шансы» при интерпретации показателя *OR*. Интуитивно более понятным становится его представление в виде отношения двух относительных рисков: $OR = RR[+] / RR[-]$. Чис-

литель $RR[+]$ здесь есть относительный риск для положительных исходов, а знаменатель $RR[-]$ для отрицательных (ср.: Норре et al., 2017 [51]). *OR* является единым интегральным показателем, который объединяет и сравнивает оба относительных риска. Однако практически более информативным является рассмотрение $RR[+]$ и $RR[-]$ по отдельности. Такой подход давно и успешно практикуется при решении задач контроля качества диагностических тестов с бинарными исходами (см., например, Тишков и др., 2013 [52]).

При $OR < 2,2$ маркер обладает заведомо низкой прогностической эффективностью во всех смыслах и при любых частотах встречаемости заболевания и маркера. Маркер может быть хорошим классификатором, если $OR > 5,4$, при условии, что его популяционная частота достаточно высока: $p_M > 0,3$. На практике это означает, что указанным неравенствам должны удовлетворять нижние границы $100(1 - \alpha)\%$ -ного доверительного интервала для оцениваемого значения OR_L , т. е. $OR_L < 2,2$ в первом случае и $OR_L > 5,4$ во втором случае [49]. Ранее близкие значения критических уровней наблюдаемых эффектов предлагались и для относительных рисков ($RR < 2$ и $RR > 5$) [53] (табл. 7).

Таблица 7

Вербальные шкалы для относительных рисков (*RR*)

Сила связи	<i>RR</i>	<i>RR</i>	<i>RR</i>
Ничтожная	1–1,5	1–1,2	< 2
Слабая	1,5–2	1,2–2	
Умеренная	2–3	2–5	2–5
Сильная	3–4	> 5	> 5
Очень сильная	> 4		
Ссылки	[54]	[53]	

В качестве иллюстрации можно привести значения *RR* для маркеров, ассоциированных с раком молочной железы (табл. 8)

Таблица 8

Высоко и умеренно пенетрантные гены предрасположенности к раку молочной железы

Ген	Варианты (аллели)	<i>RR</i>	Population Frequency (%)
<i>BRCA1</i>	Множественные мутации	> 10	0,1
<i>BRCA2</i>	Множественные мутации	> 10	0,1
<i>TP53</i>	Множественные мутации	> 10	< 0,1
<i>PTEN</i>	Множественные мутации	> 10	< 0,1
<i>ATM</i>	Укорачивающие и миссенс-мутации	2–4	< 0,5
<i>CHEK2</i>	1100delC	2–5	0,7

Ген	Варианты (аллели)	RR	Population Frequency (%)
BRIP1	Укорачивающие мутации	2–3	0,1
PALB2	Укорачивающие мутации	2–5	< 0,1

Первые четыре маркера, несомненно, заслуживают практического применения и применяются. Для сравнения можно вспомнить, как на протяжении многих лет, фактически с момента открытия групп крови, исследователи увлекались поисками их связи с различными заболеваниями. Все получаемые значения *RR* или *OR* не превышали значений 1,5, и все они не получили практического применения, к тому же результаты не воспроизводились и сейчас благополучно забыты (хотя и остались в учебниках). Основная причина: исследователи в те времена не проверяли согласие распределений частот генотипов с равновесием Харди–Вайнберга, отклонения от которого во всех случаях кажущихся ассоциаций с болезнями при ближайшем рассмотрении оказались статистически высокосignификантными [55].

Один из примеров — масштабное когортное исследование (более миллиона пациентов) по оценке риска рака ЖКТ и пептических язв в зависимости от генотипов системы групп кро-

ви АВ0 [56]. Проверка согласия с равновесием Харди–Вайнберга показала статистически высокосignификантное отклонение (рис. 7). Полученному значению статистики критерия $\chi^2 = 376$ при одной степени свободы ($df = 1$) соответствует заведомо малое значение $p = 10^{-83}$. Основной вклад в наблюдаемое отклонение внес избыток генотипов АВ. Нередко причиной таких больших отклонений от равновесия Харди–Вайнберга является процедура сбора данных на основании плохо контролируемого опроса пациентов.

Бейзов фактор

Бейзов фактор *BF* является более адекватным, нежели *p*-значение, показателем для сравнения правдоподобности той или другой гипотезы. Он принципиально отличается от значения P_{val} . Бейзов фактор не является вероятностью сам по себе, а является отношением вероятностей (или плотностей распределений), и он может варьироваться от нуля до бесконечности:

$$BF_{01} = P(D_{obs}|H_0) / P(D_{obs}|H_1);$$

$$BF_{10} = P(D_{obs}|H_1) / P(D_{obs}|H_0).$$

Здесь D_{obs} — наблюдаемые данные.

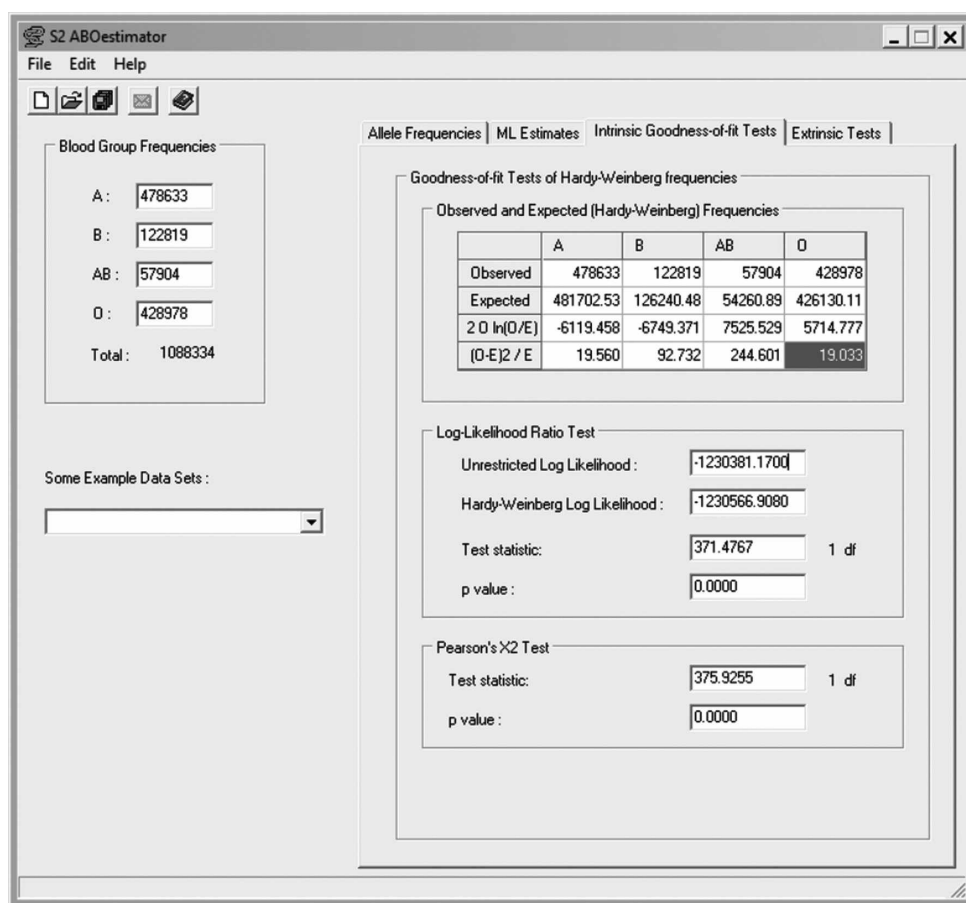


Рис. 7. Проверка согласия с РХВ для частот генотипов системы групп крови АВ0 по данным Edgrenetal., 2010 [56]. Использована программа S2 ABOestimator [57]

Это означает, что с помощью бейзового фактора проводится не однократная проверка значимости лишь нулевой гипотезы, а сравниваются вероятности получить наблюдаемые данные при обеих гипотезах. Для его значений также предложены словесные интерпретации (табл. 9а и 9б).

Таблица 9а

Вербальные шкалы для бейзовых факторов BF_{01} и BF_{10}

Свидетельство в пользу гипотезы H_0 против гипотезы H_1	Авторы				
	Elston, Johnson	Jeffreys	Jackman	Kass, Raftery	Royall
Едва стоит упоминания	> 1–2	> 1–3	> 1–3	> 1–3	> 1–8
Умеренное	2–6	3–10	3–12	3–20	8–32
Сильное	6–20	10–30	12–150	20–150	> 32
Очень сильное	20–60	30–100		> 150	
Убедительное	60–100	> 100	> 150	> 150	

В настоящее время наиболее общепринята следующая шкала:

Таблица 9б

Сводная консенсусная вербальная шкала для интерпретации бейзовых факторов

BF_{01}	Свидетельство в пользу гипотезы H_0 против гипотезы H_1
> 100	Убедительное
30–100	Очень сильное
10–30	Сильное
3–10	Умеренное (слабое)
1–3	Пренебрежимо малое
BF_{10}	Свидетельство в пользу гипотезы H_1 против гипотезы H_0

Программное обеспечение анализа генетических данных

Основными программами для анализа генетико-популяционных данных следует признать три: GenAlEx, GENEPOP и Arlequin. Есть еще полезная программа Power Marker, но у нее истек срок действия и приходится переустанавливать на компьютере календарную дату на май 2006 года [58]. Сравнение возможностей этих программ представлено недавно [59].

Серьезным неудобством в использовании этих и других программ является тот факт, что разные программы используют разные форматы для исходных входных файлов. Одна опечатка в использовании запятой, точки с запятой, пробела, слэша и т. п. может сделать данные нечитаемыми или неправильно классифицированными. В результате для организации данных в правильном формате часто требуется больше времени, чем для проведения самого анализа. К счастью, некоторые программы обеспечивают возможность импорта/экспорта данных из/в другие форматы, что позволяет избежать переформатирования данных вручную и облегчает и ускоряет набор данных для разных программ. Это важно, поскольку нет универсальной программы, в разных программах — разный набор инструментов анализа и для решения разных задач приходится лавировать между программами. Поэтому анализ таких данных удобно начать с программы GenAlEx [60].

Ее небольшое неудобство состоит в том, что эта программа не воспринимает буквенные обозначения генотипов, и приходится их шифровать: A = 1, C = 2, G = 3 и T = 4.

На рис. 8 представлен образец входного (input) файла.

	A	B	C	D	E	F	G	H
15	3029	2	2016	1013				
NWRus	NWRus3029	Data Sub	NWRus1	NWRus2				
sample	Pop	D8S1179		D21S11		D7S820		CSF1PO
1	NWRus1		12	14	29	29	10	10
2	NWRus1		11	12	29	302	11	13
3	NWRus1		13	14	31	322	11	12
4	NWRus1		12	14	29	29	8	13
5	NWRus1		10	15	302	322	10	10
6	NWRus1		13	15	30	322	10	11
7	NWRus1		12	14	29	29	8	13
8	NWRus1		10	15	302	322	10	10
9	NWRus1		13	15	30	322	10	11
10	NWRus1		8	15	29	302	11	11
11	NWRus1		12	14	28	29	8	11
12	NWRus1		12	13	29	30	9	10

Рис. 8. Образец входного файла для программы GenAlEx

Обозначения: A1 — число локусов; B1 — общий объем выборки; C2 — число популяций; D1–E1 — объем каждой из выборок. Представлен фрагмент данных по локусам STR, используемым в судебной генетике

Данные в этой программе легко конвертировать во входные файлы для программ GENEPOP и Arlequin и других 27 программ. GenAlEx производит высококачественные диаграммы (рис. 9).

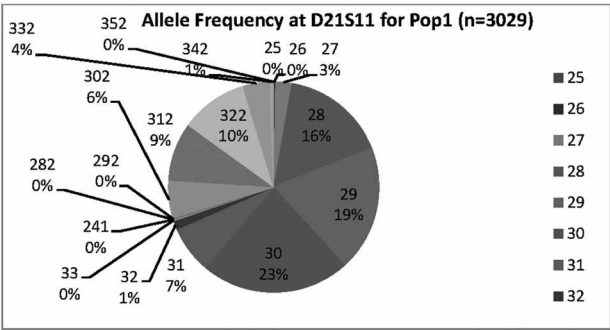


Рис. 9. Графические возможности GenAlEx

Подбор референсных групп

Оценка популяционных частот аллелей, гаплотипов и генотипов — важный этап в молекулярно-генетических исследованиях. Для получения надежных оценок таких частот необходимо иметь выборки большого объема, которые зачастую бывает затруднительно получить силами одной лаборатории. Однако возможно набрать данные из различных источников, и если они статистически однородны, то их можно суммировать. Примером такого подхода может служить оценка эталонных частот аллелей в локусах *HLA-DQA1* и *HLA-DQB1*, связанных с предрасположенностью к целиакии [61]. В этой работе были собраны данные о результатах HLA-

генотипирования в выборках относительно небольшого объема (от 70 до 290 человек) из шести различных российских местностей (Санкт-Петербург, Москва, Северо-Западный округ, Вологда, Западная Сибирь, Архангельск). Данные и результаты их статистического анализа представлены в табл. 10. Их визуализация представлена на рис. 10.

Результаты традиционной (частотной) проверки однородности этих данных с использованием точного критерия Фишера–Фримена–Холтона, казалось бы, свидетельствуют об их статистически высокосignificant неоднородности (гетерогенности): $p = 8 \cdot 10^{-5}$ для локуса *HLA-DQA1* и $p = 10^{-9}$ для *HLA-DQB1*. Соответственно, их вроде бы нельзя объединять. Однако значения бейзовых факторов BF_{01} свидетельствуют о противоположном. Получить эти данные в 10^{17} и в 10^{18} раз более правдоподобно при нулевой гипотезе H_0 об их статистической однородности, чем при альтернативной гипотезе H_1 об их неоднородности. Одновременно эти значения BF_{01} показывают, во сколько раз апостериорные шансы в пользу гипотезы H_0 против гипотезы H_1 превышают их априорные шансы. А p -значения не позволяют делать подобные выводы.

Наблюдаемое кардинальное противоречие между выводами на основе p -значений и бейзовых факторов демонстрирует то, что называют парадоксом Джеффриса–Линдли, а именно: вероятность получить наблюдаемые данные при нулевой гипотезе может быть мала (на этой вероятности зиждется вычисление p -значений), но она может быть еще меньше при альтернативной гипотезе.

Таблица 10

Распределение численностей аллелей двух локусов *HLA-DQA1* и *HLA-DQB1* в выборках здоровых детей из шести различных российских местностей и оценка их эталонных частот [61]

Аллель	Численность аллелей в выборках								«Эталонная» частота аллели с 99%-ными ДИ
	СПб и Москва	Сев.-Зап.	Вологда	Москва_1	Зап.-Сиб.	Москва_2	Арх.	Всего	
Локус <i>HLA-DQA1</i> *									
101	39	31	68	76	35	37	19	305	₁₂ 14 ₁₇ %
102	41	37	90	114	44	52	30	408	₁₇ 19 ₂₂ %
103	14	18	42	64	23	16	9	186	₇ 9 ₁₁ %
201	26	21	72	87	35	38	19	298	₁₂ 14 ₁₇ %
301	19	33	84	58	19	37	33	283	₁₁ 13 ₁₆ %
401	5	9	24	9	13	4	5	69	₂ 3 ₅ %
501	55	53	102	172	46	93	29	550	₂₃ 26 ₂₉ %
601	1	0	2	2	1	2	1	9	_{0,1} 0,4 _{1,1} %
Всего	200	202	484	582	216	279	145	2108	100%
<i>p</i>	8 · 10 ⁻⁵								
<i>BF</i> ₀₁	4 · 10 ¹⁷								
<i>V</i>	0,04 0,08 _{0,12}								

Аллель	Численность аллелей в выборках								«Эталонная» частота аллели с 99%-ными ДИ
	СПб и Москва	Сев.-Зап.	Вологда	Москва_1	Зап.-Сиб.	Москва_2	Арх.	Всего	
Локус <i>HLA-DQB1</i> *									
201	33	35	88	112	46	52	24	366	¹⁶ ₂₂ 19 %
301	45	40	76	140	41	72	18	414	¹⁸ ₂₄ 21 %
302	12	23	56	43	9	21	25	164	⁶ ₁₁ 8 %
303	10	10	24	20	5	11	16	80	³ ₆ 4 %
304	2	0	0	1	2	0	0	5	^{0,04} _{0,9} 0,3 %
305	1	0	4	0	0	0	0	5	^{0,04} _{0,9} 0,3 %
401	4	8	34	10	5	4	5	65	² ₅ 3 %
501	32	26	62	64	32	27	15	243	¹⁰ ₁₅ 12 %
502	13	14	8	41	12	18	4	106	⁴ ₇ 5 %
503	6	5	6	13	1	6	2	37	¹ ₃ 2 %
601	8	3	2	19	2	7	1	41	¹ ₃ 2 %
602	34	38	124	119	54	51	35	420	¹⁹ ₂₅ 22 %
Всего	200	202	484	582	209	269	145	2091	100%
<i>p</i>	10 ⁻⁹								
<i>BF</i> ₀₁	4 · 10 ¹⁸								
<i>V</i>	0,12 _{0,08 0,16}								

Обозначения и примечания: ДИ – доверительный интервал; BF_{01} – байесов фактор в пользу нулевой гипотезы H_0 против альтернативной гипотезы H_1 ; V – коэффициент сопряженности Крамера.

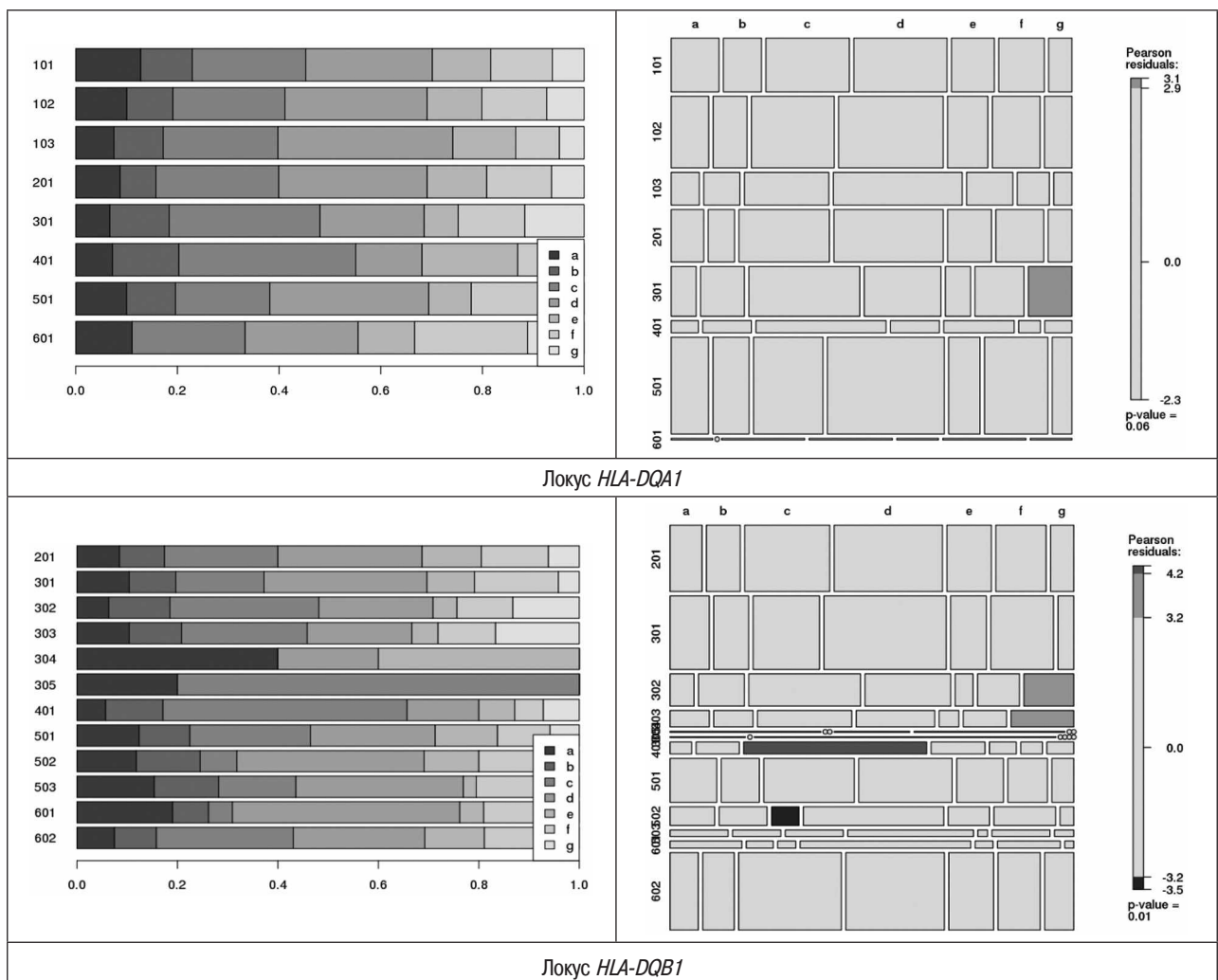


Рис. 10. Мозаичные диаграммы для таблиц сопряженности из табл. 10. Визуализация взаимоотношений между строками и столбцами в таблицах сопряженности. Использована интерактивная программа Chi-squareTest [64]

Для оценки точных p -значений (методом Монте-Карло) и коэффициента сопряженности Крамера (V) использован признанный эталонным пакет программ для точных непараметрических методов статистики StatXact [62]. Для вычисления бейзова фактора (BF_{01}) использован пакет программ JASP [63]. Инструментом для визуализации результатов анализа таблиц сопряженности являются мозаичные диаграммы (рис. 10).

В последнем столбце табл. 10 указаны «эталонные» частоты аллелей с границами точных

99%-ных доверительных интервалов (ДИ). Использована процедура вычисления так называемых *совместных* (*simultaneous*) ДИ, при которой наблюдаемые численности рассматриваются как выборка из полиномиального распределения. Для вычислений использована программа StatXact-8. Аналогичная ситуация наблюдается и в случае подбора референсной группы для частот генотипов по диморфизму $308G>A$ в гене *TNFA* на основе 18 независимых выборок здоровых россиян-европеоидов. Результаты представлены в табл. 11 (Плоткин и др., 2007) [65].

Таблица 11

Численность генотипов по диморфизму $308G>A$ в гене *TNFA* в 18 независимых выборках здоровых россиян-европеоидов

Местность	Генотипы			Всего
Уфа_1	GG	GA	AA	
Сибирь	53	65	2	159
СПб	47	4	1	91
Белгород	108	8	1	289
Москва_1	125	31	3	222
Москва_2	70	19	2	258
Москва_3	217	70	2	214
Новосибирск	169	50	3	400
Омск_1	204	49	5	117
Омск_2	167	43	4	52
Томск	290	100	10	189
Уфа_2	135	47	7	120
Уфа_3	122	54	4	180
Уфа_4	177	64	5	246
Уфа_5	273	103	5	381
Уфа_6	78	21	2	101
Уфа_7	133	65	6	204
Уфа_8	13	6	1	20
p	$8 \cdot 10^{-10}$			
BF_{01}	$7 \cdot 10^{10}$			
Всего	2381	799	63	3243
Частоты	$_{72}^{73}_{75} \%$	$_{23}^{25}_{26} \%$	$_{1}^{2}_{3} \%$	100%
Равновесие Харди–Вайнберга, mid- p	0,69			

Проверка согласия с равновесием Харди-Вайнберга

Генетика — уникальная наука в том смысле, что позволяет проверить, являются ли анализируемые данные случайной выборкой из панмиктической (свободно скрещивающейся) популяции. Процедура называется проверкой согласия с равновесием Харди–Вайнберга и должна быть обязательной при анализе таких данных. Полезным введением в проблему может служить приложение к университетскому учебнику для вузов Л. З. Кайданова «Генетика популяций» [66].

Проверка того, соответствуют ли частоты генотипов контролей (субъектов без болезней) рав-

новесию Харди–Вайнберга, дает представление о качестве проводимых молекулярно-генетических исследований. Отклонение от равновесия Харди–Вайнберга может быть вызвано такими факторами, как инбридинг, ассортативное скрещивание, отбор или миграция. Можно ожидать, что влияние этих факторов на равновесие Харди–Вайнберга будет незначительным в большинстве человеческих популяций, хотя отбор может играть важную роль при инфекционных заболеваниях. Другая причина — популяционная стратификация, которая вызывает дефицит гетерозигот. К избытку гетерозигот могут приводить также вариации числа копий генов (CNV). Наконец, отклонение от РХВ может быть вызвано просто ошибками генотипирования.

Ошибки генотипирования

Ошибки генотипирования неизбежны и являются скорее правилом, нежели исключением, и их частота для некоторых локусов может достигать 15%. Такие ошибки, даже если они происходят с очень низкой частотой, могут исказить результаты многих генетических анализов, таких, например, как оценка происхождения или родства или изучение связей/ассоциаций.

Ошибки генотипирования могут привести к серьезным смещениям и потере точности, и, кроме того, снижается статистическая мощность обнаруживаемых эффектов. Существует целый ряд типов ошибок, которые могут произойти в процессе генотипирования, включая ошибки при отборе образцов биоматериала для генотипирования, качество образца, взаимодействие между молекулами ДНК, перебои или неисправность оборудования и реагентов. Тем не менее основная причина ошибки генотипирования связана с человеческими факторами. Оценка качества метода генотипирования позволяет избежать ошибок генотипирования, которые приводят к ложным оценкам риска и отклонениям от равновесия Харди–Вайнберга. В современных исследованиях генотипирование основано на полностью автоматизированных методах, и эти методы подвержены ошибкам при определении генотипов из-за экспериментальной вариации и качества образцов ДНК, их загрязнения примесями и др. Очевидно, что анализы генотипирования чувствительны к загрязнению ДНК из планшет, пробирок, праймеров и других компонентов среды лаборатории, особенно для наиболее чувствительных новых технологий. Тем не менее другие типы систематических ошибок могут быть особенно тревожными, потому что они влияют на весь процесс исследования, и степень проблемы может отличаться в зависимости от случая и контроля. Например, весьма распространено группировать ДНК-планшеты отдельно для случаев и контролей или кодировать образцы структурированной системой нумерации, которая позволяет произвольно заранее идентифицировать случаи и контроли. Члены исследовательской группы, от лаборантов до статистиков, не должны пренебрегать рандомизацией и «ослеплением» образцов из групп случаев и контролей.

В недавней рукописи, депонированной в архиве bioRxiv, авторы сообщают: «Чипы ОНВ крайне ненадежны для генотипирования очень редких патогенных вариантов и не должны использоваться для принятия решений о здоровье без проверки» [67].

Если довольствоваться нахождением лишь p -значений, то вполне дружественной является интерактивная веб-страница, которая вычисля-

ет точные p -значения, как традиционные, так и mid- p , с поправкой на консервативность точных критериев [68].

Результат проверки согласия с равновесием Харди–Вайнберга с использованием этой программы для «референсной» группы из табл. 11 показан на рис. 11.

Hardy-Weinberg equilibrium exact test

799	heterozygote count
2381	homozygote count 1
63	homozygote count 2
clear	

☒ mid-p adjustment

P-value: 0.6930993494385

Рис. 11. Результат проверки согласия с PXB для «референсной» группы, подобранной на основе данных о численности генотипов по диморфизму 308G>Ав гене TNFA в 18 независимых выборках здоровых россиян-европеоидов

Очевидно, что имеет место удовлетворительное согласие с равновесием Харди–Вайнберга. Об этом же свидетельствует и бейзовский подход: $BF_{01} = 43,2$. Согласно вербальной шкале (табл. 4) такое значение BF_{01} можно интерпретировать как весомое свидетельство в пользу согласия анализируемых данных с равновесием Харди–Вайнберга против их отклонения от равновесия Харди–Вайнберга. Вычисления можно реализовать в известной программной среде R, в библиотеке которой есть специализированный пакет HWEbayes. После его инсталляции и запуска

```
library("HWEbayes",
lib.loc="~/R/win-library/3.5")
```

вводится команда с анализируемыми численностями генотипов:

```
HWEDirichBF2(nvec=c(2381,799,63),
bvec0=c(1,1),bvec1=c(1,1,1))
```

Результат:

```
[1] 43.22327
```

На рис. 12 представлен другой пример с результатами сравнения частот генотипов в двух сравниваемых группах школьников (контроль) и детей с артериальной гипертензией (АГ) и их согласия с PXB для диморфизма 3123C>T (rs1191046) в гене AGTR2 (ген рецептора типа 2 для ангиотензина II) [69]).

Fisher's exact test p-value calculator, 2x2 and 2x3

19	68	85
88	121	79

clear

☒ mid-p adjustment

P-value: 4.839749010924e-8

Hardy-Weinberg equilibrium exact test

19	heterozygote count	88	heterozygote count
68	homozygote count 1	121	homozygote count 1
85	homozygote count 2	79	homozygote count 2

clear

$F_{IS} = 0,78$

clear

$F_{IS} = 0,38$

☒ mid-p adjustment

☒ mid-p adjustment

P-value: 3.371272430225e-27

P-value: 1.62442472472e-10

Рис. 12. Результаты сравнения распределений частот генотипов с диморфизмом 3123C>T(rs1191046) в гене AGTR2 в двух группах детей — с АГ и здоровых школьников и проверка их согласия с равновесием Харди–Вайнберга

Обозначения вида 4.8e-8 есть так называемая экспоненциальная запись чисел в компьютере, и в обычной записи это: $4,8 \cdot 10^{-8}$; mid- p — p -значение с поправкой на консервативность точных вероятностных критериев.

Распределения частот генотипов имеют статистически высокосignificant различие ($p = 4,8 \cdot 10^{-8}$), однако в каждой из двух групп наблюдаются статистически высокосignificant отклонения от РХВ ($p = 3,4 \cdot 10^{-27}$ и $p = 1,6 \cdot 10^{-10}$). Более подробный анализ можно провести с помощью широко известной программы GENEPOP [70]. Для этого в текстовом редакторе надо создать входной файл вида (для группы детей с АГ):

```
AGTR2_HT
2
68
19 85
```

где первая строка — произвольное название, в данном примере AGTR2_HT, вторая строка — число аллелей, в данном примере 3, далее — численность генотипов.

Программа запускается путем задания команд в командной строке для директории, в которой находится исполнимый файл genepop.exe:

```
Genepop HWFile=AGTR2_HT.txt
HWfileOptions=1,2,3,4
```

Результат сохраняется в этом же файле. Программа GENEPOP уникальна тем, что позволяет проверять нулевую гипотезу о согласии с равновесием Харди–Вайнберга при трех альтернативных гипотезах: 1) отклонение от РХВ за счет дефицита гетерозигот (инбридинг); 2) из-за избытка гетерозигот (аутбридинг); 3) общее отклонение от равновесия Харди–Вайнберга.

Heterozygote deficiency test:

P-value=0; S.E=0

Heterozygote excess test:

P-value=1; S.E=0

Probability test:P-value=0; S.E=0

Такие же значения получаются и для контрольной группы школьников. Здесь $S.E.$ есть стандартная ошибка для оценки p -значения методом Монте-Карло. P -value = 0 не означает, что оно равно нулю, а только что $p < 10^{-6}$. Результаты подтверждаются и бейзовским анализом. Для группы с АГ $BF_{01} = 2 \cdot 10^{-25}$ и для контрольной группы $BF_{01} = 6,8 \cdot 10^{-9}$. Эти значения свидетельствуют в пользу альтернативной гипотезы (отклонение от равновесия Харди–Вайнберга реально) против гипотезы о согласии с равновесием Харди–Вайнберга. Кроме этого, программа GENEPOP позволяет оценить коэффициент инбридинга F_{IS} (его более строгое название — индекс фиксации). В данном случае получается $F_{IS} = 0,78$. В контрольной группе также обнаруживается инбридинг с $F_{IS} = 0,38$. Для оценки F_{IS} можно использовать также программу FPtest [71]. Есть еще уникальная полезная программа HWmet [72], которая вычисляет бейзовские (credible — правдоподобные) ДИ для индексов фиксации F_{IS} , но, к сожалению, для нее имеется только код, который надо самостоятельно компилировать в исполнимый файл [73]. Скорее всего, причиной столь невероятных отклонений от равновесия Харди–Вайнберга в обсуждаемой работе являются ошибки генотипирования. На этот счет существует обширная литература [74–77].

Программа GENEPOP позволяет не просто проверить согласие данных с РХВ, но и выявить причину наблюдаемого отклонения. В обеих группах — это инбридинг с очень высокими значениями коэффициента инбридинга (индекса фиксации) $F_{IS} = 0,78$ и $0,38$. Таким значениям могут удовлетворять популяции из клонов и ближайших родственников, что, конечно же, исключено. Следовательно, сравнивать распределения частот генотипов в этих группах фактически неправомерно, т. е. наблюдаемое статистически высокосignificant различие, скорее всего, является ложным.

Программа GENEPOP пригодна также и для дифференцированного анализа избытка и дефицита гетерозигот в локусах с числом аллелей больше двух. Примером могут служить данные о частотах генотипов в локусе CYP2C9 (ген цитохрома P450 2C9) с тремя аллелями (*1, *2 и *3) из работы Шнайдер и др. [78] (табл. 12).

Таблица 12

Частота встречаемости генотипов у наблюдаемых пациентов ($N = 148$)

Генотип	Численности, n	Доли с 95%-ными ДИ
CYP2C9*1/*1	100	$_{57}^{68} 77\%$
CYP2C9*1/*2	18	$_{6}^{12} 21\%$
CYP2C9*2/*2	0	$_{0,0}^{0,0} 3,3\%$
CYP2C9*1/*3	28	$_{11}^{19} 28\%$
CYP2C9*3/*3	0	$_{0,0}^{0,0} 3,3\%$
CYP2C9*2/*3	2	$_{0,1}^{1,4} 6,1\%$

Обозначения: ДИ — доверительный интервал.

Согласно современным рекомендациям, следует указывать доверительные интервалы для всех оцениваемых показателей. Поэтому здесь в виде подстрочных индексов значения долей дополнены точными границами 95%-ных ДИ, которые можно вычислить с помощью признанного эталонным пакета статистических программ для точных непараметрических методов StatXact [62]. Вероятностной моделью этих данных является полиномиальное распределение, и эта программа является фактически единственной, которая проводит корректное вычисление точных ДИ для его параметров. Для большинства подобных оценок достаточно указывать значения долей и границ ДИ с точностью до двух значащих цифр.

Входной файл для проверки согласия с РХВ в данном случае имеет вид:

```
CYP2C9
3
100
18 0
28 0 2
```

Результат:

```
Heterozygote deficiency test:
P-value=0.70; S.E=0.0037
Heterozygote excess test:
P-value=0.32; S.E=0.0068
Probability test: P-value=0.45;
S.E=0.0060
```

Анализируемые данные статистически согласуются с равновесием Харди–Вайнберга: все три p -значения статистически незначимы. Программа GENEPOP позволяет оценивать не только общее F_{IS} , но и частные (парциальные) F_{IS} для каждой аллели по-отдельности:

Allele	Sample count	Frequency	F_{IS} ----- W&C
1	246	0.8311	-0.1036
2	18	0.0608	-0.0614
3	32	0.1081	0.0223
Tot	296		-0.0541

Тестирование неравновесности передачи аллелей и/или сравнение групп случаев и контролей

Генетика есть наука о наследственности и изменчивости. Наследственность есть фундаментальное свойство живых организмов передавать свои признаки и гены в ряду поколений. Если данная аллель и генотип чаще встречаются у особей с данным признаком и они действительно определяют до определенной степени этот признак, то это их преобладание должно проявляться также и в предыдущих и последующих поколениях. Они должны передаваться в чреде поколений неслучайно, неравновесно. Для выявления таких неравновесностей давно разработаны специальные тесты, которые так и называются ТНПА — «тесты неравновесности передачи аллели» (TDT — Transmission Disequilibrium Test).

В любом исследовании ассоциации генотип-фенотип подбор контрольной группы имеет решающее значение. Как бы исследователи ни старались подобрать равноценные контроли и случаи, невозможно быть абсолютно уверенным, что ничто не упущено. Таким образом, когда обнаруживается связь генотип-фенотип, всегда возникает беспокойство, что она может быть вызвана не неравновесностью по сцеплению, а неадекватно подобранными контролями. Тест на неравновесную передачу аллелей позволяет избежать проблемы подбора группы контролей, сопоставимой с группой случаев. Один из способов полностью избежать проблемы сопоставления — это использовать внутренние контроли, данные о которых поступают от тех же людей, что и данные о случаях. ТНПА реализует эту идею. ТНПА проводят с семьями, в которых есть один или несколько пораженных потомков. При этом неважно, поражен ли кто-либо из родителей. Этот тест позволяет проверить гипотезу о том, что аллель M_1 маркера M связана с интересующим признаком (например, с заболеванием). Для этого отбирают случаи, в которых родитель является носителем гетерозиготной комбинации аллели M_1 с любой другой аллелью M_x маркера M . Если аллель M_1 не обязательно связана с раз-

витиём болезни, то пораженные потомки не обязательно унаследуют эту аллель. Однако, если аллель M_1 играет какую-либо роль в предрасположенности к данной болезни, то можно ожидать, что пораженные дети в этой группе будут наследовать от родителя аллель M_1 чаще, чем другую его аллель. Поэтому поступают следующим образом:

Выявляют пораженных пробандов, и их и их родителей генотипируют по маркеру M . Рассматривают данные только от тех родителей, которые гетерозиготны по маркерной аллели M_1 . Не имеет значения, какой является другая аллель (при условии, что это не аллель M_1).

В каждом случае показан только один родитель.

ТНПА, основанная на стандартной статистике критерия χ^2 , предельно проста:

$$(a - b)^2 / (a + b),$$

где a — число случаев, когда гетерозиготный родитель передал аллель M_1 пораженному потомству, а b — число случаев, когда передается другая аллель M_x .

На рис. 13 представлена схема проведения теста неравновесности передачи аллелей на примере диабета 1-го типа. 94 семейства были исследованы на предмет наличия связи между диабетом 1-го типа и конкретной аллелью в повторяющейся последовательности перед геном инсулина. Среди этих 94 семей было 57 родителей, которые были гетерозиготными по исследуемой аллели M_1 (на рисунке это I), и некоторыми другими аллелями M_x (на рисунке это X). Эти 57 родителей передали 124 аллели потомству с диабетом (у некоторых было более одного больного ребенка), из них у 78 были аллели M_1 и 46 аллели M_x .

При равновесной передаче аллелей M_1 и M_x от родителей к пораженным болезнью потомкам

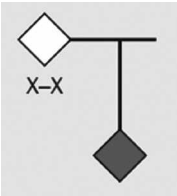
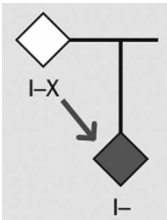
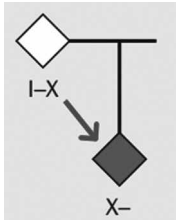
94 семьи, имеющие детей (одного ребенка и более), страдающих диабетом		
37 семей, в которых ни один родитель не является гетерозиготой по аллели M_1	57 семей, в которых один родитель является гетерозиготой по аллели M_1	
		
	Аллель M_1 перенесена от родителя к ребенку, 78 случаев	От родителя к ребенку перенесена другая аллель M_x , 46 случаев

Рис. 13. Пример данных для ТНПА [79]

их число должно быть примерно одинаковым. В данном примере статистика критерия для проверки их равенства принимает значение $\chi^2 = (78 - 46)^2 / (78 + 46) = 8,26$, чему соответствует $p = 0,0041$. Точный (биномиальный) критерий дает $p = 0,0052$. Таким образом, анализируемые данные демонстрируют статистически значимую связь между аллелью M_1 и диабетом 1-го типа.

Однако значения бейзова фактора получаются малыми: $BF_{01} = 0,42$; $BF_{10} = 2,38$ (их можно вычислить с помощью интерактивной программы Bayes factor for a binomially distributed observation [80]. Согласно вербальной шкале (табл. 86) такие значения BF принято интерпретировать как пренебрежимо малые. То есть, несмотря на статистическую значимость, эти результаты не позволяют ни подтвердить, ни отклонить гипотезу о равновесности передачи

аллелей M_1 и M_x от гетерозиготных родителей к детям, пораженным диабетом 1-го типа. Подобные тесты разработаны и для более сложных сценариев, когда отсутствуют сведения об одном из родителей (общая проблема в случаях с поздними заболеваниями). Если отсутствуют сведения об обоих родителях, то можно сравнивать частоты аллелей между пораженными и непораженными sibсами и т. п. [79].

Метаанализ

Недавно Bousman et al., 2019 [81] провели систематический обзор и метаанализ пяти проспективных рандомизированных контролируемых исследований (РКИ), в которых изучались инструменты поддержки принятия решений (ИППР) с помощью фармакогене-

тики, относящиеся к ремиссии депрессивных симптомов при большом депрессивном расстройстве (БДР). По мнению авторов, лица, получавшие терапию на основе фармакогенетического ИППР ($n = 887$), в 1,71 раза (95% ДИ: 1,17–2,48; $p = 0,005$) имели больше шансов

достичь ремиссии симптомов по сравнению с лицами, получавшими лечение в обычном режиме ($n = 850$) (рис. 14). Авторы утверждают, что фармакогенетически управляемые ИППР могут улучшить ремиссию симптомов у пациентов с БДР.

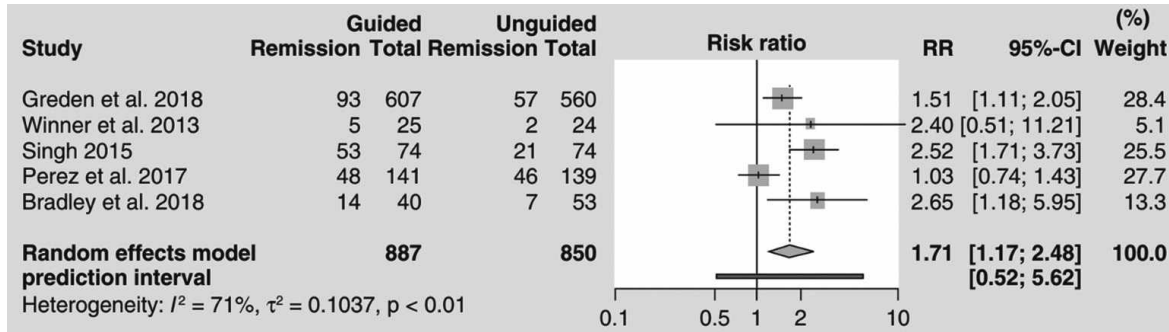


Рис. 14. Данные пяти РКИ и результаты их метаанализа [81]

Здесь мы наблюдаем проявление «синдрома статистической снисходительности». Авторы по инерции продолжают ориентироваться на преодоление традиционно и повсеместно используемого уровня значимости $\alpha = 0,05$ и, соответственно, используют уровень доверия (доверительную вероятность) $(1 - \alpha) = 95\%$. При этих условиях 95%-ный ДИ (от 1,17 до 2,48) для оценки сводного значения RR (отношения рисков или относительного риска) не покрывает безразличное значение $RR_{\text{indiff}} = 0$, и формально эффект, оцениваемый этим интервалом, можно признать статистически значимым, но всего лишь на самом снисходительном уровне значимости $\alpha = 0,05$. Одна-

ко уже 99%-ный ДИ для RR (от 0,71 до 4,16) покрывает безразличное значение $RR_{\text{indiff}} = 0$. Это означает, что уже на уровне значимости $\alpha = 0,01$ эффект следует признать статистически незначимым. Согласно современным требованиям, предъявляемым к метаанализу, полагается вычислять не только доверительные интервалы (ДИ), но и предсказательные интервалы (ПИ). Авторы приводят границы 95%-ного ПИ: 0,52; 5,62, а 99%-ный ПИ, естественно, получается еще более широким: от 0,26 до 11,42 (рис. 15).

Таким образом, очевидно, следует относиться более скептически к излишне оптимистичным утверждениям авторов.

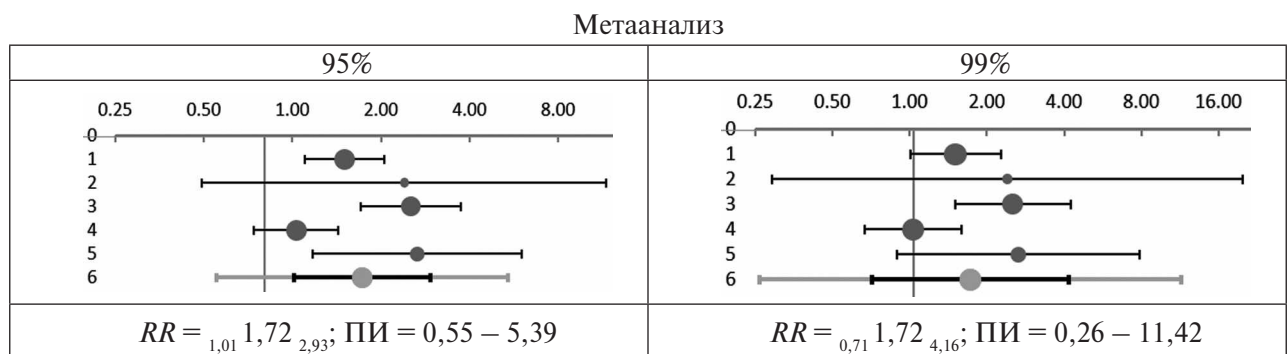


Рис. 15. «Древовидные» графики (forest plot) для результатов метаанализа

пяти проспективных рандомизированных контролируемых исследований, в которых изучалось влияние фармакогенетической терапии на ремиссию при большом депрессивном расстройстве

Обозначения: 95% и 99% — уровни доверия (доверительные вероятности) и/или предсказания, RR — относительный риск, ПИ — предсказательный интервал. Черные отрезки — ДИ, более длинные серые отрезки — ПИ. Использована программа Meta-Essentials [82].

Рекомендуется интерпретировать предсказательный интервал как описание диапазона наблюдаемых величин эффекта, а не как прогноз диапазона значений эффекта, который будет наблюдаться в будущих исследованиях (несмотря на его название «предсказательный») [80]. Тонкость, конечно, едва уловимая. Точно также следует помнить и осознавать, что доверительный интервал (ДИ) не является интервалом, в котором находится истинное значение оцениваемого (неизвестного) параметра с заданной доверительной вероятностью. Это легко понять, если представить, что мы повторим вычисления ДИ для других независимых наборов данных (других выборок из той же генеральной совокупности). Очевидно, что получаемые ДИ будут варьироваться, т. е. ДИ являются не фиксированными, а случайными (вероятностными). Доверительная вероятность, скажем, 95% означает, что при многократных повторениях эксперимента 95% полученных 95%-ных ДИ накроют оцениваемое (неизвестное) значение параметра, а в 5% случаев не накроют. И мы никогда не узнаем, какие из полученных ДИ накроют или не накроют оцениваемое значение.

От статистических описаний к вероятностно-статистическим предсказаниям

Критериями научной истины являются доказательная и предсказательная способности аналитических инструментов. Возможность предсказания в науке представляет ее сущность. Ведущую роль в экспериментальной науке играет теория эксперимента, которая базируется на вероятностно-статистической методологии. Соответственно, вероятностно-статистические предсказания представляются чрезвычайно важными с практической точки зрения. Статистика изучает повторяющиеся (взаимозаменяемые по де Финетти) данные. Повторяемость составляет суть статистики и это оснащает ее предсказательной способностью. Современная статистика способна оценивать не только то, что есть, но и то, что будет, т. е. как настоящая наука она обладает предсказательной силой.

Современная прикладная статистика довольно искусственно ограничивает себя изучением фактических данных, однако она обладает и предсказательной способностью, столь необходимой для прогресса науки и практического использования ее достижений. Традиционно слишком часто считается достаточным объявить о статистической значимости наблюдаемого эффекта и отождествить таковую с его практической значимостью. Однако этого явно недостаточно — требуется оценить предсказательную ценность

полученных данных. Достаточно эффективными инструментами для этого являются *предсказательные вероятности* и *предсказательные интервалы* (ПИ). Для этого в статистике разработаны соответствующие методы и вычислительные алгоритмы, реализованные в специализированных программных продуктах. Проблема предсказаний при анализе статистических данных представляется исключительно важной при практическом применении.

Ведущие специалисты по статистике давно, но, возможно, не слишком настойчиво, призывают исследователей обращать внимание на способность статистики предсказывать будущие результаты экспериментов и наблюдений. Соответствующие подходы, алгоритмы и формулы давно известны, и разработаны соответствующие программы. Продемонстрировать эти возможности можно на примере оценки диагностической информативности теста генерации тромбина (ТГТ) [84, 85]. Этот тест позволяет оценить динамику образования и инактивации *in vitro* ключевого фермента гемостаза — тромбина (КФ 3.4.21.5), и его используют для оценки действия антиагрегантных препаратов. Результаты теста характеризует ряд временных и количественных параметров (рис. 16). Целью исследования было оценить возможности использования теста генерации тромбина в богатой и бедной тромбоцитами плазме для оценки эффективности двойной антиагрегантной терапии у больных ИБС после интракоронарного стентирования.

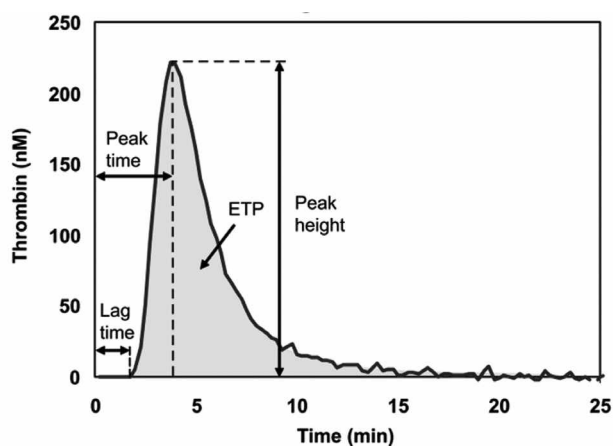


Рис. 16. Параметры кривой генерации тромбина [86]

Lag time (LT) — время задержки (лаг-период), мин; Peak time (PT) — время достижения пика, мин; Peak height (PH) — высота пика, нМ; Endogenous thrombin potential (ETP) — потенциал эндогенного тромбина (площадь под кривой), нМ·мин; V — максимальная скорость генерации тромбина — тангенс угла наклона левой прямолинейной части кривой — $V = PH / (PT - LT)$, нМ/мин.

В исследование были включены 54 больных ИБС после плановой чрескожной транслюминальной коронароангиопластики со стентированием и 40 человек без клинических проявлений ИБС и не получавших данные препараты с какой-либо другой целью. Исследование проводилось на фоне двойной антиагрегантной терапии в среднетерапевтических дозировках (клопидогрел и аспирин в дозах 75 и 75–100 мг в сутки соответственно) спустя 6 месяцев после вмешательства.

Для измерения разноименных эффектов могут использоваться различные единицы измерения, и их бывает трудно (практически невозможно)

сопоставлять. Например, в данном случае трудно понять, какая из разностей средних является клинически более значимой: $_{34}^{167} \text{М} \cdot \text{мин}$ для параметра ЕТР или $_{3,6}^{5,1} \text{нМ/мин}$ — для параметра V? Для унификации подобных сравнений существуют различные способы преобразования исходных данных в безымянные (безразмерные) величины. Таковыми, например, являются непараметрическая мера эффекта *AUC* — площадь под *ROC*-кривой — измеряемая в долях, и параметрический стандартизированный размер эффекта по Коэну, измеряемый в единицах стандартного отклонения. Результаты представлены в табл. 13.

Таблица 13

Сопоставление абсолютного и стандартизированного размеров эффекта

Параметр кривой ТГТ	Разность средних					Стандартизированный размер эффекта					Уровень доверия
	D_{01}	Границы ДИ		Границы ПИ		d_c	Границы ДИ		Границы ПИ		
LT	–1,1	–3,5	0,86	–4,3	1,7	–0,25	–0,66	0,16	–0,84	0,33	95%
		–4,2	1,6	–5,3	2,7		–0,79	0,29	–1,0	0,51	99%
		(мин)									
ETP	167	34	299	–21	355	0,52	–0,10	0,94	–0,07	1,1	95%
		–9	343	–82	416		–0,03	1,1	–0,25	1,3	99%
		(нМ · мин)									
PT	–4,5	–7,1	–2,0	–8	–0,91	–0,73	–1,2	–0,31	–1,3	–0,14	95%
		–7,9	–1,1	–9	0,26		–1,3	–0,18	–1,6	0,04	99%
		(мин)									
PH	28	18	39	13	43	1,09	0,65	1,5	0,49	1,8	95%
		14	43	8,0	48		0,51	1,7	0,30	2,0	99%
		(нМ)									
V	5,1	3,6	6,7	2,9	7,3	1,35	0,90	1,8	0,76	2,1	95%
		3,0	7,2	2,2	8,0		0,76	2,0	0,57	2,3	99%
		(нМ/мин)									

Обозначения: D_{01} — разность средних значений; d_c — стандартизированный размер эффекта по Коэну; ДИ — доверительный интервал; ПИ — предсказательный интервал. Использованы программы LePreg и JASP.

Можно видеть, как использование стандартизированного размера эффекта d_c позволяет объективизировать и универсализировать результаты измерений, произведенных в разных единицах измерения, и ранжировать разноименные параметры кривой ТГТ по «силе» эффекта (стандартизированной разности средних в сравниваемых группах пациентов). Согласно вербальной шкале (табл. 3б) только эффекты для параметров РН и V можно интерпретировать как умеренные, остальные эффекты следует признать малыми (ЕТР и ТТР) и даже ничтожным (LT). Полезную информацию несут ПИ (предсказательные интервалы). Они позволяют проследить, как могут измениться интервальные оценки, если повторить подобные измерения в независимом эксперименте.

Стандартизация эффектов и вычисление не только ДИ, но и ПИ позволяют оценить и объективно сравнить информативность изученных параметров кривой ТГТ (табл. 14).

Можно видеть, что значения параметра «время задержки» практически не различаются в сравниваемых группах. И если пытаться достигнуть статистически значимого различия, то для этого потребуются объемы выборок более 400 человек в каждой группе при $\alpha = 0,05$ или в два раза бóльшие — более 800 человек — при $\alpha = 0,01$. Очевидно, что такой гармонизированный анализ позволяет предостеречь исследователя от дальнейших попыток добиваться статистической значимости этого эффекта (любой ценой).

Таблица 14

Информативность параметров теста генерации тромбина

Пара- метр кривой ТГТ	d_c	ДИ для d_c		P_{val}	ПИ для d_c		ПИ для P_{val}		Дов. уровни	$n_1 = n_2$
LT	-0,25	-0,66	0,16	0,37	-0,84	0,33	$5 \cdot 10^{-5}$	0,94	95%	437
		-0,79	0,29		-1,04	0,52	10^{-6}	0,99	99%	808
ETP	0,52	0,10	0,94	0,015	-0,07	1,13	10^{-7}	0,63	95%	102
		-0,03	1,07		-0,26	1,34	10^{-9}	0,87	99%	188
PT	-0,73	-1,15	-0,30	0,0012	-1,35	-0,14	10^{-9}	0,26	95%	53
		-1,29	-0,17		-1,56	0,05	10^{-11}	0,59	99%	97
PH	1,09	0,65	1,53	10^{-6}	0,48	1,75	10^{-13}	0,012	95%	24
		0,51	1,67		0,29	1,97	10^{-15}	0,084	99%	45
V	1,35	0,89	1,80	10^{-8}	0,72	2,03	10^{-16}	$4 \cdot 10^{-4}$	95%	17
		0,75	1,95		0,54	2,26	$< 10^{-17}$	0,0062	99%	30

Обозначения: d_c — стандартизированный размер эффекта по Коэну; P_{val} — p -значение; $n_1 = n_2$ — минимальные объемы выборок для повторения. Использованы программы: LePrep, G*Power, JASP.

Даже параметр PT (время достижения пика) на кривой ТГТ оказывается малоинформативным. Несмотря на исходную его, казалось бы, высокую статистическую значимость ($p = 0,0012$), в повторном эксперименте может получиться 99%-ный ПИ, накрывающий безразличное значение d_c от $-1,56$ до $+0,05$, и верхняя граница ПИ для p -значения может достигнуть значений 0,26 (при уровне предсказания 95%) и 0,59 (при уровне предсказания 99%).

Наиболее информативным оказывается параметр V (максимальная скорость генерации тромбина). В табл. 14 эти данные выделены жирным шрифтом.

При повторении эксперимента (в том же объеме) он с высокой вероятностью останется статистически значимым. Однако столь же несомненно, что для большей убедительности эксперимент необходимо повторить и, желательно, с выборками большего объема.

Итак, подробный статистический анализ, гармонично сочетающий все возможности современной статистической методологии с обязательным вычислением стандартизированных размеров эффекта с их интервальными оценками, как доверительными, так и предсказательными, оказывается плодотворным подходом в современных биомедицинских исследованиях.

Заклучение

Предсказание наличия или отсутствия того или иного признака или той или иной особенности развития посредством генетического тестирования является вожеленной мечтой исследователей. Однако существуют фундаментальные и непреодолимые ограничения для точности и надежности предсказания вследствие сложной архитектуры распространенных болезней, эпистатических взаимодействий, а также множества негенетических факторов. Неоднозначность генетического тестирования кроется уже в неопределенности при определении (диагностике) изучаемого признака (фенотипа). Например, как отличить «спортсмена» от «не спортсмена»? Сегодня человек здоров и может быть отобран в контрольную группу, но завтра он заболел и попадет в группу случаев.

Генотип + среда = фенотип

Известная формула — фенотип (P) есть продукт взаимодействия генотипа (G) и среды (E) — является одним из фундаментальных положений современной биологии. Простейшей символьной схемой взаимодействия генотипа (G) и среды (E) при становлении фенотипа (P) является аддитивная (линейная) модель:

$$P_i = G_i + E$$

Однако данное линейное уравнение излишне упрощенное и далеко от истины; на самом деле эти взаимодействия являются нелинейными, мультипликативными, компенсаторными, нейтрализующими, противоположными и т. д. Некоторые из них могут быть молчащими, скрытыми, которые не проявляются в обычных (нормальных) условиях, и поэтому их трудно обнаружить. Схематично их можно представить в виде:

$$P_i = G_i + E + (G_i \times E) + (G_i \times G_k) + \dots$$

Кроме того, уже давно стало понятно, что эта схема должна быть дополнена случайными «ошибками (шумом) развития» (С):

$$P_i = G_i + E + (G_i \times E) + (G_i \times G_k) + C \dots$$

Иногда их называют «нематериальной изменчивостью» (Falconer) или «третьим компонентом» (Gärtner). Источником третьего компонента могут служить, например, случайные колебания концентрации морфогенов или случайная неравномерность клеточных делений на разных этапах эмбрио- и органогенеза.

Фенотип является продуктом запутаннейших (нелинейных) взаимодействий множества генов друг с другом, со средой и с шумовыми факторами, которые могут быть как гармоничными, так и конфликтующими, конкурирующими, взаимно нейтрализующими, разнонаправленными. Гены, среда и шумы действуют совместно на всех уровнях биологической организации, и количество их взаимосвязей не поддается учету. Сообщая о том, что за определенный признак или свойства человека отвечает определенный ген, исследователи каждый раз должны делать оговорку: не он один. Жизнедеятельность любого организма зависит от слаженного взаимодействия тысяч генов. Поэтому, зная генотип, нельзя однозначно предсказать фенотип, и наоборот: зная фенотип, невозможно однозначно предсказать генотип.

Многопричинность становления фенотипа

Большинство признаков (фенотипов) являются мультифакторными, что в переводе на русский язык означает их многопричинность, и тогда понятно, что их причинами в таком случае являются как их полигенное наследование, так и многочисленные факторы среды, которые в естественных условиях трудно (или даже невозможно) учесть и проконтролировать. В силу этого большинство многопричинных признаков имеют ограниченную генетическую обусловленность.

По отношению к каждому гену в данном геноме все остальные гены образуют своеобразную генетическую (генотипическую) среду, фон. Генотипическая среда оказывается принципиаль-

но не учитываемой. Даже при применении технологий широкогеномного и полногеномного генотипирования (GWAS, WGAS), позволяющей анализировать сразу до 1,2 миллионов аллелей, за бортом оказываются десятки миллионов вариантов. На фоне столь неопределенной, неконтролируемой исследователями генотипической среды вряд ли можно делать сколько-нибудь определенные предсказания о наличии или развитии данного многопричинного признака с заведомо ограниченной генетической обусловленностью.

Одновременно с этим под средой следует понимать не только факторы, внешние по отношению к данному организму, но и внутренние (экспозом, эпигеном и микробиом), начиная с оплодотворения и внутриутробного развития. При генетическом тестировании фактически предается забвению гомеостаз, обусловленный во многом избыточностью (дублированием) и взаимозаменяемостью многих жизненно важных генов и компенсаторными и регуляторными генными взаимодействиями. Редко удается выявить межallelные (внутригенные) и межгенные взаимодействия и эпигенетические явления. Мало внимания уделяется классической проблеме взаимодействия хромосомных и митохондриевых генов. Упускаются из виду повсеместно распространенные энтеровирусы и прочие инфекционные агенты. Сейчас становится ясным, что генетика человека в немалой степени становится генетикой его микробиоты, т. е. населяющих его микробов.

Все это проявляется в таких явлениях, как пенетрантность, экспрессивность, норма реакции, эпистаз, межallelная комплементация и, в особенности, плейотропия. Чем большим числом факторов, как генетических, так и средовых, определяются признаки, такие как предрасположенность к заболеванию, изменение скорости метаболизма ксенобиотиков (например, лекарств), склонность к талантам или, напротив, устойчивость к заболеваниям, отсутствие талантов, тем ниже предсказуемость этого признака. Одно и то же заболевание может определяться разными вариантами разных генов. Один и тот же ген может быть задействован в становлении разных признаков (заболеваний, синдромов и других). Одни варианты данного гена (варианты последовательности ДНК) могут предрасполагать к одной болезни, а другие варианты — к другой болезни или быть протективными в отношении другой болезни.

Плейотропия является одним из основных источников неопределенности и низкой предсказательной способности генетического тестирования, особенно когда она разнонаправленная («компромиссная» или «компенсаторная») и/или антагони-

стическая. Ее следствием могут быть внутригенные (межаллельные) и межгенные конфликты.

Важнейший вопрос в генетическом тестировании: усиливаются ли или хотя бы складываются «вредные» (или «полезные») эффекты тестируемых аллелей и генотипов при их объединении в одном генотипе, или же, напротив, они взаимно нейтрализуются?

При генетическом тестировании (за исключением раковых заболеваний) изучается и используется природный сбалансированный генетический полиморфизм, т. е. не генетические новообразования (мутации), а аллели, прошедшие естественный отбор и закрепившиеся в популяциях человека. Иными словами, изучаются не аномалии генома, не патологические или патогенные его варианты, а его бесконечно разнообразные, но естественные, «нормальные» вариации. Нейтральностью и сбалансированностью объясняется тот факт, что предрасполагающие генотипы встречаются как у больных, так и у здоровых, и различаются лишь их частоты в группах субъектов с данной болезнью и без нее. Поэтому заведомо следует ожидать, что их вклад в становление того или иного признака будет малым в силу действия плейотропных компенсаторных механизмов.

С точки зрения эволюционного учения, одним из первейших вопросов, на который должна дать ответ геномика, может звучать так: является ли генетический полиморфизм, выявляемый современной геномикой, результатом нейтральной эволюции или же он является отягчающим генетическим (мутационным) грузом, определяющим предрасположенность к распространенным болезням, который почему-то естественный отбор вовремя не отбраковал?

Нарождающаяся на наших глазах новая область молекулярной медицины — эволюционная медицинская геномика — свидетельствует, что *подавляющее большинство вариантов (аллелей), которые наблюдаются в геномах современных популяций человека, являются селективно нейтральными*. Действительно, оказывается, что кодирующие, т. е. функционально значимые районы в геноме человека демонстрируют гораздо меньшую степень варьирования, чем некодирующие, т. е. участки, функция которых неизвестна. Абсолютное число синонимичных вариантов превышает число несинонимичных (миссенс) вариантов, несмотря даже на то, что число позиций, в которых могут происходить несинонимичные варианты, в три раза превышает число позиций, в которых могут происходить синонимичные мутации. Доля синонимичных вариантов в 4 раза больше доли несинонимичных: 80% и 20% соответственно. В целом, нейтралистские

эволюционные воззрения приводят к выводу, что исторические адаптивные эволюционные события не являются источником болезней. Напротив, эволюция является источником устойчивости и причиной того, что люди столь успешно существуют в широко изменяющихся условиях.

Разнообразие уникальностей, не поддающееся учету

Число аллелей (вариантов последовательностей) астрономическое. Число только новых обнаруживаемых однонуклеотидных вариантов последовательностей удваивается каждые 13 месяцев и превышает уже полмиллиарда. Число их комбинаций (гаплотипов и генотипов) не поддается учету и изучению, но их гораздо больше, чем число жителей Земли — 7,7 млрд, а возможно, даже больше, чем число атомов во Вселенной — 10^{67} . Поэтому заведомо невозможно доказать, что именно данный уникальный генотип является причиной данного заболевания или данной склонности. Для этого надо иметь большие выборки субъектов с таким генотипом, но он уникален.

Персонализированную медицину популярно сравнивают с подгонкой портным одежды клиента по фигуре. Мол, если секвенировать полностью геном данного человека, то удастся подобрать ему сугубо индивидуальные средства профилактики, лечения, диеты, образ жизни и проч. Однако если продолжить эту аналогию, то ясно, что «подгонять одежду по фигуре» придется на протяжении всей его жизни.

Предрасполагающие генотипы (аллели, гаплотипы) встречаются и у больных, и у здоровых, а различаются лишь их частоты в группах субъектов с данной болезнью («случаев») и без нее («контролей»). То есть заведомо наличие в генотипе данного человека предрасполагающей аллели не свидетельствует о неизбежном наличии у него болезни или иной склонности в настоящем или о ее возникновении в будущем. Равно как и наличие у данного человека данной болезни или склонности не свидетельствует о наличии в его геноме предрасполагающих к ним вариантов генов. Это означает, что генетическое тестирование не способно однозначно предсказать наличие или развитие того или иного признака или предрасположенности к тому или иному свойству, оно лишь оценивает ее вероятность.

Каждый человек уникален, различаются даже монозиготные близнецы: числом копий генов (CNV), структурой иммуноглобулинов, отпечатками пальцев. Очевидно, что этот факт является одним из основных источников низкой воспроизводимости и прогностической способности

результатов генетического тестирования (фармакогенетического, в частности). Таким образом, генетическая уникальность каждого человека вступает в противоречие с методологией клинических испытаний и статистической методологией, которые требуют анализировать выборки случаев и контролей большого объема (тысяч и даже сотен тысяч) *идентичных* лиц для достижения убедительных выводов.

Кризис воспроизводимости и кризис биостатистики

Прогнозируемые на основе генетического тестирования риски существенно различаются между лабораториями и/или компаниями и нередко противоречат друг другу. Основная причина противоречивости результатов в геномике и противоречивости их интерпретаций — слабая воспроизводимость результатов генетического тестирования. Основная причина плохой воспроизводимости — статистически высокая значимая гетерогенность распределений частот изучаемых генотипов (даже в одной и той же популяции). Одна из причин статистически высокой значимой гетерогенности распределений частот изучаемых генотипов в изучаемых группах индивидуумов — неслучайность выборок. Одно из последствий неслучайности выборок — статистически высоко значимые отклонения от равновесия Харди–Вайнберга. Например, доноры не могут быть случайной выборкой. Причинами значимых отклонений от равновесия Харди–Вайнберга могут быть ошибки при формировании выборки; дублирование данных об одном и том же субъекте; пренебрежение рандомизацией; расслоенность (стратификация) популяций (эффект Валунда); игнорирование «нулевых» (неамплифицируемых) аллелей; трудность определения редких аллелей; эпигенетические эффекты и даже ошибки в вычислениях.

Многочисленные (мультифакторные) признаки (болезни, метаболизм ксенобиотиков и т. п.) суть количественные признаки с ограниченной генетической обусловленностью. Поэтому ведущая роль в их изучении принадлежит популяционной генетике, генетике количественных признаков и генетической эпидемиологии.

Диагностика на основе только результатов поперечно-срезовых ассоциативных исследований может оказаться чреватой, мягко говоря, недоразумениями. Генетика есть наука о наследственности. Наследственность есть свойство передавать гены, признаки и особенности развития в ряду поколений. Поэтому для убедительного доказательства наследственной природы того или иного многопричинного признака

(заболевание, метаболизм ксенобиотиков) надо вовлекать в исследование, как минимум, два поколения индивидуумов — детей и родителей или других ближайших родственников. Нужно изучать неравновесность передачи генетических вариантов (аллелей, гаплотипов, генотипов) от родителей к детям или между другими близкими родственниками (transmission disequilibrium test — TDT).

Любая ассоциация, если она не подтверждена тестом неравновесности передачи аллелей, должна рассматриваться как сугубо предварительная до тех пор, пока не будут представлены дополнительные доказательства причинно-следственной связи. Еще один источник беспокойства: множественная проверка гипотез, усугубленная публикационным смещением. Кроме того, давно пора внедрить во врачебную практику на всех уровнях медицинской помощи ведение родословных и приучать к этому пациентов. Семейные истории являются более сильными предикторами многопричинных признаков, чем результаты поперечно-срезовых генетических тестирований.

С неопределенностью генетического тестирования связаны такие явления, как «проклятие победителя» (Winner's curse), публикационное смещение, ассортативное смещение, аргументационное смещение, удачная остановка сбора данных. «Проклятие победителя-первооткрывателя» состоит в том, что в первоначальных (пилотных, разведочных) публикациях авторы («победители», «первооткрыватели») сообщают о размере эффекта, который в последующих проверочных исследованиях часто оказывается гораздо меньшим или даже вообще не подтверждается (не воспроизводится). Публикационное смещение — авторы публикуют не все результаты, а только те, которые их устраивают. Ассортативное смещение — предвзятое (не случайное) формирование сравниваемых групп случаев и контролей. Аргументационное смещение (Confirmation bias или myside bias) — стремление исследователей предпочитать только ту информацию, которая подтверждает их аргументацию или гипотезы, игнорируя опровергающие данные или принижая их («тянуть одеяло на себя»). «Удачная своевременная» остановка сбора данных: сбор данных прекращают, когда достигнут результат, близкий к желаемому. Но если продолжить или повторить сбор данных, то эффект может оказаться меньшим или вообще исчезнуть.

Повторение — мать познания, или доверия, проверй.

Изучение воспроизводимости результатов исследований и их предсказательной способности составляет суть науки, без него исследование вряд ли можно признать научным. Мировое

научное сообщество взбудоражено и озабочено многочисленными научными публикациями, неоспоримо свидетельствующими о плохой воспроизводимости результатов биомедицинских исследований. На страницах научных журналов и в блогах разыгрываются нешуточные баталии. Кризис воспроизводимости обсуждается как системная проблема современной биомедицины. Одной из основных причин этого кризиса является кризис биостатистики.

Традиционно статистической идеологией и методологией остается проверка статистической значимости нулевой гипотезы H_0 и так называемое p -значение как финальный результат этой процедуры. Для проверки гипотезы H_0 выбирают такую случайную переменную (статистику критерия), распределение которой известно, когда H_0 истинна (например, t , F , χ^2 , U , r , z и др.). Это позволяет вычислять долю тех из всех возможных значений статистики критерия, которые равны наблюдаемому и еще менее вероятны (или еще более отклоняются от ожидаемого значения согласно H_0). Такую долю и называют p -значением.

Синдром статистической снисходительности и доверчивости

В течение последнего столетия исследователи от «правила трех сигм» снизошли до «правила двух сигм». Одной из причин кризиса воспроизводимости является повальное пренебрежение интервальными оценками p -значений. Подавляющее большинство исследователей-пользователей не знают (или забывают), что p -значение есть реализация соответствующей случайной переменной P . Легко понять, что если многократно повторять данное исследование, то получаемые (наблюдаемые) p_{obs} будут варьироваться. Это означает, что такое их варьирование должно подчиняться некоторому вероятностному закону, т. е. иметь определенное распределение вероятностей. Одним из основных приемов анализа («обработки») получаемых данных является статистическое оценивание параметров распределения, из которого извлечены эти данные. Процедура статистического оценивания параметров заключается в вычислении двух типов оценок — точечной (выражаемой одним числом) и интервальной (которая отображается двумя числами — границами интервала). Существуют два основных типа интервальных оценок: предсказательные интервалы (ПИ) и доверительные интервалы (ДИ). В биомедицинских исследованиях ПИ играют важную роль и известны под названием референсных интервалов (РИ). В случае с p -значениями $(1 - \alpha)100\%$ -ный ПИ с предсказательной вероятностью $(1 - \alpha)100\%$

есть интервал, в который с вероятностью $(1 - \alpha)100\%$ попадет новое p -значение, полученное после однократного повторения исследования.

Как известно, p -значения ничего не говорят ни о вероятности проверяемой нулевой гипотезы, ни о размере эффекта, ни о вероятности воспроизведения полученного результата. Поэтому в руководствах и методических статьях по статистике давно рекомендуется наряду с p -значениями или вместо них указывать ДИ для размера эффекта, например для разности сравниваемых параметров. **Современный статистический анализ не должен ограничиваться сообщением одного-единственного показателя согласия с нулевой гипотезой, в качестве которого наиболее часто выступает p -значение, даже если оно скорректировано на множественность сравнений. Как уже было сказано, следует указывать не только и не столько ДИ, но и ПИ для оцениваемого эффекта. Пользователи-исследователи должны знать, что традиционно и повсеместно используемые пороговые уровни значимости $\alpha = 0,05$ и $0,01$ обладают неприемлемой предсказательной способностью. Этим объясняется, почему ведущие биостатистики рекомендуют ориентироваться на значение $\alpha = 0,005$ и ниже. В геномике давно уже используют уровни значимости 10^{-7} и даже 10^{-8} .**

Непредсказуемость генетических предрасположенностей

Плохая воспроизводимость экспериментальных результатов осознана как системная проблема биомедицины. Причин этого может быть несколько: 1) образцы исследованы без использования приема ослепления; 2) в отчете представлены не все результаты; 3) эксперименты не были повторены; 4) использованы не сертифицированные реагенты (например, поликлональные антитела); 5) не приведены результаты позитивных и негативных контролей; 4) использованы неадекватные статистические методы и др.

Неадекватный статистический анализ является, возможно, одной из основных из перечисленных причин. Статистический анализ должен быть исчерпывающим и гармонично сочетающимся как статистические доказательства и предсказания, так и частотнический и байесовский подходы. Недостаточно проводить проверку значимости нулевой гипотезы, сообщая лишь p -значения.

Статистическая значимость не должна восприниматься как синоним клинической значимости. Поэтому не следует ограничиваться сообщением в публикациях только наблюдаемых p -значений. Следует указывать значения размера

эффекта с доверительными и предсказательными интервалами и сопровождать их значениями байесовых факторов BF_{01} или $BF_{10} = 1/BF_{01}$, которые интуитивно более понятны, чем p -значения. Наблюдаемый на наших глазах переход (или даже возврат) на рельсы байесовской статистики отнюдь не означает полного отрицания частотнического подхода. Оба они являются взаимодополняющими (комплементарными), они должны применяться совместно.

Размер эффекта с доверительными и предсказательными интервалами

Критериями научной истины являются доказательная и предсказательная способности аналитических инструментов. Возможность предсказания в науке представляет ее сущность. Ведущую роль в экспериментальной науке играет теория эксперимента, которая базируется на вероятностно-статистической методологии. Соответственно, вероятностно-статистические предсказания представляются чрезвычайно важными с практической точки зрения. Статистика изучает повторяющиеся (взаимозаменяемые по де Финетти) данные. Повторяемость составляет суть статистики, и это оснащает ее предсказательной способностью. Современная статистика способна оценивать не только то, что есть, но и то, что будет, т. е. как настоящая наука она обладает предсказательной силой. Проблема предсказаний при анализе статистических данных представляется исключительно важной при практическом применении.

Часто считается, что если получен «статистически значимый» результат, то он обязательно воспроизведется, и потому нет необходимости повторять исследование. Повторность (воспроизведение) часто рассматривается как нечто суетное и мирское. Однако эффект (различие, связь, риск, польза, ассоциация и т. п.) может быть статистически значимым, но его практическая (например, клиническая) ценность может оказаться ничтожной. «Статистически значимый» не означает «значительный», «практически (клинически) важный», «ценный». Вопрос о клинической (практической) ценности (важности) наблюдаемого размера эффекта является ключевым при интерпретации результатов биомедицинских исследований, таких как диагностические исследования, клинические испытания и т. п. Размер эффекта можно выражать в реальных единицах, а можно сделать его безразмерным — стандартизованным.

Статистическая значимость не должна восприниматься как синоним клинической значимости. Следует указывать размер эффекта с доверительными и предсказательными интерва-

лами. Эксперименты и/или наблюдения следует повторять многократно и исследовать их воспроизводимость. Лучше всего повторять эксперименты независимо в разных лабораториях (и в разных странах).

Воспроизводимость и предсказательность

Научные выводы воистину становятся научными, когда они обладают предсказательной силой или прогностической (эвристической) способностью. Выводы, сделанные на основе анализа единичной повторности эксперимента, не могут обладать предсказательной способностью. В традиционных руководствах и учебных пособиях по статистике эта проблема даже не упоминается. «Повторение составляет суть науки: ученый должен всегда задумываться о том, что произойдет, если он или другой ученый повторят его эксперимент» (Guttman). «Ученые разработали метод определения надежности (валидности) своих результатов. Они научились задавать вопрос: воспроизводимы ли они?» (Scherr).

Исследователь почти всегда вынужден решать, повторять ли ему эксперимент или нет. Первую помощь в принятии того или иного решения могут оказать вероятности воспроизведения P_{rep} и $P_{\text{сrep}}$. Однако в настоящее время существуют более изощренные предсказательные процедуры, которые позволяют принимать более тонкие решения и получать ответы, например, на такие вопросы: «Какова вероятность того, что при повторении опыта наблюдаемый эффект будет иметь то же направление?» и «Если наблюдаемый эффект статистически значим, то какова вероятность того, что при повторении результат также будет статистически значимым (хотя бы на том же уровне значимости)?». Такие вопросы и ответы на них выходят далеко за рамки ритуальных статистических процедур и влияют на способы, которыми авторы научных работ интерпретируют получаемые экспериментальные данные и проводят свои эксперименты. Предсказательные вероятности являются (или должны стать) неизбежной и неотъемлемой частью статистического мышления, и настало время применять и воспринимать их серьезно. Современные программы позволяют вычислять ПИ для размера эффекта и для p -значения.

Идея предсказания становится центральной в биомедицинских исследованиях. Исследование без предсказательной силы скорее всего окажется бесперспективным, бесплодным и ущербным.

Таким образом, явно не следует слепо применять инструментарий процедур проверки значи-

мости нулевой гипотезы и основывать свои выводы исключительно на получаемых p -значениях. Осмысленные выводы должны основываться на разумном взвешивании p -значений и на использовании дополнительной информации о других не менее важных показателях, таких как Бейзов фактор, мощность, размер эффекта, предсказательные вероятности и предсказательные интервалы.

Бейзовский и частотнический (ортодоксальный, классический) подходы взаимодополнительны (комплементарны), и совместно они обеспечивают значительно лучшее понимание статистических проблем, чем каждый из них в отдельности.

Для описания статистических методов и программ в генетических исследованиях предлагается следующая табличная форма:

Таблица 15

Программное обеспечение статистического анализа генетических данных

Программа	Версия и/или дата	Использованные процедуры и методы	URL
StatXact	12.0 2018	Точные непараметрические методы, 160 методов и процедур	https://www.cytel.com/software/statxact
Statistical tests	2015	Проверка согласия с PXB, точный критерий, mid- p -значения точного критерия Фишера для таблиц 2x2 and 2x3	https://www.cog-genomics.org/software/stats
GenAIEx	6.51b2 06.2018	Разносторонний генетико-популяционный анализ для широкого спектра генетических маркеров. Анализ последовательностей ДНК. F -статистики, PXB, родство, совпадающие профили [60]	http://biology-assets.anu.edu.au/GenAIEx/Welcome.html
GENEPOP	4.7.2 23.06.2019	Точные критерии согласия с PXB, популяционной дифференциации и генотипической неравновесности [70]	https://kimura.univ-montp2.fr/~rousset/Genepop.html
Arlequin	3.5.2.2 02.08.2015	Анализ генетико-популяционных данных [88]	http://cmpg.unibe.ch/software/arlequin35/
G*Power	3.1.9.4 06.02.2019	Оценка достигнутой мощности критерия и необходимых объемов выборок [89]	http://www.gpower.hhu.de/
HWMET	1998	Бейзовская интервальная оценка индекса фиксации F_{is} [72, 73]	http://www.reading.ac.uk/Statistics/genetics/software.html
hwde	0.67 01.01.2016	Критерии согласия с PXB и независимости между локусами	https://cloud.r-project.org/web/packages/hwde/index.html
Hardy Weinberg	1.6.3 29.06.2019	Исследование PXB для би- и полиаллельных маркеров. Критерии: хи-квадрат, точные отношения правдоподобия и перестановочные, вычисление мощности	https://cloud.r-project.org/
HWEBayes	1.4 18.12.2013	Оценки и проверки согласия с PXB бейзовскими методами	https://cloud.r-project.org/web/packages/HWEBayes/index.html
HWEIntrinsic	1.2.2 08.07.2015	Объективное бейзовское тестирование PXB	https://cloud.r-project.org/web/packages/HWEIntrinsic/index.html
FPtest	20.12.2012	Анализ ассоциаций в группах случаев и контролей [71, 90]	http://www.jurgott.org/linkage/FPtest.html
DiagStat	—	Статистический контроль качества диагностических тестов с бинарными исходами, графики предсказательности	https://www.researchgate.net/publication/332727723_DiagStatRUS
MetaboAnalyst	4.0 23.07.2019	Метабономика, ROC-анализ, метаанализ, анализ мощности и др. [87]	https://www.metaboanalyst.ca/faces/home.xhtml
EasyROC		ROC-анализ [91]	http://www.biosoft.hacettepe.edu.tr/easyROC/

Редакционная политика

Желательно, чтобы редакторы научных журналов, рецензенты и эксперты не принимали к публикации статьи, если в них сообщаются результаты одиночных исследований генетических тестирований и отсутствуют результаты повторных независимых испытаний. В состав редколлегий должны входить эксперты по статистике. Рецензентов желательно обязывать проверять результаты вычислений. По этой причине следует обеспечивать свободный доступ к исходным («сырым») данным. Прозрачность и открытость являются краеугольным камнем научных исследований.

Программа	Версия и/или дата	Использованные процедуры и методы	URL
Episheet	20.10.2015	Анализ эпидемиологических данных, мета-анализ	http://www.drugapi.org/wp-content/uploads/2012/10/Episheet.xls
Meta-Essentials	1.4	MetaEssentials [82]	https://www.erim.eur.nl/research-facilities/metaessentials/
Effect Size Calculators	2017	Калькуляторы размеров эффекта [92]	https://effect-size-calculator.herokuapp.com/
Chi-square Test	1.0 2015	Составление таблиц сопряженности, их анализ и мозаичные диаграммы [64]	http://langtest.jp/shiny/chi/
JASP	0.10.2 17.07.2019	Программы нового поколения с гармонизированным сочетанием традиционных и байзовских подходов [63, 93]	https://jasp-stats.org/
jamovi	0.9.2.3 18.08.2018		https://www.jamovi.org/
LePAC	2.0.41	Бейзовский анализ долей и таблиц сопряженности	http://lmrs.univ-rouen.fr/Persopage/Lecoutre/PAC.htm
LePrep	2.1.0 01.04.2011	Вероятности воспроизведения и доверительные (ДИ) и предсказательные интервалы для размеров эффекта и p -значений	
MOVER-D	—	ДИ для разности долей на основе индивидуальных ДИ	http://profrobertnewcomberesources.yolasite.com/
MOVER-R	—	ДИ для отношения долей на основе индивидуальных ДИ	
Prediction Intervals	—	Предсказательные интервалы для мета-анализа	https://medstats.github.io/ranefpredict.html
Reference Value Advisor	2.1 09.12.2015	Референсные (предсказательные интервалы) с 95%-ными ДИ для их границ [94]	http://www.biostat.envt.fr/reference-value-advisor/
PAST	3.26 08.2019	Непараметрические методы на основе бутстрепа и Монте-Карло [95]	https://folk.uio.no/ohammer/past/
BoxPlotR	11.01.2017	Построение коробов с декольте и с усами	http://shiny.chemgrid.org/boxplotr/
RegressIt	24.02.2019	Описательный анализ многомерных данных и линейная регрессия	https://regressit.com/index.html
Bayes Factor Calculators	06.2010	Бейзовы факторы BF для парного и двухвыборочного t -критериев, регрессии и биномиального критериев	http://pcl.missouri.edu/bayesfactor
RNG	1998–2019	Генерирование истинно случайных чисел	https://www.random.org/
p-value Empirical Null calibration	—	Калибровка p -значений	https://stefano-cabras.shinyapps.io/p-value_calibration/
AgreeStat	2015.6.1	Коэффициенты надежности межэкспертной согласованности	http://agreestat.com
Real Statistics Using Excel	6.5	Большое разнообразие современных методов	http://www.real-statistics.com/
FDR	2011	Оценка доли ложных отклонений (FDR – False Discovery Rate) нулевой гипотезы при множественных сравнениях [96]	https://besjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fj.2041-210X.2010.00061.x&file=MEE3_61_sm_Supmat.xls
Estimation statistics	2019	Оценка стандартизированного размера эффекта и его визуализация [97]	https://www.estimationstats.com/#/
WINPEPI	11.65 23.08.2016	122 модуля статистических процедур [98]	http://www.brixtonhealth.com/pepi4windows.html
Statistical Intervals for P-value Evidence and Power	2.0 2016	Предсказательные интервалы для p -значений [32]	https://media.nature.com/original/nature-assets/nmeth/journal/v13/n2/extref/nmeth.3741-S2.xls
AtteStat	2017	Универсальный пакет статистических программ	https://sourceforge.net/projects/attestat/

Список литературы

1. Sivakumaran S., Agakov F., Theodoratou E., Prendergast J. G., Zgaga L., Manolio T. et al. Abundant pleiotropy in human complex diseases and traits // *Am. J. Hum. Genet.* 2011; 89(5): 607–618. doi: 10.1016/j.ajhg. 2011.10.004 (383 цитирования).
2. Nawaz S. K., Hasnain S. Pleiotropic effects of ACE polymorphism. *Biochem. Med.* 2009; 19(1): 36–49. doi: org/10.11613/bm.2009.004.
3. Li T., Shen X. Pleiotropy complicates human gene editing: CCR5Δ32 and beyond // *Front. Genet.* 2019; 10:669. doi: 10.3389/fgene.2019.00669.
4. Luban J. The hidden cost of genetic resistance to HIV-1 // *Nat. Med.* 2019; 25(6): 878–883. doi: 10.1038/s41591-019-0481-8.
5. Mauvais-Jarvis F. Estrogen and androgen receptors: regulators of fuel homeostasis and emerging targets for diabetes and obesity // *Trends Endocrinol. Metab.* 2011; 22(1): 24–33. doi: 10.1016/j.tem.2010.10.002.
6. Алексеев И. В., Плешкан В. В., Кузьмич А. И., Снежков Е. В., Монастырская Г. С., Дидыч Д. А., Свердлов Е. Д. Принципиально низкая воспроизводимость молекулярно-генетических исследований рака // *Генетика.* 2016. Т. 52. № 7. С. 745–760. doi: 10.7868/s0016675816070031.
7. Webb N., Williams A., McNamee M., Bouchard C., Pitsiladis Y., Ahmetov I. et al. Direct-to-consumer genetic testing for predicting sports performance and talent identification: Consensus statement // *Br. J. Sports Med.* 2015; 4(9): 1486–1491. doi: 10.1136/bjsports-2015-095343.
8. Vlahovich N., Fricker P. A., Brown M. A., Hughes D. Ethics of genetic testing and research in sport: a position statement from the Australian Institute of Sport // *Br. J. Sports Med.* 2017; 51(1): 5–11. doi: 10.1136/bjsports-2016-096661.
9. Loland S. Against genetic tests for athletic talent: The primacy of the phenotype // *SportsMed.* 2015; 45(9): 1229–1233. doi: 10.1007/s40279-015-0352-5.
10. Корбут А. Генетический тест или маркетинговая уловка? Ученый-генетик объясняет, почему генетические тесты могут быть бесполезными, а рекомендации по их результатам — бессмысленными // *НОЖ. Интеллектуальный журнал о культуре и обществе.* 09.09.2019. <https://knife.media/gen-test/>
11. Lenzini P., Wadelius M., Kimmel S., Anderson J. L., Jorgensen A. L., Pirmohamed M. et al. Integration of genetic, clinical, and INR data to refine warfarin dosing // *Clin. Pharmacol. Ther.* 2010; 87(5): 572–578. doi: 10.1038/clpt. 2010.13 (215 цитирований).
12. Kimmel S. E., French B., Kasner S. E., Johnson J. A., Anderson J. L., Gage B. F. et al. A pharmacogenetic versus a clinical algorithm for warfarin dosing // *N. Engl. J. Med.* 2013; 369(24): 2283–2293. doi: 10.1056/NEJMoa1310669 (609 цитирований).
13. Verhoef T. I., Ragia G., de Boer A., Barallon R., Kolovou G., Kolovou V. et al. A randomized trial of genotype-guided dosing of acenocoumarol and phenprocoumon // *N. Engl. J. Med.* 2013; 369(24): 2304–12. doi: org/10.1056/nejmoa1311388.
14. Johnson J. A., Cavallari L. H. Warfarin Pharmacogenetics // *Trends Cardiovasc. Med.* 2015; 25(1): 33–41. doi: 10.1016/j.tcm.2014.09.001.
15. Horne B. D., Lenzini P. A., Wadelius M., Jorgensen A. L., Kimmel S. E., Ridker P. M. et al. Pharmacogenetic warfarin dose refinements remain significantly influenced by genetic factors after one week of therapy // *Thromb. Haemost.* 2012; 107(2): 232–40. doi: org/10.1160/th11-06-0388.
16. Saffian S. M., Wright D. F. B., Roberts R. L., Duffull S. B. Methods for Predicting Warfarin Dose Requirements // *Ther. Drug Monit.* 2015; 37: 531–538. doi: org/10.1097/ftd.000000000000177.
17. Johnson J. A., Caudle K. E., Gong L., Whirl-Carrillo M., Stein C. M. et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for pharmacogenetics-guided warfarin dosing: 2017 Update // *Clin. Pharmacol. Ther.* 2017; 102(3): 397–404. doi: 10.1002/cpt.668.
18. Bousman C. A., Dunlop B. W. Genotype, phenotype and medication recommendation agreement among commercial pharmacogenetic-based decision support tools // *Pharmacogenomics J.* 2018; 18(1): 613–622. doi: 10.1038/s41397-018-0027-3.
19. Editorial. Freely associating // *Nature Genet.* 1999; 22(1): 1–2. doi: 10.1038/8702.
20. Хромов-Борисов Н. Н. Гармонизация статистических доказательств и предсказаний в биомедицине: сб. ст. Открытого российского статистического конгресса. Новосибирск: НГУЭУ, 2016. 499 с. ISBN 978-5-7014-0730-3.
21. Королёв В. Ю. Теория вероятностей и математическая статистика // М.: ТК Велби; Изд-во Проспект, 2008. 160 с. ISBN 978-5-482-01946-7.
22. GraphPad Software. GraphPad Prism 8 Statistics Guide. Asterisks. Graphpad.com. https://www.graphpad.com/guides/prism/8/statistics/extremely_significant_results.htm. Published 2019. Accessed September 10, 2019.
23. Рубанович А. В. Пересмотр критического уровня значимости (0,005 вместо 0,05): байесовский след // *Радиационная биология, радиоэкология.* 2018. Т. 58. № 5. С. 453–462. DOI: 10.1134/S0869803118050156.
24. Benjamin D. J., Berger J. O., Johannesson M., Nosek B. A., Wagenmakers E.-J., Berk R. et al. Redefine statistical significance // *Nat. Hum. Behav.* 2018; 2(1): 6–10. doi: 10.1038/s41562-017-0189-z. (674 цитирования)
25. P-value Empirical Null calibration. stefano-cabras.shinyapps.io. https://stefano-cabras.shinyapps.io/p-value_calibration/. Published 2019. Accessed September 10, 2019.
26. Cabras S., Castellanos M. E. P-value calibration in multiple hypotheses testing // *Stat. Med.* 2017; 36(18): 2875–2886. doi: 10.1002/sim.7330.
27. Глотов Н. В., Животовский Л. А., Хованов Н. В., Хромов-Борисов Н. Н. Биометрия. Л.: ЛГУ, 1982. 264 с.
28. Хромов-Борисов Н. Н. Биостатистические программы свободного доступа // *Травматология и ортопедия России.* 2015. Т. 4. № 78. С. 54–159.
29. Colquhoun D. The reproducibility of research and the misinterpretation of p-values // *R. Soc. Open Sci.* 2017; 4(12): 171085. doi: 10.1098/rsos.171085.
30. Colquhoun D. The false positive risk: A proposal concerning what to do about p-values // *Am. Stat.* 2019; 73(1): 192–201. doi: 10.1080/00031305.2018.1529622.
31. Held L., Ott M. On p-values and Bayes factors // *Annu. Rev. Stat. Appl.* 2018; 5: 393–419. doi: 10.1146/annurev-statistics-031017-100307.
32. Lazzeroni L. C., Lu Y., Belitskaya-Levy I. Solutions for quantifying P-value uncertainty and replication power // *Nature Methods.* 2016; 13(2): 107–108. doi: 10.1038/nmeth.3741.
33. Vsevolozhskaya O., Ruiz G., Zaykin D. Bayesian prediction intervals for assessing P-value variability in prospective replication studies // *Transl. Psychiatry.* 2017; 7(12): 1271. doi: 10.1038/s41398-017-0024-3.
34. Media.nature.com. <https://media.nature.com/original/nature-assets/nmeth/journal/v13/n2/extref/nmeth.3741-S2.xls>. Published 2019. Accessed September 10, 2019.
35. LePAC. <http://lmrs.univ-rouen.fr/Persopage/Le-cou-tre/PAC.htm>. Published 2019. Accessed September 10, 2019.

36. Bayesian Calibrations of P-Values [R package pCalibrate version 0.1–1]. Cran.r-project.org. <https://cran.r-project.org/web/packages/pCalibrate/index.html>. Published 2019. Accessed September 10, 2019.
37. bayesian-PValue-Prediction-Intervals. GitHub. <https://github.com/dmitri-zaykin/bayesian-PValue-Prediction-Intervals>. <https://github.com/gabriel-ruiz/bayesian-PValue-Prediction-Intervals>. Published 2019. Accessed September 10, 2019.
38. False Positive Risk calculator (FPR). Fpr-calc.ucl.ac.uk. <http://fpr-calc.ucl.ac.uk>. Published 2019. Accessed September 10, 2019.
39. Sawilowsky S. S. New effect size rules of thumb // J. Mod. Appl. Stat. Methods. 2009; 8(2): 597–599. doi: org/10.22237/jmasm/1257035100. 587 цитирований.
40. Hopkins W. G. A scale of magnitudes for effect statistics. A new view of statistics // URL <http://sportsci:resource/stats/effectmag.html>.
41. Cohen J. Statistical power analysis for the behavioral sciences // NJ: Erlbaum, 1998. 567 p. <http://www.utstat.toronto.edu/~brunner/oldclass/378f16/readings/CohenPower.pdf>.
42. Reichel C. Statistics for journalists: Understanding what effect size means // Journalists's Resources, June 25, 2019. <https://journalistsresource.org/tip-sheets/research/effect-size-statistics-risk-ratio/>
43. Гржибовский А. М., Иванов С. В., Горбатова М. А. Анализ номинальных и ранговых переменных данных с использованием программного обеспечения Statistica и SPSS // Наука и Здоровоохранение. 2016. Т. 6. № 1. С. 5–39.
44. Phi Coefficient (Mean Square Contingency Coefficient) — Statistics How To. <https://www.statisticshowto.datasciencecentral.com/phi-coefficient-mean-square-contingency-coefficient>. Published 2019. Accessed September 10, 2019.
45. FAQ/effectSize — CBU statistics Wiki. Imaging.mrc-cbu.cam.ac.uk. <http://imaging.mrc-cbu.cam.ac.uk/statwiki/FAQ/effectSize>. Published 2019. Accessed September 10, 2019.
46. Kotrlík J. W., Williams H. A. The incorporation of effect size in information technology, learning, and performance research // Information Technology, Learning, and Performance Journal. 2003; 21(1): 1–7.
47. Kotrlík J. W., Atherton J. C., Williams H. A., Jabor K. Reporting and interpreting effect size in quantitative agricultural education research // J. Agric. Educ. 2011; 52(1): 132–142. doi: 10.5032/jae.2011.01132 (377 цитирований).
48. https://www.researchgate.net/post/How_do_you_interpret_the_odds_ratio_OR. Published 2019. Accessed September 10, 2019.
49. Рубанович А. В., Хромов-Борисов Н. Н. Теоретический анализ показателей предсказательной эффективности бинарных генетических тестов // Экол. генетика. 2013. Т. 9. № 1. С. 77–90.
50. Насырова Р. Ф., Иванов М. В., Незнанов Н. Г. Введение в психофармакогенетику // СПб.: Издательский центр СПб НИПНИ им. В. М. Бехтерева, 2015. 272 с. ISBN 978-5-7452-0020-5.
51. Hoppe F. M., Daniel J., Hoppe D. J., Walter S. D. Odds ratios deconstructed: A new way to understand and explain odds ratios as conditional risk ratios // J. Clin. Epidemiol. 2017; 82: 87–93. doi: 10.1016/j.jclinepi.2016.08.007.
52. Тишкова В., Хромов-Борисов Н. Н., Комашня А. В., Марченкова Ф. Ю., Семенова Е. М., Эюбова Н. И. и др. Статистический анализ таблиц 2×2 в диагностических исследованиях. СПб.: Издательство СПбГМУ, 2013. 20 с.
53. Schield M. Confound and Confounding: back to the future // Sorto M. A., White A., Guyot L. (Eds.) Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan. iase-web.org.
54. Ioannidis J. Commentary: Grading the credibility of molecular evidence for complex diseases // Int. J. Epidemiol. 2006; 35(3): 572–577.
55. Хромов-Борисов Н. Н., Рубанович А. В. Генетика предрасположенностей — разбитые мечты и утраченные грезы. Труды 3-й Московской международной научно-практической конференции // Иммунофизиология: аутоиммунитет в норме и патологии и вопросы предиктивно-превентивной медицины. М.: 1–3 октября г. 388 с. http://www.spsl.nsc.ru/FullText/konfe/3IMPH_2012.pdf.
56. Edgren G., Hjalgrim H., Rostgaard K., Nordsa R., Wikman A., Melbye M., Nyré O. Risk of gastric cancer and peptic ulcers in relation to ABO blood type: a cohort study // Am. J. Epidemiol. 2010; 172(11): 1280–1285. doi: 10.1093/aje/kwq299.
57. S2 ABOestimator. Webpages.fc.ul.pt. <http://webpages.fc.ul.pt/~pjns/Soft/ABOestimator/>. Published 2019. Accessed September 10, 2019.
58. Liu J. PowerMarker homepage. Brcwebportal.cos.ncsu.edu. <https://brcwebportal.cos.ncsu.edu/powermarker/index.html>. Published 2019. Accessed September 10, 2019.
59. Pagnotta M. A. Comparison among methods and statistical software packages to analyze germplasm genetic diversity by means of codominant markers // J. 2018; 1: 197–215. doi: 10.3390/j1010018.
60. Peakall R., Smouse P. E. Genetic analysis in Excel. Population genetic software for teaching and research — an update // Bioinformatics. 2012; 28(19): 2537–2539. doi: 10.1093/bioinformatics/bts460 (16255 цитирований).
61. Вохмянина Н. В. Опыт HLA-генотипирования больных целиакией. Диагностическая значимость HLA-DQ2 и HLA-DQ8 // Медицинская генетика. 2010. Т. 9. № 7. С. 33–41.
62. StatXact® 12. Cytel.com. <https://www.cytel.com/software/statxact>. Published 2019. Accessed September 10, 2019.
63. JASP — A Fresh Way to Do Statistics. JASP — Free and User-Friendly Statistical Software. <https://jasp-stats.org/>. Published 2019. Accessed September 10, 2019.
64. Chi-square Test. Langtest.jp. <http://langtest.jp/shiny/chi/>. Published 2019. Accessed September 10, 2019.
65. Плоткин В. Я., Иващенко Т. Э., Воронель В. Л., Зарипова З. А., Азанчевская С. В., Хромов-Борисов Н. Н. Острый период инфаркта миокарда и генетическая предрасположенность. Сообщение I. Полиморфизмы гена фактора некроза опухолей альфа (TNFA) // Вестник СПбГУ. Сер. 11. 2007. Т. 3. № 1. С. 1–21.
66. Хромов-Борисов Н. Н. Приложение: Биометрические аспекты популяционной генетики // Кайданов Л. З. Генетика популяций. М.: Высшая школа, 1996. 320 с. ISBN 5-06-002575-6. www.academia.edu/1622053/Biometrical_aspects_of_population_genetics.
67. Weedon M. N., Jackson L., Harrison J. W., Ruth K. S., Tyrrell J., Hattersley A. T., Wright C. F. Very rare pathogenic genetic variants detected by SNP-chips are usually false positives: implications for direct-to-consumer genetic testing // <https://www.biorxiv.org/content/10.1101/696799v1>.
68. Statistical tests. Cog-genomics. <https://www.cog-genomics.org/software/stats>. Published 2019. Accessed September 10, 2019.
69. Глотов А. С. Генетические и средовые факторы риска развития гестоза у женщин, артериальной гипертензии и метаболического синдрома у детей: дисс. д-ра биол. наук. 03.02.07. СПб., 2017. 420 с. disser.spbu.ru/disser/dissertatsii-dopushchennyye-k-zashchite-i-svedeniya-o-zashchite/details/12/1309.html.
70. Rousset F. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol. Ecol. Resour. 2008; 8(1):103–106. <https://doi.org/10.1111/j.1471-8286.2007.01931.x> (6934 цитирования).

71. Jurgott. org. <http://www.jurgott.org/linkage/FPtest.html>. Published 2019. Accessed September 10, 2019.
72. Ayres K. L., Balding D. J. Measuring departures from Hardy-Weinberg: a Markov chain Monte Carlo method for estimating the inbreeding coefficient // *Heredity*. 1998; 80(6): 769–777. doi: 10.1046/j.1365-2540.1998.00360.x.
73. Statistical Genetics Group — Software. Reading.ac.uk. <http://www.reading.ac.uk/Statistics/genetics/software.html>. Published 2019. Accessed September 10, 2019.
74. Anderson C. A., Pettersson F. H., Clarke G. M., Cardon L. R., Morris A. P., Zondervan K. T. Data quality control in genetic case-control association studies // *Nat. Protoc.* 2010; 5(9): 1564–73. doi: 10.1038/nprot.2010.116. (647 цитирований).
75. Chen B., Cole J. W., Grond-Ginsbach C. Departure from Hardy Weinberg equilibrium and genotyping error // *Front. Genet.* 2017; 8: 167. doi: 10.3389/fgene.2017.00167.
76. Xu J., Turner A., Little J., Bleeker E. R., Meyers D. A. Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? // *Hum. Genet.* 2002; 111(6): 573–574. doi: 10.1007/s00439-002-0819-y.
77. Zintzaras E. Impact of Hardy-Weinberg equilibrium deviation on allele-based risk effect of genetic association studies and meta-analysis // *Eur. J. Epidemiol.* 2010; 25(8): 553–560. doi: 10.1007/s10654-010-9467-zff.
78. Шнайдер Н. А., Дмитренко Д. В., Говорина Ю. Б., Муравьева А. В., Котловский Ю. В., Бочанова Е. Н. и др. Влияние полиморфизмов гена CYP2C9 на уровень вальпроевой кислоты в крови у женщин репродуктивного возраста с эпилепсией // *Фармакогенетика и фармакогеномика*. 2015. № 2. С. 24–28.
79. Strachan T., Read A. *Human Molecular Genetics*. 4th ed., Garland Science. 2011. 812 p.
80. Bayes Factor Calculators. Perception and Cognition Lab. Pcl.missouri.edu. <http://pcl.missouri.edu/bayesfactor>. Published 2019. Accessed September 10, 2019.
81. Bousman C. A., Arandjelovic K., Mancuso S. G., Dunlop B. W., Eyre H. A. Pharmacogenetic tests and depressive symptom remission: a meta-analysis of randomized controlled trials // *Pharmacogenomics*. 2019; 20(1): 37–47. doi: 10.2217/pgs-2018-0142.
82. Suurmond R., van Rhee H., Hak T. Introduction, comparison, and validation of *Meta-Essentials*: A free and simple tool for meta-analysis // *Res. Syn. Meth.* 2017; 8(4): 537–553. DOI: 10.1002/jrsm.1260.
83. Hak T., Van Rhee H. J., Suurmond R. How to interpret results of meta-analysis // Netherlands: Erasmus Rotterdam Institute of Management. 2016. www.irim.eur.nl/researchsupport/meta-essentials/downloads.
84. Березовская Г. А., Смирнова О. А., Петрищев Н. Н., Папаян Л. П., Карпенко М. А., Головина О. Г., Хромов-Борисов Н. Н. Тест генерации тромбина в оценке действия антиагрегантов у больных ишемической болезнью сердца после чрескожного коронарного вмешательства // *Атеротромбоз*. 2015. № 1. С. 40–51. doi: 10.21518/2307-1109-2015-1-40-51.
85. Berezovskaya G., Smirnova O., Malev E., Khromov-Borisov N., Klokov E., Karpenko M. et al. Thrombin generation test for evaluation of antiplatelet treatment in patients with coronary artery disease after percutaneous coronary intervention // *Platelets*. 2018; 29(2): 185–191. doi: 10.1080/09537104.2017.1294680.
86. Castoldi E., Rosing J. Thrombin generation tests // *Thrombosis Res.* 2011; 127(3): 21–25, doi: 10.1016/S0049-384S(11)70007-X (138 цитирований).
87. Chong J., Soufan O., Li C., Caraus I., Li S., Bourque G., Wishart D. S., Xia J. *MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis*. *Nucl. Acids Res.* 2015; 46: W486–494 (393 цитирования).
88. Excoffier, Lischer H. E. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*. 2010; 10: 564–567 (10245 цитирований).
89. Faul F., Erdfelder E., Lang A.-G., Buchner A. G. Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences // *Behavior Research Methods*, 2007; 39(2): 175–191. <https://doi.org/10.3758/BF03193146> (24 587 цитирований).
90. Zhang Q., Wang S., Ott J. Combining identity by descent and association in genetic case-control studies // *BMC Genet.* 2008; 9:42. <https://doi.org/10.1186/1471-2156-9-42>.
91. Goksuluk D., Korkmaz S., Zararsiz G., Karaoglu A. E. (2016). easyROC: An interactive Web-tool for ROC curve analysis using R language environment // *The R Journal*, 2019; 8(2): 213–230. doi: 10.32614/RJ-2016-042.
92. Uanboro J. O. (2017). Effect size calculators. Available online at: <https://effect-size-calculator.herokuapp.com/>.
93. The jamovi project (2019). *jamovi* (Version 0.9) [Computer Software].
94. Geffré A., Concordet D., Braun J.-P., Trumel C. Reference Value Advisor: a new freeware set of macroinstructions to calculate reference intervals with Microsoft Excel // *Vet. Clin. Pathol.*, 2011; 40(1)07–112. doi: 10.1111/j.1939-165X.2011.00287.x (246 цитирований).
95. Hammer O., Harper D. A. T., Ryan P. D. PAST: Palaeontological statistics software package for education and data analysis // *Palaeontologia Electronica*, 2001; 4(1): 9pp. http://palaeo-electronica.org/2001_1/past/issue1_01.htm (12 439 цитирований).
96. Pike N. Using false discovery rates for multiple comparisons in ecology and evolution // *Methods Ecol. Evol.* 2011; 2(30): 278–282. doi: 10.1111/j.2041-210X.2010.00061.x (330 цитирований).
97. Ho J., Tumkaya T., Aryal S., Choi H., Claridge-Chang A. Moving beyond P values: Everyday data analysis with estimation plots // *Nature Methods*. 2019; 16(7): 1548–7105. DOI: 10.1038/s41592-019-0470-3.
98. Abramson J. H. WINPEPI updated: computer programs for epidemiologists, and their teaching potential // *Epidemiol. Perspect. Innov.* 2011; 8(1): 1–9. DOI: 10.1186/1742-5573-8-1 (504 цитирования).