

Четырехпольная таблица частот – Бритва Оккама в мире статистики

Часть 1. Как рассчитывать относительный риск и другие параметры из четырехпольной частотной таблицы

Г. П. Тихова

МИА «ИнтелТек Медиа», Петрозаводск

Fourfold frequency table – Occam's Razor in the world of statistics. Part 1. Calculating relative risk and other parameters from fourfold frequency table

G. P. Tikhova

MIA IntelTeck Media, Petrozavodsk

В медицине достаточно часто возникают задачи, в которых необходимо выяснить, насколько сильно наличие или отсутствие какого-либо свойства или события влияет на частоту возникновения заболевания (осложнения, того или иного исхода и т. д.). Для решения подобных проблем в клинической эпидемиологии разработан достаточно четкий алгоритм, позволяющий сформулировать задачу в терминах статистической гипотезы и затем проверить ее с помощью статистического метода. Поскольку изложение теоретической основы метода можно найти во многих книгах, учебниках и на тематических интернет-сайтах, посвященных статистике и клинической эпидемиологии, мы не будем останавливаться на его технических деталях, а разберем на конкретном примере его смысл и последовательность шагов при решении конкретной задачи подобного рода.

Допустим, что в процессе наблюдения за беременными женщинами, которым проводилось кесарево сечение (КС) под спинномозговой анестезией (СА), было отмечено, что артериальная гипотония, как осложнение СА, в ходе операции чаще регистрировалась у тех пациенток, которые имели исходную склонность к гипотонии. Эта идея посетила наблюдательного доктора, который тщательно проводил предоперационный осмотр каждой пациентки и просто заметил, что при исходной склонности к гипотонии операция реже проходила без проблем со стороны центральной гемодинамики, чем у пациенток с нормальными исходными цифрами давления. Возможно, это просто ошибка, порожденная излишней бдительностью специалиста, но поскольку это осложнение слишком неприятно и может иметь тяжелые последствия, не стоит пренебрегать любыми

возможностями предсказать и предотвратить его. Нам необходимо проверить, является ли исходная склонность к гипотонии у беременной женщины фактором риска развития артериальной гипотонии как осложнения СА при КС, и если да, то насколько серьезно (клинически значимо) она увеличивает этот риск у отдельно взятой пациентки.

Итак, наша гипотеза заключается в следующем: наличие у беременной исходной склонности к гипотонии клинически значимо увеличивает вероятность развития артериальной гипотонии во время операции КС под СА. Гипотеза интересная, но ее формулировка не дает нам никакого руководства к действию. Какие аргументы за или против увеличения вероятности мы можем предоставить, если само это понятие никак нами не определено? И что мы понимаем под термином «клинически значимо»? Для этого ответим на вопрос: почему специалист заподозрил, что исходная склонность к гипотонии увеличивает риск развития указанного осложнения? Потому что ему показалось, что среди женщин с исходной склонностью к гипотонии это осложнение регистрировалась чаще, чем среди тех, кто не имел этой особенности в анамнезе. Иными словами наша гипотеза может быть переформулирована следующим образом: у беременных с исходной склонностью к гипотонии артериальная гипотония при СА развивается чаще, чем у беременных с нормальными исходными цифрами давления. Для проверки (доказательства или опровержения) такой гипотезы все еще нельзя применить статистические методы, поскольку этими методами можно проверить только статистические гипотезы, т. е. такие, которые сформулированы в терминах статистических параметров, но все же

она открывает путь к решению нашей проблемы. Чтобы понять, чаще или реже какое-то событие происходит при наличии известного признака, чем при его отсутствии, нужно разбить имеющуюся выборку на две группы, одна из которых содержит все исследуемые объекты с данным признаком, а другая – все объекты, у которых этот признак отсутствует. Затем необходимо каким-то образом сравнить эти две группы в отношении исследуемого события. Таким образом, возвращаясь к нашей задаче, нам необходимо разбить выборку беременных женщин на две группы по признаку наличия или отсутствия исходной склонности к гипотонии и подсчитать, сколько раз артериальная гипотония во время операции регистрировалась в каждой группе. (Отметим в скобках, что в данном случае мы имеем дело с ретроспективным наблюдением, к результатам которого, особенно ошеломляющим, необходимо относиться очень осторожно и взвешенно, поскольку в таких исследованиях отсутствует фактор рандомизации, равновероятно распределяющий по группам неизвестные или неучтенные побочные влияния на исследуемый признак. Этот вопрос слишком емкий и требует отдельного освещения, поэтому здесь мы не будем останавливаться на нем подробно. Стоит лишь сказать в защиту нашего примера, что мы не можем по своему усмотрению управлять исходными особенностями гемодинамики пациентов, поэтому вынуждены воспользоваться таким дизайном. Несмотря на лавину критики, обрушиваемой на наблюдательные исследования, они все же обладают большой ценностью и информативностью, поскольку в основном именно они служат источником новых идей и гипотез, подлежащих проверке. Просто, как любой инструмент, использование этого вида исследований требует знаний и опыта.)

Итак, в результате мы имеем следующую таблицу частот (табл. 1).

Расположение строк и столбцов этой таблицы должно находиться в определенном порядке для того, чтобы показатели, рассчитанные из нее

в дальнейшем, имели смысл и могли быть корректно интерпретированы. В первой строке всегда должны быть записаны результаты, касающиеся группы, в которой имел место исследуемый фактор воздействия. Иными словами, первая строка отводится для группы, на которую было оказано исследуемое воздействие или объекты которой имели исследуемый фактор риска. Эта группа называется «основной группой» или «группой исследования (вмешательства, риска)». Вторая строка содержит результаты так называемой «контрольной группы» (или «группы сравнения»), хотя это не всегда действительно контрольная группа в строгом смысле этого термина. Как правило, это группа, где исследуемый фактор риска отсутствовал. Иначе говоря, первой должна быть группа, представляющая исследовательский интерес, т. е. та группа, объекты которой подвергались (или наоборот не подвергались, если нас интересует не наличие, а отсутствие воздействия) фактору риска или какому-то другому воздействию, влияние которого нам надо исследовать. Во второй строке должны быть результаты, полученные в той группе, с которой будет происходить сравнение эффекта. Аналогично в первом столбце записывается, сколько раз интересующее нас событие или явление было зарегистрировано в первой и второй группах, во втором столбце – сколько раз оно отсутствовало, т. е. результаты, касающиеся интересующего нас события должны находиться в первом столбце.

Такая таблица называется иногда четырехпольной таблицей сопряженности признаков. Она позволяет выявить наличие связи между двумя признаками объекта, а также между фактором риска и исследуемым событием, которое по гипотезе должно быть связано с ним. Признаки, которые исследуются с помощью четырехпольной частотной таблицы, называются дихотомическими, т. е. принимающими только два значения. Это очень важно отметить, поскольку метод нацелен именно на изучение связи между такими признаками. Как правило, это те случаи, когда исследуют наличие или отсутствие какого-то фактора

Таблица 1. Четырехпольная частотная таблица, сформированная для статистического анализа зависимости частоты артериальной гипотонии во время операции КС под СА от исходной склонности к гипотонии

	Арт. гипотония есть	Арт. гипотонии нет	Всего в строке
Исходная склонность к гипотонии есть	46	121	167
Исходная склонность к гипотонии отсутствует	324	1787	2111
Всего в столбце	370	1908	2278

или события, но, вообще говоря, любой показатель, даже имеющий непрерывную числовую шкалу, можно преобразовать в дихотомический признак, определив точку разбиения его шкалы на две части, например: $ИМТ \geq 26$ и $ИМТ < 26$ (индекс массы тела), равносильно дихотомическому показателю, принимающему значения: «ожирение есть» и «ожирения нет». Этот нехитрый прием позволяет применять данный метод достаточно широко, не заботясь о параметрических характеристиках выборок.

Для указания расчетных формул, которые нам понадобятся, перепишем четырехпольную таблицу в общем виде (табл. 2).

Интуитивно понятно: первое, что мы можем получить из табл. 1, это частота артериальной гипотонии в группе, где отмечалась исходная склонность к гипотонии, и в группе, где этот исходный признак не был зарегистрирован.

Частота события в основной группе, ЧОГ (Experimental Event Rate, EER) – это частота события в группе, подвергавшейся воздействию исследуемого фактора (если исследуемым фактором является отсутствие риска, значит – не подвергавшейся этому воздействию). Она рассчитывается как отношение количества объектов в основной группе, у которых было зарегистрировано событие, к общему числу объектов в этой группе. В нашем случае параметр ЧОГ равен количеству беременных с исходной склонностью к гипотонии, у которых была зафиксирована артериальная гипотония во время СА, деленному на общее число беременных с исходной склонностью к гипотонии. Итак, делим первое число первой строки на последнее число этой же строки и затем умножаем на 100, чтобы получить относительную частоту в процентах. Результат равен 27,5%.

Частота события в контрольной группе, ЧКГ (Control Event Rate, CER) – это частота события в группе контроля или той группе, с которой проводится сравнение. Она также рассчитывается делением количества объектов с зарегистрированным событием в контрольной группе на общее число членов этой группы. В нашем

примере параметр ЧКГ равен количеству беременных без исходной склонности к гипотонии, но с артериальной гипотонией во время СА, деленному на общее число беременных, у которых не отмечалась исходная склонность к гипотонии. Получаем 15,3%.

Мы получили так называемые точечные оценки относительных частот артериальной гипотонии в группе риска (основной группе) и контрольной группе. Казалось бы, осталось их сравнить, и анализ закончен. Но не все так просто! Поскольку мы имеем дело не с точными значениями этих частот, а с так называемыми точечными оценками, которые подвержены статистической ошибке, то вполне вероятно, что в следующей выборке мы получим другие значения. Это происходит потому, что мы рассчитываем эти частоты на основе не всей популяции, а лишь ее репрезентативной части, приблизительно отражающей свойства исследуемой популяции. Обратим внимание на слово «приблизительно»: именно этот факт и является источником статистической ошибки точечных оценок, но в то же время именно он позволяет нам применять статистические методы. В первую очередь необходимо рассчитать **стандартную ошибку доли**, т. е. статистическую ошибку каждой из частот, которая дает представление о точности нашей оценки. И так как полученные частоты могут измениться при расчете на другой выборке, нам, конечно, хотелось бы узнать, насколько существенны могут быть эти изменения и, вообще, какие минимальные интервалы значений «гарантированно» покрывают реальные точные значения искомым частот. Термин «гарантированно» очень неудачен и несовместим с самой природой статистики. Он специально использован здесь, чтобы обратить внимание на этот важнейший факт. Статистика никогда ничего не может гарантировать безоговорочно и точно, поскольку является необычным разделом точной науки математики, в основе которого лежит неточность, имеющая, однако, свое измерение. Эта измеренная неточность (неопределенность) носит название вероятности. Итак, сформулируем вопрос корректно: каков минимальный интервал,

Таблица 2. Общий вид четырехпольной таблицы частот

	Арт. гипотония есть	Арт. гипотонии нет	Всего в строке
Исходная склонность к гипотонии есть	a	b	a+b
Исходная склонность к гипотонии отсутствует	c	d	c+d
Всего в столбце	a+c	b+d	a+b+c+d = N

содержащий реальное значение искомой частоты с вероятностью, скажем, 95%? Если учесть, что такой интервал называется в статистике **95% доверительным интервалом (95% ДИ)**, то задача превращается в нахождение 95% доверительных интервалов для относительных частот артериальной гипотонии в группе с исходной склонностью к гипотонии и в группе без таковой. Для этого существуют хорошо известные стандартные статистические методы. С практической точки зрения 95% доверительный интервал означает, что 95% всех потенциальных выборок дадут нам значения частот, попадающие в полученный интервал, и лишь в 5% случаев значения частот выйдут за найденные пределы, и эти 5% мы будем считать артефактами, не отражающими реалии исследуемой популяции. Смысл доверительного интервала на самом деле достаточно глубок и достоин отдельной статьи, но и такого приблизительного понимания достаточно, чтобы адекватно оценить результаты расчетов. Понимание это крайне необходимо, так как доверительные интервалы сопровождают очень многие статистики, в том числе и в рассматриваемой методике.

Точечные и интервальные оценки частот в двух группах дают первое впечатление об эффекте воздействия фактора риска на исследуемое событие, но впечатление – это очень субъективное понятие. Действительно ли одна частота отличается от другой или, может быть, их разница находится в пределах статистической ошибки, т. е. обусловлена случайностью и на самом деле эти частоты равны? Если они все-таки различны, т. е. влияние фактора риска на частоту возникновения события имеет место, то сколь значительно это влияние? Говоря математическим языком, насколько присутствие фактора риска увеличивает частоту появления события? Поскольку речь идет о расчетах, производимых на выборке, то хотелось бы получить как точечную (выборочное значение частоты \pm ее стандартная ошибка), так и интервальную (доверительный интервал значений с заданной вероятностью, например, 95% или 99%) оценки. Обратимся сначала к первому вопросу. Как понять и доказать, что одна частота действительно (в статистике используется термин «значимо» или «достоверно») отличается от другой. Мы подходим вплотную к самому важному моменту трансформации нашей гипотезы из терминов предметной области в термины статистики, что позволит нам применить стандартные методы проверки статистических гипотез и численно выразить степень нашей уверенности в том, что мы утверждаем. Итак,

если частоты в двух группах (А и В) не отличаются друг от друга, то, значит, они равны: Частота (А) = Частота (В) или то же самое Частота (А)/Частота (В) = 1. Значит, если фактор риска не оказывает влияния на частоту возникновения события, то частоты в обеих группах должны быть приблизительно одинаковы, а их частное не должно отличаться от 1. Но крайне мало вероятно, что мы получим на выборках такие частоты, которые будут точно равны друг другу, а их частное – единице, даже если в действительности никакого влияния нет. Заметим, что поскольку, как мы выяснили, выборочные частоты имеют стандартные ошибки, то и их частное будет иметь ошибку, обусловленную случайностью, а значит, для этого частного можно найти и доверительный интервал, в пределах которого с заранее заданной вероятностью будет находиться реальное значение этого частного. Рассмотренное отношение Частота (А)/Частота (В) (где А – основная группа, а В – контрольная группа) получило название **относительный риск, ОР (Relative Risk, RR)**. Мы хотим узнать, действительно ли относительный риск отличается от 1 или это отличие чисто случайное, обусловленное конкретной выборкой. Обычное слово «действительно» заменяется в статистике термином «значимо» или «достоверно». Итак, мы хотим проверить, значимо ли отличается выборочное значение относительного риска от 1. Поскольку относительный риск величина случайная, имеющая нормальный закон распределения, то наше предположение о том, что ее выборочное значение не отличается значимо от 1 (нуль-гипотеза), является уже статистической гипотезой, к которой применимы статистические методы проверки. Таким образом, мы выразили нашу гипотезу в терминах статистических параметров и имеем в нашем распоряжении весь арсенал статистических методов для ее доказательства или опровержения, но с заранее заданной степенью уверенности, которая сколь угодно близко может подойти к 100%, но никогда не может превратиться в 100% уверенность. Аргументы за или против этой гипотезы (нуль-гипотезы об отсутствии значимого различия) достаточно очевидны. Как для любого параметра, рассчитанного по выборке, для данного значения относительного риска можно найти доверительный интервал с заданной вероятностью (например, 99%). Смысл этого интервала мы обсуждали выше, но коротко еще раз отметим: 99% ДИ означает, что реальное значение искомого относительного риска не выходит за пределы полученного интервала и вероятность, что мы,

принимая эту гипотезу, допускаем ошибку, составляет лишь 1%. Таким образом, если в результате расчетов мы получаем 99% доверительный интервал, не содержащий 1, то можем говорить о том, что частоты в группах статистически значимо различаются с уровнем значимости 1%, т. е. фактор риска действительно оказывает влияние на частоту появления события, и мы в этом уверены на 99%.

Вернемся к нашему примеру о влиянии исходной склонности к гипотонии на частоту развития артериальной гипотонии как следствия СА. Находим 95% ДИ параметров ЧОГ и ЧКГ, т. е. относительной частоты артериальной гипотонии в группе риска и группе сравнения. Граничные значения этих интервалов представлены ниже – 95% ДИ ЧОГ (относительная частота артериальной гипотонии в группе риска): (20,9%; 34,1%); 95% ДИ ЧКГ (относительная частота артериальной гипотонии в группе сравнения): (13,8%; 16,8%).

Относительный риск по этим данным составил 1,79 с 95% доверительным интервалом от 1,38 до 2,34. Поскольку 1 не входит в этот интервал, мы можем утверждать, что значение относительного риска **СТАТИСТИЧЕСКИ** значимо отличается от 1. Иными словами, тот факт, что полученное значение относительного риска не равно 1, является реальным свойством исследуемой популяции, а не случайной флуктуацией, обусловленной нашей выборкой. Поскольку относительный риск в популяции отличается от 1, то мы можем утверждать (с вероятностью ошибки менее 5%), что исходная склонность к гипотонии **статистически значимо** (или **статистически достоверно**), т. е., говоря обычным языком, действительно увеличивает частоту развития артериальной гипотонии как осложнения СА во время КС. Но насколько существенно это влияние с клинической точки зрения? Насколько значимый вклад вносит этот фактор в увеличение частоты артериальной гипотонии? В данном случае слово значимый имеет уже не статистический, а клинический смысл. Несмотря на то, что статистическая достоверность влияния доказана при заданном уровне значимости, с клинической (научной) точки зрения сила (величина) этого влияния может быть малоинтересна или несущественна. Например, предположим, что статистически значимый рост частоты артериальной гипотонии под влиянием фактора риска составил 0,5%. С клинической (научной) точки зрения это едва ли имеет большее значение, хотя статистически наличие влияния доказано на заданном уровне значимости,

равном 5%. Итак, как узнать насколько существенный вклад вносит фактор риска в увеличение частоты события? На этот вопрос может ответить полученное значение абсолютного увеличения (уменьшения) частоты события. Этот показатель называется еще **атрибутивным риском (Absolute Risk Reduction, ARR)**. Если в группе риска частота события составляет, например, 60%, а в контрольной группе – 35%, то атрибутивный риск в этом случае равен 25%, т. е. фактор риска увеличивает вероятность возникновения события на 25%. Много это или мало – решать исследователю, поскольку только ему известны все особенности того процесса, который он изучает. Возвращаясь к нашему примеру, атрибутивный риск исходной склонности к гипотонии в отношении развития артериальной гипотонии как осложнения СА при КС составил $12,2\% \pm 3,6\%$ с 95% ДИ от 5,2% до 19,3%. Иными словами, для каждой беременной с исходной склонностью к гипотонии вероятность такого осложнения, как артериальная гипотония во время КС под СА, возрастает в среднем на 12,20%, минимум – на 5,2%, максимум – на 19,3%. Для практикующего врача увеличение или снижение риска, выраженное в терминах вероятности, хотя и дает некоторое представление о процессе, но все же является достаточно теоретическим. Чтобы перевести полученные результаты в практическую плоскость, рассчитывается показатель NNT (Number Need to Treat), или **число больных, которых необходимо лечить** (или подвергнуть фактору риска), **ЧБНЛ**, чтобы получить один желаемый (или исследуемый) исход. ЧБНЛ рассчитывается из атрибутивного риска и в нашем случае составляет 8,2 с 95% ДИ от 5,2 до 19,4. Это означает, что в среднем из каждых восьми беременных, имеющих в анамнезе исходную склонность к гипотонии, у одной следует ожидать развития артериальной гипотонии во время операции КС под СА. Такая информация дает картину, более приближенную к практическому принятию решений.

В табл. 3 сведены воедино все те расчетные величины, которые мы здесь разобрали.

Но это далеко не вся информация об исследуемом процессе, которую можно извлечь из четырехпольной таблицы сопряженности признаков. На самом деле возможности, которые эта маленькая таблица предоставляет внимательному и непредвзятому исследователю для всестороннего и глубокого изучения интересующей проблемы, достойны удивления. На сайте «CRITICAL» компании «ИнтелТек» (www.critical.ru) в разделе

Таблица 3. Формулы для расчета различных статистических параметров из полей четырехпольной таблицы частот

Наименование статистики	Формула	Пример
Частота в основной группе, ЧОГ	$a/(a+b) * 100\%$	$46/167 * 100\% = 27,5\%$
Частота в контрольной группе, ЧКГ	$c/(c+d) * 100\%$	$324/2111 * 100\% = 15,3\%$
Стандартная ошибка ЧОГ, с.о.(ЧОГ)	$(\text{ЧОГ} * (100 - \text{ЧОГ}) / (a+b))^{1/2}$	$(27,5 * (100 - 27,5) / 167)^{1/2} = 3,5\%$
Стандартная ошибка ЧКГ, с.о.(ЧКГ)	$(\text{ЧКГ} * (100 - \text{ЧОГ}) / (c+d))^{1/2}$	$(15,3 * (100 - 15,3) / 2111)^{1/2} = 0,8\%$
Пределы 95% ДИ для ЧОГ	$\text{ЧОГ} \pm 1,96 * \text{с.о. (ЧОГ)}$	$27,5 \pm 1,96 * 3,5 = (20,7; 34,3)$
Пределы 95% ДИ для ЧКГ	$\text{ЧКГ} \pm 1,96 * \text{с.о. (ЧКГ)}$	$15,3 \pm 1,96 * 0,8 = (13,8; 16,8)$
Относительный риск, ОР*	$\text{ЧОГ} / \text{ЧКГ}$	$27,5 / 15,3 = 1,79$
Атрибутивный риск, АР*	$\text{ЧОГ} - \text{ЧКГ}^{\&}$	$27,5 - 15,3 = 12,2\%$
Стандартная ошибка АР	$(\text{ЧОГ} * (100 - \text{ЧОГ}) / (a+b) + \text{ЧКГ} * (100 - \text{ЧОГ}) / (c+d))^{1/2}$	$(3,5 * 3,5 + 0,8 * 0,8)^{1/2} = 3,6$
Пределы 95% ДИ для АР	$\text{АР} \pm 1,96 * \text{с.о. (АР)}$	$12,2 \pm 1,96 * 3,6 = (5,2; 19,3)$
ЧБНЛ	$100 / \text{АР}$	$100 / 12,2 = 8,2$
Нижний предел 95% ДИ для ЧБНЛ	$100 / \text{ВП (АР)}^{\$}$	$100 / 19,3 = 5,2$
Верхний предел 95% ДИ для ЧБНЛ	$100 / \text{НП (АР)}^{\$}$	$100 / 5,2 = 19,4$

Примечания: * – формулы для расчета стандартной ошибки и 95% ДИ ОР и АР достаточно сложны, поэтому не приводятся в этой таблице. Для расчета этих параметров лучше обратиться к стандартным статистическим пакетам или к помощи специалиста.

& – значение атрибутивного риска будет положительным, если частота растет под воздействием исследуемого фактора, и отрицательным, если частота снижается.

\$ – ВП(АР) и НП(АР) означают верхний и нижний пределы 95% ДИ для АР соответственно.

«Форум» находится программное обеспечение, с помощью которого можно провести полный расчет и анализ результатов четырехпольной таблицы с последующей распечаткой текстовой интерпретации всех рассчитанных величин. Этот программный инструмент много раз проверен на практике, в том числе и при проведении большого мульти-центрового исследования осложнений спинномозговой анестезии при операции кесарева сечения, и может быть использован для анализа выборок любого объема. В реальном исследовании провести статистическую обработку данных означает не только рассчитать какие-то интегральные величины по выборке, но и, самое главное, правильно их проинтерпретировать, потому что именно на интерпретации полученных результатов строятся выводы и гипотезы. Правильная и исчерпывающая интерпретация результатов статистической обработки данных – это ключ к получению по-настоящему новой и объективной информации, поэтому так важно уделять этой части исследования не меньше внимания, чем разработке дизайна и протокола. Кроме того, для успешного проведения этой финальной части исследования требуются знания статистических методик, но не только и не столько самих расчетных формул, сколько понимания сути и смысла, который содержится в их результатах. Об этом в основном

мне и хотелось рассказать в данной и последующих публикациях по статистике.

Коротко:

1. Чтобы иметь возможность применить методы математической статистики, необходимо перевести исследуемую проблему в термины, параметры и гипотезы математической статистики.
2. Четырехпольная таблица сопряженности признаков позволяет изучать взаимосвязь двух дихотомических показателей (свойств, признаков). Дихотомической называется любая переменная, принимающая только 2 значения (да/нет, есть/нет, норма/патология и т. д.).
3. Чтобы сформировать четырехпольную таблицу частот в обсервационном исследовании, надо стратифицировать полученную выборку по наличию и отсутствию изучаемого фактора, сформировав, таким образом, основную и контрольную группы и подсчитать частоту изучаемого события в каждой из этих групп.
4. Расчет интегральных статических величин можно проводить с помощью формул, указанных в табл. 3, или с использованием специально созданных статистических программ.
5. Все величины, рассчитанные статистическими методами, имеют стандартные ошибки, которые также должны быть рассчитаны и указаны

- в обязательном порядке. Вместо них рядом с рассчитанной величиной можно указать доверительный интервал с заданным уровнем значимости.
6. Необходимо различать статистическую и клиническую значимость полученной зависимости. Это совершенно разные и независимые понятия. Увеличение статистической значимости никак не влияет на клиническую значимость. Статистическая значимость определяется стандартными статистическими методами и не зависит от воли исследователя. Клиническая значимость устанавливается только специалистом в исследуемой области и не подлежит алгоритмизации, это исключительно эвристическая процедура.
 7. Содержательная (текстовая) интерпретация полученных чисел в терминах исследуемой проблемы (а не простое перечисление средних значений, стандартных отклонений, относительных рисков и других статистических параметров) является финальной, наиболее сложной, но и наиболее важной частью получения значимых и понятных выводов.
В данной статье мы рассмотрели несложный, но достаточно информативный метод, позволяющий анализировать качественные признаки дихотомической природы.