

федеральное государственное бюджетное образовательное
учреждение высшего образования
«Приволжский исследовательский медицинский университет»
Министерства Здравоохранения Российской Федерации

Жукова О.В.

**Применение методов атрибутивной статистики в
фармацевтических и медико-биологических исследованиях
(таблицы сопряженности)**

Учебное пособие

Нижний Новгород, 2020

УДК

Автор:

Жукова О.В., к.фарм.н., доцент, заведующий кафедрой фармацевтической химии и фармакогнозии ФГБОУ ВО ПИМУ Минздрава России

Рецензенты:

Петров Сергей Борисович - заведующий кафедрой гигиены ФГБОУ ВО «Кировский государственный медицинский университет» Минздрава России, к.м.н., доцент

Гордцев Александр Сергеевич - заведующий кафедрой общей химии ФГБОУ ВО «Приволжский исследовательский медицинский университет» Минздрава России, д.х.н., профессор

Применение методов атрибутивной статистики в фармацевтических и медико-биологических исследованиях (таблицы сопряженности) : Учебное пособие / О.В. Жукова. – Нижний Новгород: ПИМУ, 2020. – 74 с.

В учебном пособии представлены теоретические и методологические основы использования математико-статистических методов атрибутивной статистики при проведении исследования.

Учебное пособие предназначено для студентов, обучающихся по основной профессиональной образовательной программе по специальности 33.05.01 «Фармация».

ОГЛАВЛЕНИЕ

1	Принципы и роль математико-статистических методов и компьютерных технологий в современном научном знании.....	4
2	Математико-статистический анализ и компьютерные технологии в здравоохранении (на примере бронхолегочных заболеваний).....	10
3	Таблица сопряженности – элемент атрибутивной статистики.....	16
4	Статистический анализ таблиц сопряженности.....	19
4.1.	Анализ с использованием критерия χ^2 Пирсона.....	19
4.2.	Особенности расчета критерия χ^2 -квадрата с поправкой Йейтса..	23
5.	Примеры определения критерия χ^2 -квадрата в атрибутивной статистике при анализе таблиц сопряженности.....	24
5.1.	Определение статистической зависимости между курением и частыми обострениями хронической обструктивной болезни легких (ХОБЛ).....	24
5.2.	Определение статистической зависимости между включением антагониста лейкотриеновых рецепторов в схемы противоастматической терапии и частотой наступления положительных клинических исходов	27
6	Точный тест Фишера.....	31
7	Критерий ϕ и критерий Крамера (V) для оценки силы связи между воздействием анализируемого фактора и наступлением события..	33
8	Определение математико-статистической связи (количественные характеристики) между курением и развитием частых обострений у пациентов с ХОБЛ.....	38
9	Определение математико-статистической связи (количественные характеристики) между конкретными возбудителями и развитием воспалительных заболеваний респираторного тракта (внебольничная пневмония, острый бронхит).....	44
10	Определение клинической эффективности лекарственного препарата на основании математико-статистического анализа....	60
11	Тестовые задания.....	67
12	Ситуационные задачи.....	71
13	Эталоны ответов для тестовых заданий и ситуационных задач.....	72
14	Рекомендуемая литература.....	73

1. Принципы и роль математико-статистических методов и компьютерных технологий в современном научном знании

Для оценки результатов как научной, так и практической деятельности очень широко в последнее десятилетие стали применяться методы количественного и качественного анализа. Данный процесс сопровождается интенсивным внедрением количественных методов, основанных на использовании математического аппарата.

В настоящее время имеется множество компьютерных технологий (программы Excel, MathCad, MathLab, Statistica, SPSS, Statgraphics, Stadia и др.), позволяющие обрабатывать массивы данных и представлять информацию как в числовом значении, так и графически. Но использование программных средств при отсутствии методологии использования методов математико-статистического анализа в исследованиях может привести к некорректному, формальному использованию математического аппарата, не позволяя сформулировать правильные выводы и снижая качество обоснованности и достоверности полученных результатов.

Основными задачами использования методов математико - статистического анализа в любой области исследований являются:

- проверка достоверности различий между полученными результатами, сравнение одновременно нескольких групп результатов измерений (выборок), объединенных в единый статистический комплекс (дисперсионный анализ);
- выявление силы связи между отдельными явлениями, объектами (корреляционный анализ);
- изучение влияния одних признаков на другие (регрессионный анализ);
- распределение данных по факторам, позволяющий выявлять весомости каждого фактора (факторный анализ);

- прогнозирование и моделирование изучаемых процессов на основе компьютерных технологий и математико-статистических методов, например, диаграмма «Дерево решений», модель Маркова, нейронные сети и др.

В последнее время в нашей стране, как и во всем мире, возрастает количество математико-статистических исследований в различных областях научного знания, отраслях экономики и народного хозяйства. Точные и естественные науки давно используют методы математического анализа при проектировании материалов, продуктов, прогнозировании их эффективности. Также математические методы являются научным направлением экономики, направленным на исследование с помощью математических моделей экономических процессов и систем. Также математические методы применяются специалистами в сфере управления. Использование математических методов позволяет снизить временные затраты на проведение экономического анализа, при этом данные методы способствуют более полному учету факторов, оказывающих влияние на результат деятельности организации. Другим важным преимуществом математических методов выступает значительное снижение риска ошибочного вывода и точность расчетов и вычислений¹²³.

Исследуется и анализируется глобальное развитие экономики страны, регионов, отдельных отраслей и предприятий. Данное направление прочно вошло в систему подготовки специалистов в области финансов и экономики. Широко проводится анализ данных в социальных науках, педагогике и психологии. Математико-статистический и компьютерный анализы проводятся в спорте.

¹ Алесинская Т.В., Сербин В.Д., Катаев А.В. «Экономико-математические методы и модели. Линейное программирование», учебно-методическое пособие. Таганрог, 2013. 79 с.

² Балдин, К.В. Математические методы и модели в экономике: Учебник / К.В. Балдин, В.Н. Башлыков, А.В. Рукосуев. - М.: Флин- та, МПСИ, 2012. - 328с.

³ Белолипецкий, А.А. Экономико-математические методы: Учеб- ник для студ. высш. учеб. заведений / А.А. Белолипецкий. - М.: ИЦ Академия, 2014. - 368 с.

По мнению Орлова А.И. , одним из отличительных признаков новой парадигмы математической и прикладной статистики, анализа данных и других статистических методов, математических методов экономики является широкое применение компьютерно-статистических методов⁴. Согласно новой парадигме, о которой говорит Орлов А.И., информационные технологии – эффективные инструменты получения выводов (имеются в виду датчики псевдослучайных чисел, размножение выборок, в т.ч. бутстреп, автоматизированный системно - когнитивный анализ и др.).

Методы статистических испытаний или методы Монте-Карло включают две составляющие: база и надстройка. Базой являются датчики псевдослучайных чисел. Результатом работы таких датчиков являются последовательности чисел, которые обладают некоторыми свойствами последовательностей случайных величин (в смысле теории вероятностей). Надстройкой являются различные алгоритмы, использующие последовательности псевдослучайных чисел.

При имитационном моделировании зачастую используется метод статистических испытаний (Монте-Карло) .

Определенное место компьютерная статистика заняла в контроллинге (современный подход к управлению организацией).

В математико-компьютерном обеспечении принятия решений в контроллинге используются методы теории нечеткости, относящиеся к системной нечеткой интервальной математике.

Поиск оптимального решения, оптимальной стратегии, использование которых приводит к наилучшим из возможных результатов в заданных условиях, является целью моделирования.

Многие работы направлены на разработку рекомендаций по формированию последовательности в процессе принятия решений. Так, коллективом авторов (Надеждин Е.Н., Смирнова Е.Е., Варзаков В.С., 2011) обобщены эти

⁴ Орлов А.И. Компьютерно-статистические методы: состояние и перспективы// Научный журнал КубГАУ, №103(09), 2014. URL: <http://ej.kubagro.ru/2014/09/pdf/12.pdf> (дата обращения 05.08.2017г.)

рекомендации и предложен алгоритм типового процесса принятия решений:

- 1) предварительное формулирование проблемы;
- 2) определение целей операции и выбор соответствующих критериев оптимальности;
- 3) выявление и формулирование дисциплинирующих условий;
- 4) составление возможно более полного списка альтернатив и предварительный их анализ с целью отбрасывания явно неэффективных;
- 5) сбор необходимой информации и прогнозирование изменений параметров операции в будущем;
- 6) точное формулирование постановки задачи;
- 7) разработку математической модели операции, позволяющей оценивать эффективность каждой альтернативы;
- 8) выбор метода решения задачи и разработка алгоритма решения;
- 9) оценку альтернатив и определение наиболее эффективных;
- 10) принятие решения ответственным руководителем;
- 11) выполнение решения и оценка результатов⁵.

Переработка информации происходит в виде последовательных трёх фаз, для каждой из которых характерен свой уровень определенности решения задачи: 1) структуризация; 2) характеристика; 3) оптимизация (табл.1).

Таблица 1.

Фазы переработки информации в процессе поиска оптимального решения

Фаза переработки информации	Характеристика фазы переработки информации
структуризация	выделение основных элементов решаемой задачи и установление зависимости между ними. Результат - упорядоченная система, позволяющая определить необходимую информацию и распределить её по элементам структуры (схемы, таблицы)
характеризация	определение системы характеристик, параметров и

⁵ Надеждин Е.Н., Смирнова Е.Е., Варзаков В.С. Математические методы и модели в экономике: учебное пособие для студентов экономических специальностей. - Тула: Автономная некоммерческая организация ВПО «Институт экономики и управления», 2011. - 249 с.

	показателей, количественно описывающих поставленную задачу. Результат более полное и точное описание задачи в сравнении со структуризацией и подготовка исходных данные для оптимизации
оптимизация	поиск наилучшего решения задачи. Результат - имеющаяся информация преобразуется в конечную форму, описывающую решение и показывающую зависимость характеристики решения от исходных данных. Проведение оптимизации приводит к полной определенности решения задачи

Основой операционного исследования является математическая модель. Построение и расчет математической модели позволяют проанализировать ситуацию и выбрать оптимальные решения по управлению ею или обосновать предложенные решения. Применение модели необходимо в тех случаях, когда проблема сложна, зависит от большого числа факторов, по-разному влияющих на её решение. В этом случае непродуманное и научно не обоснованное решение может привести к серьезным последствиям. Примеров этому имеется немало, в частности в экономике. Использование математических методов и моделей позволяет осуществить предварительный выбор оптимальных или близких к ним вариантов решений по определенным критериям. Они научно обоснованы, и лицо, принимающее решение, может руководствоваться ими при выборе окончательного решения. В настоящее время математические модели применяются в задачах анализа, прогнозирования и выбора оптимальных решений в различных областях экономики. Это планирование и оперативное управление производством, управление персоналом, управление запасами, распределение материальных ресурсов, планировка и размещение активов, руководство инновационным проектом, формирование портфеля заказов и т. п.

Существует множество математических моделей. Классификация их подробно представлена в работе Надеждина Е.Н., Смирновой Е.Е., Варзакова В.С. По числу критериев эффективности математические модели делятся на однокритериальные и многокритериальные. Многокритериальные математические модели содержат два и более критерия. По учёту неизвестных факторов математические модели делятся на детерминированные, стохастические и модели с элементами неопределённости. В стохастических моделях неизвестные факторы – это случайные величины, для которых известны функции распределения и различные статистические характеристики (математическое ожидание, дисперсия, среднеквадратическое отклонение и т. п.). К стохастическим относятся модели стохастического программирования, в которых либо в целевую функцию, либо в ограничения входят случайные величины (рис.1).

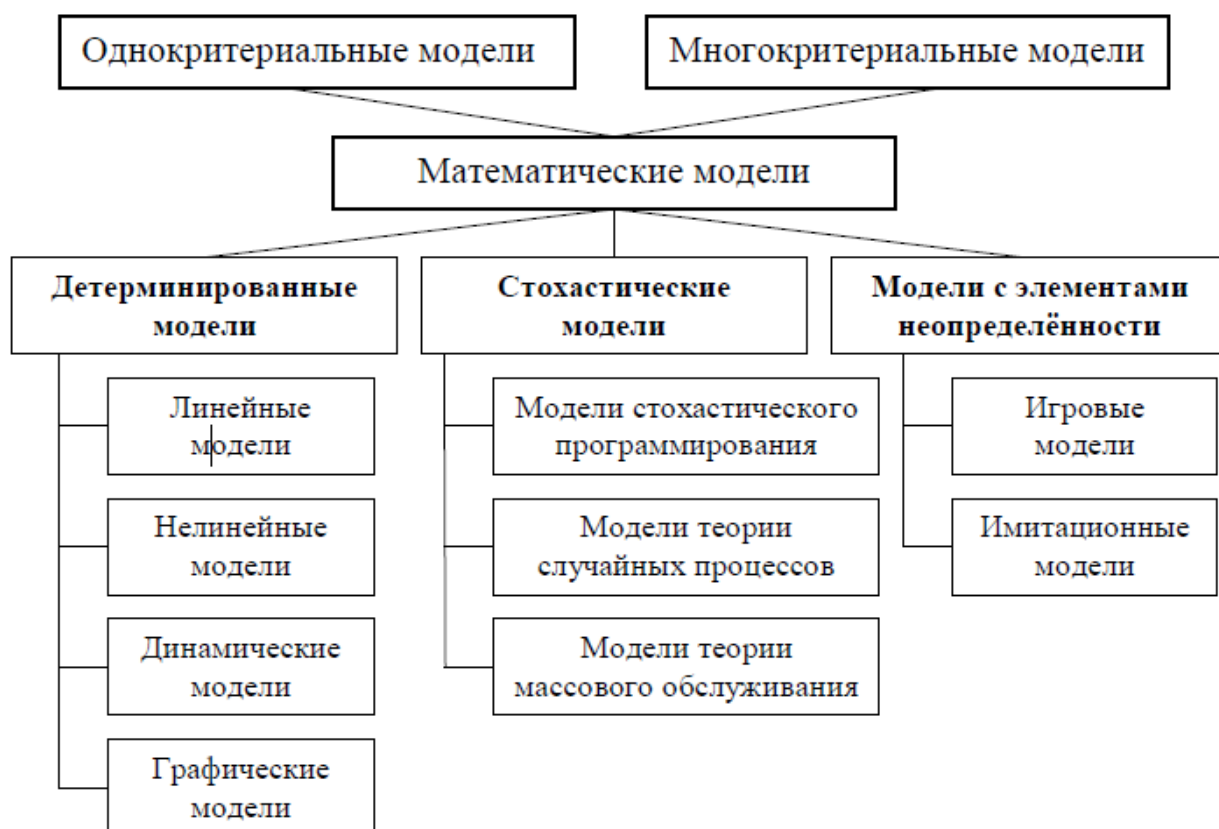


Рисунок 1. Классификация математических моделей⁶.

⁶Надеждин Е.Н., Смирнова Е.Е., Варзаков В.С. Математические методы и модели в экономике: учебное пособие для студентов экономических специальностей. - Тула: Автономная некоммерческая организация ВПО «Институт экономики и управления», 2011. - 249 с.

2. Математико-статистический анализ и компьютерные технологии в здравоохранении (на примере бронхолегочных заболеваний)

В медицине и здравоохранении часто используются различные статистические концепции при принятии решений по таким вопросам как оценка состояния здоровья, его прогноз, выбор стратегии и тактики профилактики и лечения, оценка отдаленных результатов и выживаемости. Знания принципов статистики абсолютно необходимо для планирования, проведения и анализа научных исследований в медицине. Особенно это актуально для области общественного здоровья и здравоохранения, базирующихся на данных популяционных исследований. Состоятельность таких исследований и их результатов зависит от применения на всех стадиях разумных статистических принципов. Медико-статистическая информация является основой подготовки решений по совершенствованию стратегии развития отрасли здравоохранения. По мнению Какориной Е.П. (2012) медицинская статистика в России характеризуется ведомственной разобщенностью и требует пересмотра всей организации отчета и учета медицинской документации с целью совершенствования работы данной службы. В целях оптимизации работы службы медицинской статистики необходима система непрерывного повышения квалификации врачей – статистиков, увеличение полномочий и функций медицинских статистиков в плане учета и контроля деятельности практикующих врачей, а также увеличение их роли в планировании деятельности МО, в том числе и в выборе стратегий терапии.

Использование методов математической и прикладной статистики, а также компьютерных технологий является в настоящее время закономерным требованием качественного развития системы здравоохранения в плане оптимизации технологий здравоохранения. Современные технические и программные средства информатизации обеспечивают возможность создания электронных баз данных в достаточных объемах для получения достоверных

результатов. Наличие соответствующих пакетов прикладных программ по математико-статистической обработке и анализу данных позволяют применять многомерные математические методы и процедуры специалистами предметных областей, далеких от фундаментальных знаний прикладной математики. Создаваемые на этой основе многомерные математико-статистические модели сложных медицинских систем позволяют перейти от интуитивных представлений врача о медицинских процессах к математически обоснованному количественному анализу и моделированию. Но такая возможность в нашей стране используется недостаточно.

Нами был проведен наукометрический анализ публикаций по использованию моделирования в медицине и в частности в терапии заболеваний (на примере заболеваний пульмонологического профиля). Для поиска были выбраны РИНЦ – база данных, отражающая преимущественно ситуацию в России, и PubMed – англоязычная текстовая [база данных](#) медицинских и биологических публикаций, отражающая ситуацию в мире.

Поисковый запрос «математическое моделирование» по тематике 76.00.00 «Медицина и здравоохранение» в РИНЦ дал результат в 3337 публикаций (дата обращения 05.02.2020г.). Всего на момент обращения в базе РИНЦ имелось 3 714 997 публикации по медицинской тематике. Доля использования «математического моделирования» в них составила менее 0,1% (0,09%). При этом основная масса публикаций затрагивала вопросы моделирования в фундаментальной и прикладной медицине, например, такие как математическое моделирование флуоресценции биологических тканей⁷, моделирование альтерации структурных элементов глаза⁸, моделированию

⁷ Дрёмин В.В. [Аналитический обзор подходов к математическому моделированию флуоресценции биологических тканей](#) // [Фундаментальные и прикладные проблемы техники и технологии](#). 2016. № 6 (320). С. 92-102.

⁸ Баринов Э.Ф., Голубов К.Э., Сухина Л.А., Сулаева О.Н., Канзюба И.А. Контузионная травма органа зрения: от экспериментального к математическому моделированию альтерации структурных элементов глаза // [Офтальмология Восточная Европа](#). 2013. № 3 (18). С. 56-63.

действительных размеров остаточной костной полости при хирургическом лечении околокорневых кист челюстей⁹ и т.д.

Однако в последнее время появляются работы по моделированию, разработке методик, алгоритмов для системы здравоохранения, направленные на автоматизацию принятия решений и совершенствование системы здравоохранения в целом.

Направлением, связанным с оптимизацией схем лечения, лекарственного обеспечения, в медицине и фармации занимался ряд ученых. Ростова Н.Б. разработала организационные подходы, в т.ч. на основе фармакоэкономических расчетов, к оптимизации процесса лекарственного обеспечения на основе принципов рационального использования лекарственных средств при оказании медицинской помощи^{10 11}.

Шакирова Д.Х. занималась математическим моделированием потребности в ЛП при различных нозологиях¹².

Моделированием в фармакоэкономике, в т.ч. и математическим очень широко занимается коллектив кафедры организации лекарственного обеспечения и фармакоэкономики и лаборатории фармакоэкономических исследований [Первого Московского государственного медицинского университета им. И.М. Сеченова](#) (Сеченовский Университет) под руководством Ягудиной Р.И. Авторами разрабатываются и методология проведения фармакоэкономических исследований с применением

⁹ Ефимов Ю.В. Хирургическое лечение околокорневых кист челюстей с математическим моделированием действительных размеров остаточной костной полости. Автореферат диссертации на соискание ученой степени кандидата медицинских наук / Московский государственный медико-стоматологический университет им. А.И. Евдокимова. Москва, 1994.

¹⁰ Ростова Н.Б. [Рациональное использование лекарственных средств как основа совершенствования лекарственного обеспечения населения](#)//[Медицинский альманах](#). 2012. № 1. С. 206-208.

¹¹ Ростова Н.Б. [Результаты изучения организации процесса лекарственного обеспечения и внедрения формулярной системы в медицинских учреждениях](#)//[Проблемы стандартизации в здравоохранении](#). 2011. № 1-2. С. 30-34.

¹² Дун О.А., Шакирова Д.Х., Хабиров Ф.А., Хайбуллин Т.И., Бабичева Н.Н., Аверьянова Л.А., Гранатов Е.В. Математическое моделирование потребности в препаратах, изменяющих течение рассеянного склероза//[Фармация](#). 2012. № 2. С. 34-37.

инструментов моделированиях¹³ и моделирование фармакоэкономической эффективности терапии различных нозологий¹⁴.

Что касается результатов поиска публикаций, освещающих методики, проблемы и результаты математического моделирования в клинко-экономической оценке терапии заболеваний, в PubMed они (исследования) представлены намного шире (табл.2).

Таблица 2.

Результаты наукометрического поиска по использованию моделирования в здравоохранении в PubMed(дата обращения 05.02.2020 г.)

Формулировка поискового запроса	Количество наеденных результатов, соответствующих запросу
mathematical model	1551217
mathematical model in pharmaco-economic	1624
markov model cost-effectiveness	4832
markov model cost-effectiveness COPD	67
markov model cost-effectiveness bronchial asthma	40
markov model cost-effectiveness pneumonia	49 1993 по 2018 гг., из них 39 посвящено клинко-экономической оценке вакцинации

¹³ Ягудина Р.И., Серпик В.Г. [Методологические основы фармакоэкономического моделирования//Фармакоэкономика: теория и практика. 2016. № 1. С. 7-12.](#)

¹⁴ Дедов И.И., Шестакова М.В., Сунцов Ю.И., Ягудина Р.И., Крысанов И.С., Куликов А.Ю., Аринина Е.Е. [Фармакоэкономическое моделирование отдаленных результатов лечения сахарного диабета 2 типа у пациентов, получавших современные аналоги инсулина по сравнению с терапией пероральными сахароснижающими препаратами//Сахарный диабет. 2010. № 1. С. 101-110.](#)

Марковское моделирование используется при клинико-экономической оценке терапии заболеваний (4832 публикация). Марковское моделирование в клинико-экономической оценке ХОБЛ использовалось в 67 исследованиях. Основное количество посвящено оценке базисной терапии ХОБЛ в виде сравнения используемых схем, новых комбинаций.

Британские исследователи (McLean, Barbour Sarah, Wild, Simpson) в области организации здравоохранения в 2014-2015 гг. провели поиск и анализ эпидемиологических моделей для оценки распространенности и бремени ХОБЛ. Поиск включал базы данных Medline, Embase, CAB Abstracts и ВОЗ с 1980 по ноябрь 2013 года. Авторы определили 22 модели, в которых используется множество методов для расчета распространенности и / или бремени ХОБЛ. Модели рассчитывали распространенность и / или смертность или другие аспекты бремени болезней с использованием демографических факторов и факторов риска или тенденций, моделирование типа Маркова и моделирование микросимуляции. Определены шесть моделей, которые были высококачественными: модель Peabody, которая представляла оценки распространенности ХОБЛ¹⁵; Модель ВОЗ DISMOD II, в которой были сделаны оценки бремени с точки зрения инвалидности, скорректированных с годами с ХОБЛ, и лет жизни, утраченных в ХОБЛ; Модель Atsou, которая дала прирост ожидаемой продолжительности жизни отдельных курильщиков, которые бросили курить и связанные с этим расходы¹⁶; две голландские модели ХОБЛ, в которых оценивались расходы на здравоохранение, связанные с ХОБЛ; модель Pichon-Riviere, которая представляла издержки и экономическую эффективность программ по отказу от курения¹⁷.

¹⁵ Peabody JW, Schau B, Lopez-Vidriero M, . COPD: A prevalence estimation model. *Respirology* 2005; 10: 594–602.

¹⁶ Atsou K, Chouaid C, Hejblum G. Simulation-based estimates of effectiveness and cost-effectiveness of smoking cessation in patients with chronic obstructive pulmonary disease. *PLoS One* 2011; 6: e24870–e24870

¹⁷ Pichon-Riviere A, Augustovski F, Bardach A, . Development and validation of a microsimulation economic model to evaluate the disease burden associated with smoking and the cost-effectiveness of tobacco control interventions in Latin America. *Value Health* 2011; 14(5 Suppl.): S51–S9.

Экономическая и клиническая оценка на основе математического моделирования, в т.ч. и Марковского, отказа от курения, как главного фактора риска ХОБЛ, широко освещена в зарубежной литературе. Все исследователи отмечают улучшение функции легких при отказе от курения при ХОБЛ и, как следствие, снижение затрат на терапию.

При анализе БА также проводится моделирование как течения заболевания, так и влияния терапии и ее режимов. В 2016 г. в Чили и Колумбии проводилось исследование, в основе которого лежало моделирование, сопоставления затрат и полезности дозирования ингаляционных глюкокортикостероидов (один раз в день и два раза в день) при БА у детей. Было показано, что по сравнению с дозированием два раза в день стратегия дозирования один раз в день включала более низкие затраты (1529,3 долл. США против средней стоимости 1709,1 долл. США на одного пациента в течение 12 месяцев) и имела наибольший прирост в QALY (0,8284 против 0,8084 QALY в среднем на одного пациента в течение 12 месяцев).

Математическое моделирование в оценке терапии и течения заболевания, в прогнозировании эффективности от использования того или иного ЛП, медицинской технологии, в последние десятилетие активно внедряется в здравоохранение, позволяя оптимизировать систему принятия управленческих решений.

Большинство моделей в здравоохранении основано на статистически полученных в ходе исследования значениях. Анализ статистической информации является основной стадией управленческого процесса в здравоохранении, определяющего состояние здоровья населения и эффективность системы здравоохранения. С его помощью выявляются конкретные проблемы, определяются приоритеты, ставятся необходимые задачи для более полного и рационального использования кадровых, финансовых, материально-технических ресурсов.

3. Таблица сопряженности – элемент атрибутивной статистики

Статистика — отрасль знаний, в которой излагаются общие вопросы сбора, измерения и анализа массовых статистических (количественных или качественных) данных; изучение количественной стороны массовых общественных явлений в числовой форме. Статистическим рядом распределения называют упорядоченное распределение единиц совокупности на группы по изучаемому признаку. В зависимости от признака ряды могут быть вариационными (количественными) и атрибутивными (качественными). Т.е., упрощенно статистику можно разделить на вариационную (количественную) и атрибутивную (качественную).

Методы атрибутивной статистики нечасто используются в медицинских и фармацевтических исследованиях. Хотя проводимые эксперименты и получаемые в ходе них результаты имеют все предпосылки для применения методов атрибутивной статистики. Иногда исследователи непреднамеренно делают ошибки, пытаясь качественные признаки описать количественно.

Важной основой атрибутивной статистики является таблица сопряженности.

Таблица сопряженности (или таблица контингентности, факторная таблица) — важный элемент атрибутивной статистики. Средство представления совместного распределения двух переменных, предназначенное для исследования связи между ними. Таблица сопряженности представляет собой универсальное средство исследования статистической связи.

Большинство результатов медико-биологических, фармацевтических исследований возможно представить в виде таблиц сопряженности, в частности в виде четырехпольной таблицы (табл.3).

Таблица 3.

Четырехпольная таблица сопряженности

Изучаемый фактор	Гипотетическое состояние, к которому приводит действие изучаемого фактора		Всего
	Да	Нет	
Есть	(a) Группа в гипотетическом состоянии <i>с воздействием</i> изучаемого фактора риска	(b) Группа вне гипотетического состояния <i>с воздействием</i> изучаемого фактора риска	(A) Сумма a+b
Нет	(c) Группа в гипотетическом состоянии <i>без воздействия</i> изучаемого фактора риска	(d) Группа вне гипотетического состояния <i>без воздействия</i> изучаемого фактора риска	(B) Сумма c+d
Всего	(C) Сумма a+c	(D) Сумма b+d	(Q) Сумма A+B или C+D

Строки таблицы сопряженности соответствуют значениям одной переменной (воздействие изучаемого фактора), столбцы - значениям другой переменной (событие, к которому привел или не привел изучаемый фактор). На пересечении строки и столбца указывается частота совместного появления соответствующих значений двух признаков (фактор воздействовал и привел к наступлению события (a); фактор воздействовал и не привел к наступлению этого события (b); фактор не воздействовал, но событие наступило (c); фактор не воздействовал, и событие не наступило (d)). Именно так проводится сравнение с контрольной группой или группой сравнения, например, нового метода лечения, нового лекарственного препарата. Зачастую медико-биологические исследования ставят своей целью установить статистическую взаимосвязь между изучаемым методом, признаком, фактором и наступлением исследуемого события (выздоровление, усугубление течения заболевания, улучшение состояния и т.д.).

В таблицах сопряженности могут быть представлены как абсолютные, так и относительные частоты (в долях или процентах). Относительные частоты могут рассчитываться по отношению: к маргинальной частоте по строке; к маргинальной частоте по столбцу; к объему выборки.

Сумма частот по строке называется маргинальной частотой строки; сумма частот по столбцу - маргинальной частотой столбца. Сумма маргинальных частот равна объему выборки n ; их распределение представляет собой одномерное распределение переменной, образующей строки или столбцы таблицы.

4. Статистический анализ таблиц сопряженности

Таблицы сопряженности используются для проверки гипотезы о наличии связи между двумя признаками ([статистическая связь](#), [критерий «хи-квадрат»](#)), а также для измерения тесноты связи.

4.1. Анализ с использованием критерия χ^2 Пирсона

Критерий χ^2 Пирсона – это непараметрический метод, который позволяет оценить значимость различий между фактическим (выявленным в результате исследования) количеством исходов или качественных характеристик выборки, попадающих в каждую категорию, и теоретическим количеством, которое можно ожидать в изучаемых группах при справедливости нулевой гипотезы. Непараметрические критерии применяются к тем исследованиям, когда ничего неизвестно ни о законах распределения интересующей нас величины, ни об их параметрах.

Значение χ^2 - квадрат используется для оценки меры рассогласованности наблюдаемого и ожидаемого результата. Если ожидаемый результат будет сильно отличаться от наблюдаемых значений, значит стоит поставить под сомнение справедливость гипотезы.

Вычислим значение χ^2 -квадрат. Значение представляет собой сумму квадратов разностей наблюдаемой и ожидаемой частоты, деленную на соответствующую ожидаемую частоту:

При определении критерия χ^2 -квадрата необходимо сначала **определить ожидаемое количество наблюдений. Для этого на основании таблицы сопряженности, которая отражает фактические значения (табл.3), строится таблица ожидаемых количеств наблюдений** путем перемножения сумм рядов и столбцов с последующим делением полученного произведения на общее число наблюдений (табл.4).

Таблица 4.

Четырехпольная таблица ожидаемых количеств наблюдений

	Событие наступило	Событие не наступило	Всего
Фактор воздействует	$(a+b)*(a+c)/(a+b+c+d)$	$(a+b)*(b+d)/(a+b+c+d)$	a + b
Фактор отсутствует	$(c+d)*(a+c)/(a+b+c+d)$	$(c+d)*(b+d)/(a+b+c+d)$	c + d
Всего	a + c	b + d	

Далее определяем критерия χ^2 -квadrата по формуле (1):

$$\chi^2 = \sum_{i=1}^l \sum_{j=1}^c \frac{(O_{ij} - \Phi_{ij} - 0.5)^2}{\Phi_{ij}} \quad (1),$$

где i – номер строки (от 1 до l), j – номер столбца (от 1 до c), O_{ij} – фактическое количество наблюдений в ячейке ij , Φ_{ij} – ожидаемое число наблюдений в ячейке ij .

На следующем этапе после получения критерия χ^2 -квadrата необходимо правильно интерпретировать полученный результат. Для этого полученное значение критерия χ^2 необходимо сравнить с критическим, которое определяется на основании таблицы критических точек распределения Пирсона (χ^2 -квadrат) (табл. 5).

Таблица 5.

Таблица критических точек распределения Пирсона (χ^2 -квadrат)

Число степеней свободы (f)	Уровень значимости (вероятность) (p)					
	0,01	0,025	0,05	0,95	0,975	0,99
1	6,63	5,02	3,84	0	0	0
2	9,21	7,38	5,99	0,1	0,05	0,02
3	11,34	9,35	7,81	0,35	0,22	0,11
4	13,28	11,14	9,49	0,71	0,48	0,3
5	15,09	12,83	11,07	1,15	0,83	0,55

6	16,81	14,45	12,59	1,64	1,24	0,87
7	18,48	16,01	14,07	2,17	1,69	1,24
8	20,09	17,53	15,51	2,73	2,18	1,65
9	21,67	19,02	16,92	3,33	2,7	2,09
10	23,21	20,48	18,31	3,94	3,25	2,56
11	24,72	21,92	19,68	4,57	3,82	3,05
12	26,22	23,34	21,03	5,23	4,4	3,57
13	27,69	24,74	22,36	5,89	5,01	4,11
14	29,14	26,12	23,68	6,57	5,63	4,66
15	30,58	27,49	25	7,26	6,26	5,23
16	32	28,85	26,3	7,96	6,91	5,81
17	33,41	30,19	27,59	8,67	7,56	6,41
18	34,81	31,53	28,87	9,39	8,23	7,01
19	36,19	32,85	30,14	10,12	8,91	7,63
20	37,57	34,17	31,41	10,85	9,59	8,26
21	38,93	35,48	32,67	11,59	10,28	8,9
22	40,29	36,78	33,92	12,34	10,98	9,54
23	41,64	38,08	35,17	13,09	11,69	10,2
24	42,98	39,36	36,42	13,85	12,4	10,86
25	44,31	40,65	37,65	14,61	13,12	11,52
26	45,64	41,92	38,89	15,38	13,84	12,2
27	46,96	43,19	40,11	16,15	14,57	12,88
28	48,28	44,46	41,34	16,93	15,31	13,56
29	49,59	45,72	42,56	17,71	16,05	14,26
30	50,89	46,98	43,77	18,49	16,79	14,95
31	52,19	48,23	44,99	19,28	17,54	15,66
32	53,49	49,48	46,19	20,07	18,29	16,36
33	54,78	50,73	47,4	20,87	19,05	17,07
34	56,06	51,97	48,6	21,66	19,81	17,79
35	57,34	53,2	49,8	22,47	20,57	18,51
36	58,62	54,44	51	23,27	21,34	19,23
37	59,89	55,67	52,19	24,07	22,11	19,96
38	61,16	56,9	53,38	24,88	22,88	20,69
39	62,43	58,12	54,57	25,7	23,65	21,43
40	63,69	59,34	55,76	26,51	24,43	22,16
41	64,95	60,56	56,94	27,33	25,21	22,91
42	66,21	61,78	58,12	28,14	26	23,65
43	67,46	62,99	59,3	28,96	26,79	24,4
44	68,71	64,2	60,48	29,79	27,57	25,15
45	69,96	65,41	61,66	30,61	28,37	25,9
46	71,2	66,62	62,83	31,44	29,16	26,66
47	72,44	67,82	64	32,27	29,96	27,42
48	73,68	69,02	65,17	33,1	30,75	28,18

49	74,92	70,22	66,34	33,93	31,55	28,94
50	76,15	71,42	67,5	34,76	32,36	29,71

Критические точки распределения Пирсона (χ^2 -квадрат) можно определить также в Excel по формуле =ХИ2ОБР(α ; k) где α - уровень значимости (вероятность), k - число степеней свободы.

Число степеней свободы определяется как произведение разности количества столбцов/строк и единицы: $f = (l - 1) \times (c - 1)$. Т.е. для четырехпольной таблицы число степеней свободы составит $f = (2-1)*(2-1) = 1$. Далее по таблице 3 определяется критическое значение критерия χ^2 -квадрат Пирсона, которое при уровне значимости $p=0.05$ и числе степеней свободы 1 составляет 3.841, а при уровне значимости $p=0.01$ – 6.63.

Уровень значимости - это вероятность того, что различия, принятые существенными, на самом деле случайны. Уровень значимости отражает степень уверенности в точности, истинности полученных (прогнозируемых) данных. Общая статистическая значимость явления выражается коэффициентом p -value (p -уровень). В любом эксперименте или наблюдении существует вероятность, что полученные данные возникли из-за ошибок выборки. статистически значимой является величина, чья вероятность случайного возникновения крайне мала либо стремится к крайности. Крайностью в этом контексте считают степень отклонения статистики от гипотезы, которую проверяют на согласованность с полученными выборочными данными. Часто употребляемыми уровнями значимости являются 10% (0,1); 5% (0,05); 1% (0,01); и 0,1% (0,001). В научной практике уровень значимости выбирается перед сбором данных и, как правило, его коэффициент составляет 0,05 (5 %). Для систем, где крайне важны точные значения, этот показатель может составлять 0,01 (1 %) и менее. Когда указано, что различия достоверны на 5%-ом уровне значимости, или при $p < 0,05$, то вероятность того, что они недостоверны, составляет 0,05 или 5%.

4.2. Особенности расчета критерия χ^2 -квадрата с поправкой Йейтса

Метод расчета критерия χ^2 -квадрата с поправкой Йейтса аналогичен предыдущему. Поправка Йейтса делает общую оценку более умеренной. Метод можно применять, если процент доли меньше 10%, но более 5%. Т.е. поправка Йейтса вводится в том случае, если число ожидаемого явления меньше 10 (в абсолютных числах) хотя бы в одной ячейке (табл.4). Данная поправка позволяет уменьшить вероятность ошибки первого типа, т.е. обнаружения различий там, где их нет. Поправка Йейтса заключается в вычитании 0,5 из абсолютного значения разности между фактическим и ожидаемым количеством наблюдений в каждой ячейке, что ведет к уменьшению величины критерия хи-квадрат.

Формула для расчета критерия χ^2 с поправкой Йейтса следующая (2):

$$\rho = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{a!b!c!d!N!} \quad (2),$$

где i – номер строки (от 1 до l), j – номер столбца (от 1 до c), O_{ij} – фактическое количество наблюдений в ячейке ij , Φ_{ij} – ожидаемое число наблюдений в ячейке ij .

Использование поправки Йейтса ведет к уменьшению величины критерия χ^2 -квадрат.

5. Примеры определения критерия χ^2 -квадрата в атрибутивной статистике при анализе таблиц сопряженности

5.1. Определение статистической зависимости между курением и частыми обострениями хронической обструктивной болезни легких (ХОБЛ)

Материалами для исследования служили данные историй болезни пациентов, госпитализированных с обострением ХОБЛ (166 пациентов). Количество обострений составило 2 и более в течение года. Пациенты были разделены на две группы: курящие (N=110) и некурящие (N=56).

Первым этапом в анализе риска развития обострений ХОБЛ среди курящих пациентов является построение таблицы сопряженности (табл. 6). Расположение строк и столбцов этой таблицы находится в определенном порядке для того, чтобы показатели, рассчитанные из нее в дальнейшем, имели смысл и могли быть корректно интерпретированы. Первая строка отводится для группы, объекты которой имели исследуемый фактор риска, т.е. пациенты курящие пациенты с ХОБЛ. Вторая строка содержит показатели, характеризующие угрозу развития обострений ХОБЛ среди некурящих пациентов. Исследуемый фактор риска в данной группе отсутствовал. Таким образом, первой является группа, представляющая исследовательский интерес, т.е. та группа, пациенты которой подвергались фактору риска. Во второй строке должны быть результаты, полученные в той группе, с которой будет происходить сравнение эффекта. Аналогично, в первом столбце записывается, сколько раз исследуемое событие было зарегистрировано в первой и второй группах, во втором столбце – сколько раз оно отсутствовало.

Вначале формируется гипотеза о том, что курение среди пациентов с ХОБЛ является риском развития обострений данного заболевания (2 и > раз/год).

Таблица 6.

Таблица сопряженности

Наличие курения, как фактора риска	Обострение ХОБЛ (2 и более раз в год)		Всего
	Да	Нет	
Есть	79 (a)	31 (b)	110 (A)
Нет	18 (c)	38 (d)	56 (B)
Всего	97 (C)	69 (D)	166

На следующем этапе составляем на основании расчетов (табл.4) таблицу ожидаемых количеств наблюдений (табл.7).

Таблица 7.

Четырехпольная таблица ожидаемых количеств наблюдений для определения статистической зависимости между курением и частыми обострениями ХОБЛ

	Частые обострения ХОБЛ	Без частых обострений ХОБЛ	Всего
Наличие курения	64,27711	45,72289	110
Отсутствие курения	32,72289	23,27711	56
Всего	97	69	166

Далее по формуле (1) необходимо определить критерий χ^2 -квадрата, который составил 24,05.

Число степеней свободы составило 1. При уровне значимости (вероятности) $p \leq 0.05$ критическая точка распределения Пирсона (χ^2 -квадрат) составила 3.84. Полученное значение 24,05 превышает 3.84, следовательно, между курением и частыми обострениями ХОБЛ существует статистическая зависимость.

Далее рассмотрим пример с использованием поправки Йейтса. Анализируется такой же пример с развитием частых обострений у пациентов с ХОБЛ. Используем ту же таблицу сопряженности, что в рассмотренном

выше примере только с учетом того, что значение «с» в четырехпольной таблице составляет 5, что значит, что у 9 человек, страдающих ХОБЛ и некурящих, имели место частые обострения, а значение «d» составило 11, что значит у 11 некурящих пациентов не было частых обострений ХОБЛ. Таким образом, таблица несколько видоизменяется (табл. 8).

Таблица 8.

Таблица сопряженности

Наличие курения, как фактора риска	Обострение ХОБЛ (2 и более раз в год)		Всего
	Да	Нет	
Есть	79 (a)	31 (b)	110 (A)
Нет	9 (c)	38 (d)	47 (B)
Всего	88 (C)	69 (D)	157

Также на основании расчетов по таблице 4 составляем таблицу ожидаемых количеств наблюдений (табл. 9).

Таблица 9.

Четырехпольная таблица ожидаемых количеств наблюдений для определения статистической зависимости между курением и частыми обострениями ХОБЛ

	Частые обострения ХОБЛ	Без частых обострений ХОБЛ	Всего
Наличие курения	75,90361	74,09639	150
Отсутствие курения	8,096386	7,903614	16
Всего	84	82	166

Далее рассчитываем критерий χ^2 -квadrата с поправкой Йейтса по формуле (2), т.к. две ячейки в таблице ожидаемых количеств наблюдений имеют значения меньше 10. В нашем случае он составил 2.71.

Число степеней свободы составило 1. При уровне значимости (вероятности) $p \leq 0.05$ критическая точка распределения Пирсона (χ^2 -квadrат) составила 3.84. Полученное значение 2.71 меньше 3.84, следовательно,

между курением и частыми обострениями ХОБЛ нельзя установить статистическую зависимость.

Аналогично критерий χ^2 -квадрата Пирсона можно использовать при оценке таких атрибутивных данных, как результаты использования новых лекарственных препаратов, схем фармакотерапии, методов диагностики и лечения.

5.2. Определение статистической зависимости между включением антагониста лейкотриеновых рецепторов в схемы противоастматической терапии и частотой наступления положительных клинических исходов

Рассмотрим пример влияния антагониста лейкотриеновых рецепторов (монтелукаста) на частоту положительных клинических эффектов схем фармакотерапии при лечении обострений бронхиальной астмы у детей в условиях стационара.

Материалом для исследования служили данные фармакотерапии историй болезни 608 пациентов с бронхиальной астмой (БА), госпитализированных в стационар медицинской организации г. Нижнего Новгорода в 2014 -2015 г. Исследование носило ретроспективный характер.

К средствам базисной терапии относятся: ингаляционные глюкокортикостероиды (ИГКС) и системные глюкокортикостероиды (ГКС); антагонисты лейкотриеновых рецепторов (АЛР); длительно действующие β_2 -агонисты (ДДБА) в комбинации с ИГКС; кромоны (кромоглициевая кислота, недокромил натрия); пролонгированные теофиллины; антитела к IgE

В качестве базисной противоастматической терапии в стационаре в анализируемом периоде использовали следующие ЛП и их комбинации: 1) ИГКС (в качестве монотерапии); 2) комбинацию ИГКС и КДБА; 3) комбинацию ИГКС и ДДБА; 4) комбинацию ИГКС и АЛР; 5) комбинацию ИГКС, КДБА и АЛР; 6) комбинацию КДБА и АЛР.

Для выявления математико-статистической зависимости клинической эффективности схем противоастматической терапии при добавлении монтелукаста необходимо определение абсолютного количества положительных и отрицательных клинических исходов от назначенных схем лечения (табл.10).

Таблица 10.

Абсолютное значение наступления положительных и отрицательных клинических эффектов при использовании схем противоастматической терапии у детей

Схема терапии	Абсолютное количество положительных клинических исходов	Абсолютное количество отрицательных клинических исходов
ИГКС (в качестве монотерапии)	122	60
ИГКС и КДБА	55	14
ИГКС+ДДБА (комбинированные ЛП)	78	10
КДБА и АЛР	35	4
ИГКС и АЛР	84	7
ИГКС, КДБА и АЛР	64	4
ИГКС+ДДБА (комбинированные ЛП) +АЛР	63	5

Положительным клиническим эффектом от назначенной схемы лечения считалось наступление улучшения состояния (при выписке) от использования изначально назначенной схемы лечения без увеличения дозировок лекарственных препаратов и без добавления новых лекарственных препаратов к назначенной противоастматической терапии. В том случае, когда к схеме лечения были добавлены новые лекарственных препараты, такие как эуфиллин, системные ГКС и т.д., или - для достижения эффекта были увеличены дозировки и/или кратность применения лекарственным

препаратов, клинический эффект от изначально назначенной схемы терапии принимался отрицательным.

Первым этапом в анализе статистической зависимости между эффективностью и наличием монтелукаста в составе противоастматической терапии у детей являлось построение таблицы сопряженности (табл. 11).

Таблица 11.

Таблица сопряженности

Монтелукаст в составе противоастматической терапии	Наступление положительного клинического эффекта		Всего
	Да	Нет	
Есть	248 (a)	22 (b)	270 (A)
Нет	253 (c)	85 (d)	338 (B)
Всего	501 (C)	107 (D)	608 (Q)

Далее на основании таблицы 4 определяем ожидаемое количество наблюдений (табл. 12).

Таблица 12.

Четырехпольная таблица ожидаемых количеств наблюдений для определения статистической зависимости между включением монтелукаста в схемы противоастматической терапии и наступлением положительного клинического исхода

Монтелукаст в составе противоастматической терапии	Наступление положительного клинического эффекта		Всего
	Да	Нет	
Есть	222,4835526	47,51644737	270
Нет	278,5164474	59,48355263	338
Всего	501	107	608

По формуле (1) определяем критерий χ^2 -квадрата Пирсона, который составил 29.91. Сравниваем с критической точкой распределения Пирсона (χ^2 -квадрат), с 3.84 ($f=1$, $p=0.05$). Полученное значение превышает критическое, следовательно, можно утверждать о статистической

зависимости между добавлением монтелукаста в схемы противоастматической терапии и наступлением положительного клинического исхода.

6. Точный тест Фишера

Точный тест Фишера представляет собой тест статистической значимости, используемый в анализе таблиц сопряженности для выборок маленьких размеров. Назван именем своего изобретателя Р. Фишера.

Ранее было отмечено, что метод расчета критерия χ^2 -квадрата с поправкой Йейтса делает общую оценку более умеренной и может применяться, если процент доли меньше 10%, но более 5%. В том случае, когда процент доли меньше 5%, целесообразно использовать точный тест Фишера. Таким образом, критерий Фишера рассчитывается на малых выборках. Его использование на больших выборках трудоемко, требует много времени или мощных вычислительных ресурсов.

Точный критерий Фишера разделяется на односторонний и двусторонний. При одностороннем варианте точно известно, куда отклонится один из показателей. Например, во время исследования сравнивают, сколько пациентов выздоровело по сравнению с группой контроля. Предполагают, что терапия не может ухудшить состояние пациентов.

Двусторонний тест оценивает различия частот по двум направлениям. То есть, оценивается вероятность как большей, так и меньшей частоты явления в экспериментальной группе по сравнению с контрольной группой.

Для расчета критерия Фишера составляется таблица сопряженности (табл.13) аналогично рассмотренным выше, но с меньшим количеством наблюдений.

Таблица 13. Таблица сопряженности

Наличие курения, как фактора риска	Обострение ХОБЛ (2 и более раз в год)		Всего
	Да	Нет	
Есть	9 (a)	5 (b)	14 (A)
Нет	4 (c)	3 (d)	7 (B)
Всего	13 (C)	8 (D)	21

Расчет точного критерия Фишера проводим по формуле (3):

$$p = \sqrt{\frac{\chi^2}{n}} \quad (3),$$

где a, b, c, d – значения ячеек таблицы сопряженности; N - общее число исследуемых объектов/субъектов в двух группах; ! – факториал числа.

Факториал числа — математическое понятие, которое означает произведение всех натуральных чисел от 1 до этого числа. Например, факториал 4 записывается как 4! и составляет $4! = 4 \cdot 3 \cdot 2 \cdot 1$; $2! = 1 \cdot 2 = 2$.

Для расчета точного критерия Фишера можно воспользоваться различными on-line калькуляторами. Полученное значение p представляет уровень значимости различий сравниваемых групп. Далее необходимо сравнить полученное число с критическим уровнем значимости, обычно принимаемым в медицинских исследованиях за 0,05 или 0,01. Если значение точного критерия Фишера больше критического, то делается вывод об отсутствии статистически значимых различий частоты наступления события в зависимости от наличия воздействующего фактора. В том случае, если значение точного критерия Фишера меньше критического, то делается вывод о наличии статистически значимых различий частоты наступления события в зависимости от воздействия изучаемого фактора.

В представленной случае $p = 1,0000$, что $> 0,05$, следовательно, нельзя сделать вывод о наличии прямой взаимосвязи курения и вероятности развития частых обострений ХОБЛ. Частота возникновения частых обострений у курящих пациентов с ХОБЛ статистически незначимо выше, чем у некурящих.

Представленные методы: критерий χ^2 Пирсона, критерий χ^2 Пирсона с поправкой Йейтса, критерий Фишера позволяют оценить значимость различий исходов (наступления события) в зависимости от воздействия изучаемого фактора, т.е. позволяют определить, статистически значима или

незначима зависимость между воздействующим фактором и событием, к которому приводит действие этого фактора. После того как будет определено наличие статистической зависимости, встает вопрос, о силе связи между воздействующим фактором и событием, к которому он приводит.

7. Критерий ϕ и критерий Крамера (V) для оценки силы связи между воздействием анализируемого фактора и наступлением события

Критерии, оценивающие силу связи между номинальными переменными, могут принимать значения от 0 до 1.

Критерий ϕ предназначен взаимосвязи только для четырехпольных таблиц. Для многопольных таблиц целесообразнее применять критерий V Крамера.

Оба критерия основаны на критерии χ^2 и рассчитываются по формулам (4) и (5):

$$(4),$$

где χ^2 – критерий χ^2 Пирсона,

n – объем выборки ($Q, A+B+C+D$ в четырехпольной таблице сопряженности)

$$(5),$$

где χ^2 – критерий χ^2 Пирсона,

n – объем выборки ($Q, A+B+C+D$ в четырехпольной таблице сопряженности),

l – количество строк в таблице сопряженности,

r – количество столбцов в таблице сопряженности.

Для четырехпольных таблиц значения обоих критериев будут совпадать.

Для интерпретации полученных значений критериев ϕ и V Крамера используются интервалы значений (табл. 14).

Таблица 14

Интерпретация значений критериев ϕ и V Крамера (Гржибовский А.М., 2008)

Значение критериев ϕ или V Крамера	Сила взаимосвязи
<0,1	Несущественная
0,1 – <0,2	Слабая
0,2 – <0,4	Средняя
0,4 – <0,6	Относительно сильная
0,6 – <0,8	Сильная
0,8 – 1,0	Очень сильная

Так, в описанных выше примерах мы получили статистически значимую связь между фактом курения и развитием частых обострений у пациентов с ХОБЛ (табл. 6,7) ($\chi^2=24,05$), а также – между включением антагониста лейкотриеновых рецепторов в схемы противоастматической терапии и наступлением положительного клинического исхода (табл. 9, 10) ($\chi^2=29,97$).

В ходе расчета по формуле 4 критерия ϕ получаем для первого примера $\phi=0,381$; для второго - $\phi=0,179$.

Анализируя полученные значения критерия ϕ , можно говорить о средней взаимосвязи между фактом курения и развитием частых обострений ХОБЛ, однако эта связь существует и является статистически значимой. Включение антагониста лейкотриеновых рецепторов (монтелукаста) в схемы противоастматической терапии у детей статистически значимо приводит к увеличению частоты наступления положительных клинических исходов. Сила взаимосвязи является слабой для представленной выборки.

При наличии соотношения между вариацией качественных признаков говорят об их ассоциации, взаимосвязанности. Для оценки связи в этом случае используют ряд показателей, таких как коэффициент ассоциации (K_a) (формула 6), предложенный К. Пирсоном, и коэффициент контингенции (K_k) (формула 7), предложенный Дж. Юлом и Морисом Дж. Кендэллом, которые рассчитываются на основании таблицы сопряженности.

$$K_a = \frac{a \cdot d - b \cdot c}{\sqrt{(a+b) \cdot (b+d) \cdot (a+c) \cdot (c+d)}} \quad (6),$$

где a, b, c, d – значения ячеек таблицы сопряженности.

$$(7),$$

где a, b, c, d – значения ячеек таблицы сопряженности.

Коэффициент контингенции всегда меньше коэффициента ассоциации. Связь считается подтвержденной, если $K_a > 0,5$ или $K_k > 0,3$.

Анализировали связь между фактом курения и развитием частых обострений у пациентов с ХОБЛ (табл. 6,7) ($K_a = 0,687$; $K_k = 0,381$), а также – между включением антагониста лейкотриеновых рецепторов в схемы противоастматической терапии и наступлением положительного клинического исхода (табл. 9, 10) ($K_a = 0,582$; $K_k = 0,222$).

Для определения тесноты связи двух качественных признаков, каждый из которых состоит только из двух групп, применяют коэффициенты ассоциации и контингенции.

Коэффициенты ассоциации и контингенции неудобны для определения тесноты связи при анализе более двух признаков. В этом случае целесообразно использовать коэффициент сопряженности Пирсона (P), а также критерий Чупрова (K). Коэффициент сопряженности представляет собой меру оценки силы взаимосвязи, основанной на критерии χ^2 . Имеются сведения, что для малых таблиц (не более 5 x 5) более точную

оценку дает критерий Чупрова (К). Хотя эти оба коэффициента можно использовать и для четырехпольных таблиц сопряженности.

Коэффициент сопряженности Пирсона (Р) определяется по формуле (8).

$$K_{\chi} = \sqrt{\frac{\chi^2}{N \cdot (l-1)}} \quad (8),$$

где χ^2 – критерий Пирсона, рассчитанный по формуле (1),

N – объем выборки

Критерий Чупрова рассчитывается также на основе χ^2 – критерий Пирсона по формуле (9) и формуле (10).

$$K_{\chi} = \sqrt{\frac{\chi^2}{N \cdot (l-1) \cdot (c-1)}} \quad (9),$$

где χ^2 – критерий Пирсона, рассчитанный по формуле (1),

l – количество строк, при условии, что количество строк равно количеству столбцов ($l=c$).

N – объем выборки.

Формула (9) используется, если количество строк равно количеству столбцов в таблице сопряженности. В том случае, когда количество строк не равно количеству столбцов используется формула (10).

$$AP_{\chi} = \frac{a}{A} \quad (10),$$

где χ^2 – критерий Пирсона, рассчитанный по формуле (1),

l – количество строк в анализируемой таблице сопряженности,

c – количество столбцов в анализируемой таблице сопряженности,

N – объем выборки.

Для представленных примеров оценки влияния курения на развитие частых обострений ХОБЛ, а также влияния антагониста лейкотриеновых рецепторов в составе противоастматической терапии на достижение положительных клинических исходов, χ^2 составил 24,05 и 29,91, соответственно.

Коэффициент сопряженности Пирсона при оценке влияния курения на развитие частых обострений ХОБЛ, а также влияния антагониста лейкотриеновых рецепторов в составе противоастматической терапии на достижение положительных клинических исходов, P составил 0,356 и 0,380, соответственно, что указывает на умеренную связь.

Коэффициент Чупрова при оценке влияния курения на развитие частых обострений ХОБЛ, а также влияния антагониста лейкотриеновых рецепторов в составе противоастматической терапии на достижение положительных клинических исходов, K_c составил 0,217 и 0,222, соответственно, что указывает на умеренную связь.

Таким образом, представлена методика математико-статистической обработки полученных в ходе эксперимента и представленных в виде таблицы сопряженности данных, а именно определения наличия статистической связи между изучаемыми факторами и событиями, а также силы этой связи.

На основании математико-статистического анализа таблицы сопряженности могут быть получены количественные характеристики влияния изучаемых признаков на события.

8. Определение математико-статистической связи (количественные характеристики) между курением и развитием частых обострений у пациентов с ХОБЛ¹⁸

Вначале формируется гипотеза о том, что курение среди пациентов с ХОБЛ является риском развития обострений данного заболевания (2 и > раз/год). Основой проведения анализа является таблица сопряженности (табл. 3), в представленном случае – таблица 6. Первым является расчет абсолютного риска (АР) - доля больных от общего объема группы. В нашем случае это сводится к расчету частоты угрозы обострений ХОБЛ в группах курящих и некурящих пациентов. По формуле (11) находим частоту развития обострений ХОБЛ (2 и > раз/год) в экспонированной группе (курящие пациенты с ХОБЛ) 71,8%.

$$AP_n = \frac{c}{B} \quad (11)$$

То есть, в исследуемой экспонируемой выборке 0,781 - ее часть или 71,8% имеют угрозу развития обострений ХОБЛ (2 и > раз/год). Аналогично, по формуле (12) рассчитываем частоту развития обострений ХОБЛ в неэкспонированной группе (некурящие пациенты, страдающие ХОБЛ) - 32,1%.

$$S_{AP} = \sqrt{\frac{AP \times (1 - AP)}{n}} \quad (12)$$

Мы получили так называемые точечные оценки относительных частот развития обострений ХОБЛ (2 и > раз/год) в группе риска и контрольной группе. Точечные оценки подвержены статистической ошибке, поэтому возможно, что в следующей выборке будут получены другие значения, поскольку рассчитываем эти частоты на основе не всей популяции, а лишь ее репрезентативной части, приблизительно отражающей свойства популяции. Поэтому, далее необходимо рассчитать стандартную ошибку полученных

¹⁸Жукова О.В., Кононова С.В., Коньшкіна Т.М. Терапевтический архив. 2015. №3. С. 23-26.

AP, т.е. статистическую ошибку каждой из частот, которая дает представление о точности оценки. Стандартная ошибка AP экспонируемой группы, рассчитана по формуле (13) и составила 0,043.

$$ДИ_{AP} = AP \pm t \times S \quad (13)$$

где n - объем экспонируемой или неэкспонируемой группы, т.е. А или В.

Аналогично была рассчитана стандартная ошибка AP неэкспонируемой группы. Она составила 0,062.

Полученные частоты могут измениться при расчете на другой выборке. Поэтому нужно определить насколько существенны будут эти изменения, и какие минимальные интервалы значений покрывают реальные точные значения искомых частот. То есть, какой минимальный интервал содержит реальное значение искомой частоты с вероятностью, например, 95%. Такой интервал является в статистике 95% доверительным интервалом (95% ДИ). С практической точки зрения 95% доверительный интервал означает, что 95% всех потенциальных выборок дадут значения частот, попадающих в полученные интервалы, и лишь в 5% случаев значения частот выйдут за найденные пределы.

Рассчитаем по формуле (14) 95% ДИ абсолютного риска экспонируемой группы $0,718 \pm 0,084$ или $71,8\% \pm 8,4\%$.

$$AmP = AP_{\text{э}} - AP_{\text{н}} = \frac{a}{c} - \frac{c}{B}$$

(14)

где t – критическое значение для уровня статистической значимости. Для 95% ДИ $t=1,96$,

S – стандартная ошибка AP.

То есть, 95% ДИ принимает минимальное значение 63,4% и максимальное 80,2% со средним (AP_э) 71,8% и стандартной ошибкой 4,3%.

Таким образом, можно констатировать, что наличие курения в исследуемой группе обуславливает в $71,8\% \pm 4,3\%$ развитие обострений ХОБЛ (2 и >

раз/год), при этом, 95% возможных значений заболеваемости попадает в интервал от 63,4% до 80,2%.

Аналогичным образом, по формуле (14) рассчитано 95% ДИ абсолютного риска неэкспонируемой группы $0,321 \pm 0,122$ или $32,1\% \pm 12,2\%$.

То есть, при отсутствии воздействия курения в контрольной группе угроза развития обострений (2 и > раз/год) составляет $32,1\% \pm 12,2\%$. При этом, 95% возможных значений заболеваемости попадает в интервал от 19,9% до 44,3%.

Таким образом, были получены абсолютные показатели развития обострений ХОБЛ (2 и > раз/год) в группах, находящихся и не находящихся под воздействием фактора риска (курение). Отсюда следует, что вероятность развития обострений ХОБЛ под воздействием курения увеличивается. Однако, насколько существенный вклад вносит фактор риска в данное увеличение? Для этого рассчитывается атрибутивный риск (АТР), который характеризует собой именно ту часть риска (долю) развития обострений, которая связана с данным фактором риска и объясняется им. АТР был рассчитан по формуле (15) и составил 0,397 или 39,7%.

$$F = \frac{C}{Q} \quad (15)$$

То есть, в группе риска частота события составляет 71,8%, в контрольной группе – 32,1%, атрибутивный риск в этом случае равен 39,7%, то есть фактор риска увеличивает вероятность возникновения события на 39,7%.

Для расчета стандартной ошибки полученной разницы необходимо предварительно рассчитать объединенную оценку доли, которая составила $\approx 0,584$.

$$S_{AmP} = \sqrt{F \times (1 - F) \times \left(\frac{1}{A} + \frac{1}{B} \right)} \quad (16)$$

Далее, определяли стандартную ошибку АТР по формуле (17), которая составила 0,081 или 8,1%.

$$OP = \frac{AP_{\text{э}}}{AP_{\text{н}}} = \frac{a/A}{c/B} \quad (17)$$

Затем рассчитывали 95% ДИ атрибутивного риска: $0,397 \pm 0,159$ или $39,7\% \pm 15,9\%$

Таким образом, наличие курения у пациентов с ХОБЛ обуславливает увеличение развития обострений на $39,7\% \pm 15,9\%$, при этом, 95% возможных (истинных) значений разности заболеваемости попадает в интервал от 23,8% до 55,6%. Таким образом, возможные (истинные) значения, попадающие в 95% ДИ могут свидетельствовать о $AP_{\text{э}} > AP_{\text{н}}$, то есть, наличие «атипичных» инфекций увеличивает риск развития БА.

Методология оценки рисков основывается на статистических показателях (средняя, ошибка средней, ДИ), которые, в свою очередь, базируются на теории вероятности. Поэтому, рассчитывая те или иные риски, следует говорить не об абсолютной (точной) зависимости исхода от фактора, а о степени вероятности данной зависимости. В свою очередь, зависимость исхода от воздействующего фактора может быть резко выраженной и в этом случае статистически достоверная вероятность подобной зависимости обнаруживается при достаточно небольших выборках. Если же зависимость исхода от воздействия фактора имеется, но слабо выражена, то для обнаружения статистически достоверной ее вероятности необходимы большие по объему выборки, порой, даже в пределах популяции (например, население конкретного региона).

Аналогично рассмотренным особенностям анализа ДИ атрибутивного риска, необходимо рассмотрение ДИ всех рассчитываемых рисков. Например, в рассчитанных ранее $AP_{\text{э}}$ и $AP_{\text{н}}$ доверительные интервалы (63,4% и 80,2%, 19,9% и 44,3% соответственно) не включают в себя 0 или отрицательные значения и, поэтому, могут считаться статистически значимыми.

С помощью АТР показано, что наличие курения ведет к увеличению развития обострений ХОБЛ (2 и > раз/год) в среднем на 39,7%.

С помощью расчета относительных рисков (ОР) можно показать силу связи между воздействующим фактором риска и исходом, то есть, во сколько раз увеличивается вероятность обострений от наличия курения у пациентов с ХОБЛ. Если развитие обострений ХОБЛ выше в экспонируемой группе, то отношение $AR_{\text{э}}/AR_{\text{н}}$ будет больше 1, если ниже – то отношение $AR_{\text{э}}/AR_{\text{н}}$ будет меньше 1, если же вероятность в двух группах одинаковая, то их соотношение будет равно 1. Таким образом, согласно гипотезе, с математической точки зрения, необходимо доказать, что отношение развития обострений ХОБЛ в двух группах больше 1.

По формуле (18) рассчитаем ОР: 2,2.

$$PAR = \frac{C}{Q} - \frac{c}{B} \quad (18)$$

Популяционный атрибутивный риск, ПАР (англ. population attributable risk, PAR) - абсолютная разница показателей (или риска) во всей популяции и в неэкспонированной группе. ПАР аналогичен АТР, но в отличие от последнего характеризует популяционную составляющую риска и, поэтому, зависит от того, насколько широко распространены факторы риска в данной популяции, ПАР варьирует в зависимости от распространенности фактора риска в популяции.

По формуле (19) вычисляем значение ПАР: 0,263 или 26,3%.

$$S_{PAR} = \sqrt{F \times (1 - F) \times \left(\frac{1}{Q} + \frac{1}{B} \right)} \quad (19)$$

То есть, курение приводит к увеличению обострений ХОБЛ по всей популяции на 26,3%.

Для расчета стандартной ошибки полученной разницы необходимо предварительно рассчитать объединенную оценку доли. Она была рассчитана ранее согласно формулы (16) и составила 0,584.

По формуле (20) высчитываем среднюю ошибку ПАР: 0,076 или 7,6%

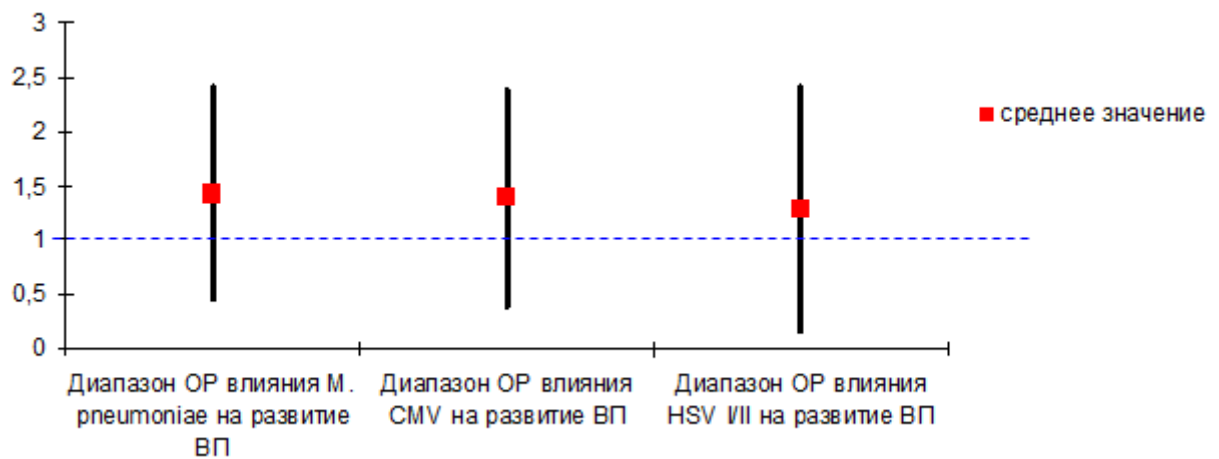
$$ДИ_{ПАР} = ПАР \pm t \times S \quad (20)$$

По формуле (21) рассчитываем 95% ДИ ПАР. Он составил $0,263 \pm 0,076$ или $26,3\% \pm 7,6\%$.

$$ИПВ(ЧБНЛ) = \frac{1}{AmP} \quad (21)$$

Таким образом, наличие курения увеличивает развитие обострений ХОБЛ в популяции на $26,3\% \pm 7,6\%$, при 95% ДИ от 18,7% до 33,9%.

Еще одним показателем, являющимся производным от АТР является индекс потенциального вреда (ИПВ) или число больных, которых необходимо лечить (ЧБНЛ). Если воздействующий фактор предположительно вызывает негативные изменения в состоянии здоровья, то речь идет о ИПВ. Если же воздействующий фактор улучшает состояние здоровья (например, новый вид лечения снижает количество рецидивов, введение профилактических мероприятий снижает заболеваемость), имеют в виду ЧБНЛ. Оба показателя представляют собой отношение 1 к АТР (22).



(22)

ИПВ курения в вышерассмотренном примере составляет: 2,5.

То есть, при наличии курения у пациентов с ХОБЛ у каждого третьего экспонируемого лица развиваются обострения (2 и > раз/год), дополнительно к фоновому уровню заболеваемости ХОБЛ.

**9. Определение математико-статистической связи
(количественные характеристики) между конкретными возбудителями и
развитием воспалительных заболеваний респираторного тракта
(внебольничная пневмония, острый бронхит)¹⁹**

В ходе исследования нами было проанализировано распространение редких и труднокультивируемых возбудителей, таких как *M. pneumoniae*, *S. pneumoniae*, *CMV*, *HSV I/II* среди детей при воспалительных заболеваниях респираторного тракта (внебольничная пневмония (ВП), острый бронхит (ОБ)). Также была выделена группа сравнения (n=127) (табл.15).

Таблица 15.

Частота выявления редких и труднокультивируемых возбудителей у детей.

	Количество случаев инфицирования			
	<i>M. pneumoniae</i>	<i>S. pneumoniae</i>	<i>CMV</i>	<i>HSV I/II</i>
ВП (n=341)	74	4	119	23
ОБ (n=360)	31	4	187	17
Группа сравнения (n=127)	2	0	19	2

На первом этапе анализа риска развития ВП среди детей были построены таблицы сопряженности (табл. 16, табл.17, табл. 18, табл. 19). Данные таблицы показывают гипотетическую зависимость развития ВП от инфицирования персистирующими инфекциями. Первая строка отводится для группы, объекты которой имели исследуемый фактор риска (персистирующая инфекция). Вторая строка содержит результаты так называемой «контрольной группы». Исследуемый фактор риска в данной группе отсутствовал. Таким образом, первой является группа, представляющая исследовательский интерес, т.е. та группа, пациенты

¹⁹ Жукова О.В., Бруснигина Н.Ф., Кононова С.В., Сперанская Е.В., Ефимов Е.И. Журнал инфектологии. 2016. Т. 8, №2. С. 79-82

которой подвергались фактору риска. Во второй строке должны быть результаты, полученные в той группе, с которой будет происходить сравнение эффекта. Аналогично, в первом столбце указывается, сколько раз исследуемое событие (ВП) было зарегистрировано в первой и второй группах, во втором столбце – сколько раз оно отсутствовало.

Таблица 16

Таблица сопряженности формирования ВП от инфицированности *M. pneumoniae*

<i>M. pneumoniae</i>	Наличие ВП		Всего
	Есть	Нет	
Есть	74 (a)	2 (b)	76 (A)
Нет	267 (c)	125 (d)	392 (B)
Всего	341 (C)	127 (D)	468 (Q)

Таблица 17

Таблица сопряженности формирования ВП от инфицированности *S. pneumoniae*

<i>S. pneumoniae</i>	Наличие ВП		Всего
	Есть	Нет	
Есть	4 (a)	0 (b)	4 (A)
Нет	337 (c)	127 (d)	464 (B)
Всего	341 (C)	127 (D)	468 (Q)

Таблица 18

Таблица сопряженности формирования ВП от инфицированности *CMV*

<i>CMV</i>	Наличие ВП		Всего
	Есть	Нет	
Есть	119 (a)	19 (b)	138 (A)
Нет	222 (c)	108 (d)	330 (B)
Всего	341 (C)	127 (D)	468 (Q)

Таблица 19

Таблица сопряженности формирования ВП от инфицированности *HSV I/II*

<i>HSV I/II</i>	Наличие ВП		Всего
	Есть	Нет	
Есть	23 (a)	2 (b)	25 (A)
Нет	318 (c)	125 (d)	443 (B)
Всего	341 (C)	127 (D)	468 (Q)

Вначале формируется гипотеза о том, что наличие *M. pneumoniae* у детей является риском развития ВП. Первым является расчет AP, который представляет собою долю больных от общего объема группы. В нашем случае это сводится к расчету частоты распространения ВП в группах обследуемых, инфицированных и не инфицированных *M. pneumoniae*, *S. pneumoniae*, *CMV*, *HSV I/II*. По формуле (11) находим частоту распространения ВП в экспонированной группе (инфицированные *M. pneumoniae* (1), *S. pneumoniae* (2), *CMV* (3), *HSV I/II* (4)) 97,37%, 100%, 92,97%, 92,00% (табл.20).

Получили, что в исследуемых экспонируемых выборках (определение развития ВП от наличия *M. pneumoniae* (1), *S. pneumoniae* (2), *CMV* (3), *HSV I/II* (4)) 0,9737; 1; 0,9297; 0,9200 - части соответствующих выборок или 97,37%; 100%; 92,97%; 92,00% имеют угрозу развития ВП. Аналогично, по формуле (12) рассчитываем частоту формирования ВП в неэкспонированной группе (не имеющей персистирующие инфекции. Она составила 68,11%; 72,63%; 67,27%; 71,78% для *M. pneumoniae* (1), *S. pneumoniae* (2), *CMV* (3), *HSV I/II* (4) соответственно.

В итоге были получены так называемые точечные оценки относительных частот развития ВП в группе риска (инфицирование персистирующими возбудителями) и контрольной группе (при отсутствии персистирующих инфекций). Точечные оценки подвержены статистической ошибке, поэтому возможно, что в следующей выборке будут получены другие значения, поскольку полученные частоты были рассчитаны на основе не всей популяции, а лишь ее репрезентативной части, лишь приблизительно отражающей свойства популяции. Поэтому, далее была рассчитана стандартная ошибка полученных AP, т.е. статистическая ошибка каждой из частот, которая дает представление о точности оценки. Стандартная ошибка AP экспонируемой группы, рассчитана по формуле (13) и составила 0,0184;

0; 0,0226; 0,0543 для *M. pneumoniae* (1), *S. pneumoniae* (2), *CMV* (3), *HSV I/II* (4) (табл.20).

Аналогично была рассчитана стандартная ошибка АР неэкспонируемой группы. Она составила 0,0235; 0,0207; 0,0258; 0,0214 для *M. pneumoniae* (1), *S. pneumoniae* (2), *CMV* (3), *HSV I/II* (4) (табл.20).

Однако полученные частоты могут измениться при расчете на другой выборке. Поэтому на следующем этапе исследования нами было определено, насколько существенны будут эти изменения, и какие минимальные интервалы значений покрывают реальные точные значения искомых частот, т.е., какой минимальный интервал содержит реальное значение искомой частоты с вероятностью 95%. Такой интервал является в статистике 95% доверительным интервалом (95% ДИ). С практической точки зрения 95% доверительный интервал означает, что 95% всех потенциальных выборок дадут значения частот, попадающих в полученные интервалы, и лишь в 5% случаев значения частот выйдут за найденные пределы. Чаще всего в исследованиях используется 95% или 99% ДИ.

Рассчитаем по формуле (14) 95% ДИ абсолютного риска экспонируемой группы 93,77% - 100,97%; 100%; 88,54% - 97,40%; 81,37% - 102,63% для *M. pneumoniae* (1), *S. pneumoniae* (2), *CMV* (3), *HSV I/II* (4) соответственно (табл.20).

Аналогичным образом, по формуле (4) рассчитано 95% ДИ абсолютного риска неэкспонируемой группы 65,50% - 72,73%; 68,57% - 76,69%; 62,21% - 72,34%; 67,59% - 75,97% для *M. pneumoniae* (1), *S. pneumoniae* (2), *CMV* (3), *HSV I/II* (4) соответственно (табл.20).

Таким образом, были получены абсолютные показатели заболеваемости в группах, находящихся и не находящихся под воздействием фактора риска. Отсюда следует, что развитие ВП под воздействием персистирующих возбудителей увеличивается. Однако, неизвестно, насколько существенный вклад вносит фактор риска в данное увеличение. Для этого рассчитывается атрибутивный риск (АТР), который представляет

именно ту часть риска (долю) развития болезни, которая связана с данным фактором риска и объясняется им. АТР был рассчитан по формуле (15) и составил 29,26%; 27,37%; 25,70%; 20,21% для *M. pneumoniae* (1), *S. pneumoniae* (2), *CMV* (3), *HSV III* (4) соответственно (табл.20).

Например, в группе риска (инфицирование *M. pneumoniae*) частота события (возникновение ВП) составляет 97,37%, в контрольной группе (возникновение ВП у детей без инфицирования *M. pneumoniae*) – 68,11%, АТР в этом случае равен 29,26%, то есть фактор риска увеличивает вероятность возникновения события на 29,26%.

Для расчета стандартной ошибки полученной разницы необходимо предварительно рассчитать объединенную оценку доли по формуле (16), которая составила $\approx 0,7286$.

Далее, определяли стандартную ошибку АТР по формуле (17) (табл.20).

Затем рассчитывали 95% ДИ АТР: 18,33% - 40,18%; -16,39% - 71,14%; 16,62% - 34,77%; 2,33% - 38,13% для *M. pneumoniae* (1), *S. pneumoniae* (2), *CMV* (3), *HSV III* (4) соответственно (табл.20).

Таким образом, наличие *M. pneumoniae* обуславливает увеличение развития ВП на $29,26\% \pm 10,93\%$. При этом, 95% возможных (истинных) значений разности заболеваемости попадает в интервал от 18,33% до 40,18%. Таким образом, возможные (истинные) значения, попадающие в 95% ДИ могут свидетельствовать о $AR_{\text{э}} > AR_{\text{н}}$. Это доказывает, что инфицирование *M. pneumoniae* увеличивает риск развития ВП.

Согласно концепции факторов риска, в случае, если $AR_{\text{э}} = AR_{\text{н}}$, то есть, воздействие фактора риска не изменяет вероятность наступления события. Если же $AR_{\text{э}} < AR_{\text{н}}$, то есть, воздействие фактора риска уменьшает вероятность наступления события.

Обращает на себя внимание нижний порог 95% ДИ для *S. pneumoniae*, который принимает отрицательное значение (-16,39%). При расчете АТР, мы предполагали, что наличие *S. pneumoniae* является фактором,

увеличивающим заболеваемость ВП, то есть разница между АР экспонируемых и АР неэкспонируемых должна быть больше 0. Полученная нами разница, то есть АТР, составляет 27,37%, однако 95% ДИ прироста заболеваемости от инфицированности *S. pneumoniae* включает в себя 0. Т.е., одним из 95% возможных (истинных) значений АТР может являться 0, что свидетельствует об отсутствии различий заболеваемости экспонируемой и неэкспонируемой групп. Таким образом, нельзя говорить с 95% уверенностью, что имеются различия заболеваемости между экспонируемыми (инфицирование *S. pneumoniae*) и неэкспонируемыми (отсутствие инфицирования *S. pneumoniae*).

Аналогично рассмотренным особенностям анализа ДИ атрибутивного риска, необходимо рассмотрение ДИ всех рассчитываемых рисков. Например, в рассчитанных ранее АРэ и АРн доверительные интервалы не включают в себя 0 или отрицательные значения и, поэтому, могут считаться статистически значимыми.

С помощью АТР показано, что инфицирование *M. pneumoniae*, *CMV*, *HSV III* ведет к увеличению развития ВП в среднем на 29,26%; 25,70%; 20,21% соответственно. Для *S. pneumoniae* полученные значения не являются статистически значимыми.

Далее нами был проведен расчет относительных рисков (ОР), с помощью которых можно показать силу связи между воздействующим фактором риска и исходом, то есть, во сколько раз увеличивается заболеваемость ВП от наличия персистирующих инфекций у детей. Если возникновение ВП выше в экспонируемой группе, то отношение АРэ/АРн будет больше 1, если ниже – то отношение АРэ/АРн будет меньше 1, если же заболеваемость в двух группах одинаковая, то их соотношение будет равно 1. Таким образом, согласно гипотезе, с математической точки зрения, необходимо доказать, что отношение развития ВП в двух группах (экспонируемые и неэкспонируемые) больше 1.

По формуле (18) рассчитаем ОР: 1,43 для *M. pneumoniae*, 1,38 для *S. pneumoniae* и для *CMV*, 1,28 для *HSV I/II*.

Однако, поскольку речь идет о расчетах, производимых на выборке, то необходимо показать статистическую значимость полученного результата. Для этого по формуле (19) была рассчитана стандартная ошибка ОР (табл. 20).

По формуле (20) рассчитываем 95% ДИ относительного риска (табл. 20).

Полученные значения ОР составляют с 95% ДИ от 0,45 до 2,41 для *M. pneumoniae*; от 0,39 до 2,37 для *CMV*, от 0,15 до 2,41 для *HSV I/II*. Для *S. pneumoniae* ОР не удалось определить. Т.о., можем предположить, что инфицирование *S. pneumoniae* не играет статистически важной роли в развитии ВП.

Таким образом, значения 95% ДИ, попадающие в диапазон менее 1, свидетельствуют о возможной несостоятельности гипотезы о влиянии персистирующих инфекций на формирование ВП.

Однако средние значения ОР для *M. pneumoniae*, *CMV*, *HSV I/II* находятся в интервале выше 1, а также большой отрезок коридоров колебаний ОР приходится на интервал больше 1 (рис.2).

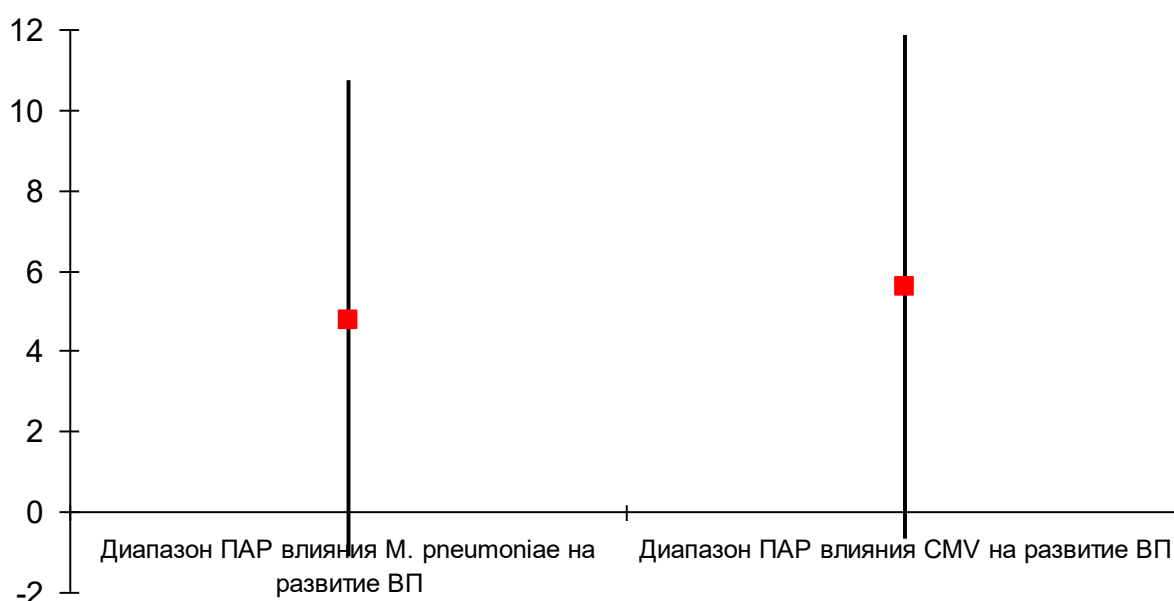


Рисунок 2. Коридоры колебаний значений ОР с 95% ДИ возникновения ВП от инфицирования персистирующими возбудителями.

По формуле (21) вычисляем значение ПАР:

То есть, наличие персистирующих возбудителей приводит к увеличению заболеваемости ВП по всей популяции на 4,75% для *M. pneumoniae*, 0,23% для *S. pneumoniae*, 5,59% для *CMV*, 1,08% для *HSV I/II*.

По формуле (22) высчитываем среднюю ошибку ПАР. По формуле (23) рассчитываем 95% ДИ ПАР.

Таким образом, согласно полученным статистическим результатам, наличие *M. pneumoniae* увеличивает заболеваемость ВП в популяции при 95% ДИ от -1,12% до 10,72%; наличие *S. pneumoniae* увеличивает заболеваемость ВП в популяции при 95% ДИ от -5,48% до 5,94%; наличие *CMV* увеличивает заболеваемость ВП в популяции при 95% ДИ от -0,67% до 11,86%; наличие *HSV I/II* увеличивает заболеваемость ВП в популяции при 95% ДИ от -4,69% до 6,86%.

Обращают на себя внимание отрицательные значения нижней границы 95% ДИ, что по аналогии с АТР указывает на возможную несостоятельность гипотезы о влиянии персистирующих инфекций на формирование ВП. Подтверждением недостоверности ПАР является высокое значение стандартной ошибки для *S. pneumoniae*, в 10 раз превышающее среднее значение и для *HSV I/II*, более чем в 2 раза превышающее среднее значение. Как следствие, такие показатели включают в свой диапазон отрицательные значения и 0. Эти положения позволят нам исключить *S. pneumoniae* и *HSV I/II* из факторов риска развития ВП. Для *M. pneumoniae* и *CMV* нижние показатели 95% ДИ имеют отрицательные значения. Но наибольший отрезок коридоров колебаний ПАР приходится на положительные значения (рис.3).



Рисунок 3. Коридоры колебаний значений ПАР с 95% ДИ возникновения ВП от инфицирования персистирующими возбудителями.

Таблица 20.

Результаты статистического анализа влияния персистирующих возбудителей на формирование ВП, полученные на основании концепции рисков

		Влияние возбудителей на развитие ВП у детей			
		<i>M. pneumoniae</i>	<i>C. pneumoniae</i>	<i>CMV</i>	<i>HSV I/II</i>
1. Абсолютный риск					
	<i>в экспонируемой группе</i>	0,973684	1	0,9296875	0,92
	<i>в не экспонируемой группе</i>	0,681122	0,726293103	0,672727273	0,717832957
2. Стандартная ошибка абсолютного риска					
	<i>в экспонируемой группе</i>	0,018362	0	0,022598509	0,05425864
	<i>в не экспонируемой группе</i>	0,023539	0,020698546	0,025829587	0,021382721
3. Доверительный интервал абсолютного риска					
	<i>в экспонируемой группе</i>	0,937695 - 1,009673	1	0,885394423 - 0,973980577	0,813653066 - 1,026346934
	<i>в не экспонируемой группе</i>	0,634987 - 0,727258	0,685723953 -	0,622101283 -	0,675922823 -

			0,766862254	0,723353262	0,759743091
4. Атрибутивный риск	0,292562		0,273706897	0,256960227	0,202167043
5. Объединенная оценка доли	0,728632		0,728632479	0,728632479	0,728632479
6. Стандартная ошибка атрибутивного риска	0,055732		0,223288902	0,046302451	0,09140801
7. Доверительный интервал атрибутивного риска	0,183327- 0,401797		-0,163939351 - 0,711353144	0,166207423- 0,347713031	0,023007343 - 0,381326743
8. Относительный риск	1,429529		1,376854599	1,381967905	1,28163522
9. Стандартная ошибка относительного риска	0,039369		0,028498888	0,045442962	0,066072525
10. Доверительный интервал относительного риска	0,448512 - 2,410545		-	0,393800127 - 2,370135684	0,153473299 - 2,409797141
11. Популяционный атрибутивный риск	0,04751		0,002339375	0,055905206	0,010799522
12. Стандартная ошибка популяционного риска	0,030445		0,029131258	0,031963531	0,029475935
13. Доверительный интервал популяционного риска	-0,01216 - 0,107182		-0,05475789 - 0,05943664	-0,006743314 - 0,118553726	-0,046973312 - 0,068572355

Т.о., в результате проведенного анализа нами было получено, что инфицирование

M. pneumoniae и *CMV* является факторами риска развития ВП. АТР *M. pneumoniae* составил 29,26% (с 95% ДИ 18,33% - 40,18%). АТР *CMV*- 25,67% (с 95% ДИ 16,62% - 34,77%). Воздействие *M. pneumoniae* приводит к увеличению заболеваемости ВП в среднем на 29,26%, а воздействие *CMV* - на 25,67%.

ОР составил 1,43 (с 95% ДИ от 0,45 до 2,41) для *M. pneumoniae*. ОР для *CMV* составил 1,38(с 95% ДИ от 0,39 до 2,37). Т.е. инфицирование *M.*

рneumonіae приводит к увеличению заболеваемости ВП у детей в 1,43 раза, а инфицирование *CMV* – в 1,38 раза.

ПАР составил 4,75% для *M. рneumonіae* (с 95% ДИ от -1,21 до 10,72%) и 5,59% для *CMV* (с 95% ДИ от -0,67% до 11,86%). Инфицирование *M. рneumonіae* приводит к увеличению заболеваемости ВП по всей популяции на 4,75%, а инфицирование – на 5,59%.

Полученные данные позволяют говорить об усугубляющей роли *M. рneumonіae* и *CMV* в развитии ВП. Однако инфицирование *M. рneumonіae*, *CMV* не является достаточной причиной возникновения ВП у детей.

Инфицирование *S. рneumonіae* и *HSV I/II* не играет статистически важной роли в развитии ВП.

Аналогично были проведены расчеты для определения влияния персистирующих возбудителей на развитие ОБ. В начале были составлены таблицы сопряженности формирования ОБ от каждого из анализируемых возбудителей (табл. 21, табл. 22, табл. 23, табл. 24).

Таблица 21

Таблица сопряженности формирования ОБ от инфицированности *M. рneumonіae*

<i>M. рneumonіae</i>	ОБ		Всего
	Есть	Нет	
Есть	31 (a)	2 (b)	33 (A)
Нет	329 (c)	125 (d)	454 (B)
Всего	360 (C)	127 (D)	487 (Q)

Таблица 22

Таблица сопряженности формирования ОБ от инфицированности *S. рneumonіae*

<i>S. рneumonіae</i>	Наличие ОБ		Всего
	Есть	Нет	
Есть	4 (a)	0 (b)	4 (A)
Нет	356 (c)	127 (d)	483 (B)
Всего	360 (C)	127 (D)	487 (Q)

Таблица 23

Таблица сопряженности формирования ОБ от инфицированности *CMV*

<i>CMV</i>	Наличие ОБ	Всего
------------	------------	-------

	Есть	Нет	
Есть	187 (a)	19 (b)	206 (A)
Нет	173 (c)	108 (d)	281 (B)
Всего	360 (C)	127 (D)	487 (Q)

Таблица 24

Таблица сопряженности формирования ОБ от инфицированности HSV I/II

HSV I/II	Наличие ОБ		Всего
	Есть	Нет	
Есть	17 (a)	2 (b)	19 (A)
Нет	343 (c)	125 (d)	468 (B)
Всего	360 (C)	127 (D)	487 (Q)

После был выполнен статический анализ определения влияния каждого возбудителя на формирование ОБ (табл. 25).

Таблица 25.

Результаты статистического анализа влияния персистирующих возбудителей на формирование ОБ, полученные на основании концепции рисков.

		Влияние возбудителей на развитие ОБ у детей			
		<i>M. pneumoniae</i>	<i>C. pneumoniae</i>	<i>CMV</i>	<i>HSV I/II</i>
1. Абсолютный риск					
	<i>в экспонируемой группе</i>	0,939393939	1	0,90776699	0,894736842
	<i>в не экспонируемой группе</i>	0,724669604	0,737060041	0,615658363	0,732905983
2. Стандартная ошибка абсолютного риска					
	<i>в экспонируемой группе</i>	0,041536025	0	0,02016029	0,0704059
	<i>в не экспонируемой группе</i>	0,02096376	0,020031163	0,029018533	0,020451882
3. Доверительный интервал абсолютного риска					
	<i>в экспонируемой группе</i>	0,85798333 - 1,020804549	1	0,868252822 - 0,947281159	0,756741277 - 1,032732407
	<i>в не экспонируемой группе</i>	0,683580634 -	0,697798962 -	0,558782039 -	0,692820293 -

		0,765758573	0,776321121	0,672534687	0,772991672
4. Атрибутивный риск		0,214724336	0,262939959	0,292108627	0,161830859
5. Объединенная оценка доли		0,739219713	0,739219713	0,739219713	0,739219713
6. Стандартная ошибка атрибутивного риска		0,079159626	0,220437293	0,040271876	0,102751673
7. Доверительный интервал атрибутивного риска		0,059571469 - 0,369877203	-0,169117136 - 0,694997053	0,21317575 - 0,371041504	-0,039562421 - 0,363224139
8. Относительный риск		1,29630653	1,356741573	1,474465458	1,22080712
9. Стандартная ошибка относительного риска		0,052838478	0,027177111	0,052104248	0,08349042
10. Доверительный интервал относительного риска		0,222925903 - 2,369687158		0,510977903 - 2,437953013	-0,003400301 - 2,44501454
11. Популяционный атрибутивный риск		0,014550109	0,002159671	0,12356135	0,00631373
12. Стандартная ошибка популяционного риска		0,028643531	0,028194992	0,03289176	0,028420939
13. Доверительный интервал популяционного риска		-0,041591212 - 0,07069143	-0,053102513 - 0,057421855	0,059093499 - 0,1880292	-0,049391312 - 0,062018771

Статистически значимыми показателями являются АТР, ОР, ПАР.

Т.о., в результате проведенного анализа нами было получено, что инфицирование *M. pneumoniae* и *CMV* является факторами риска развития ОБ. АТР *M. pneumoniae* составил 21,47% (с 95% ДИ 5,96% - 36,99%). АТР *CMV*- 29,21% (с 95% ДИ 21,32% - 37,10%). Воздействие *M. pneumoniae* приводит к увеличению заболеваемости ВП в среднем на 21,47%, а воздействие *CMV* - на 29,21%. АТР *S. pneumoniae* составил 26,29% (с 95% ДИ -16,91% - 69,50%). АТР *HSV I/II* составил 16,18% (с 95% ДИ -3,96% - 36,32%).

ОР составил 1,30 (с 95% ДИ от 0,22 до 2,37) для *M. pneumoniae*. ОР для *CMV* составил 1,47 (с 95% ДИ от 0,51 до 2,44) (рис.4).

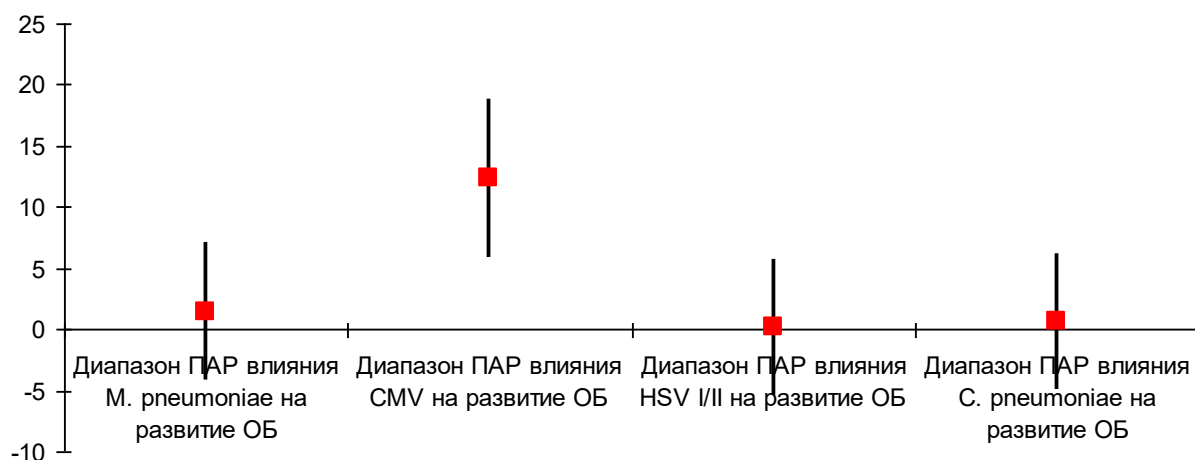


Рисунок 4. Коридоры колебаний значений ОР с 95% ДИ возникновения ОБ от инфицирования персистирующими возбудителями

Т.е. инфицирование *M. pneumoniae* приводит к увеличению заболеваемости ОБ у детей в 1,30 раза, а инфицирование *CMV* – в 1,47 раза. ОР составил 1,36 (95% ДИ не удалось определить) для *S. pneumoniae*. Для *HSV I/II* ОР составил 1,22 (с 95% ДИ от -0,34 до 2,45).

ПАР составил 1,46% для *M. pneumoniae* (с 95% ДИ от -4,16 до 7,07%) и 12,36% для *CMV* (с 95% ДИ от 5,91% до 18,80%). Значения 95% ДИ для ПАР *M. pneumoniae* не играют статистически важной роли в развитии ОБ. Инфицирование *CMV* приводит к увеличению заболеваемости ОБ по всей популяции на 12,36%. ПАР для *S. pneumoniae* составил 0,22% (с 95% ДИ от -5,31 до 5,74%) и 0,63% - для *HSV I/II* (с 95% ДИ от -4,94 до 6,20%) (рис.5).

$$AЭэ = \frac{a}{A}$$

Рисунок 5. Коридоры колебаний значений ПАР с 95% ДИ возникновения ОБ от инфицирования персистирующими возбудителями.

Инфицирование *S. pneumoniae* и *HSV I/II* не играет статистически важной роли в развитии ОБ. Большие отрезки отрицательных значений 95%

ДИ АТР, ОР, ПАР для *S. pneumoniae* и для *HSV III* указывают на невозможность включения данных возбудителей в значимые факторы риска ОБ.

Полученные данные позволяют говорить об усугубляющей роли *CMV* и относительно усугубляющей роли *M. pneumoniae* в развитии ОБ. Однако инфицирование *M. pneumoniae*, *CMV* не является достаточной причиной возникновения ОБ у детей.

Статистически значимыми показателями являются АТР, ОР, ПАР. АТР развития ВП составил 29,26% (18,33% - 40,18%); 27,37% (-16,39% - 71,14%); 25,70% (16,62% - 34,77%); 20,21% (2,33% - 38,13%) для *M. pneumoniae*, *S. pneumoniae*, *CMV*, *HSV III* соответственно. ОР составил 1,43 (0,45 – 2,41) для *M. pneumoniae*; 1,38 – для *S. pneumoniae* (не удалось определить доверительный интервал) и для *CMV* (0,39 - 2,37); 1,28 (0,15 – 2,41) – для *HSV III*. Т.о., можем предположить, что инфицирование *S. pneumoniae* не играет статистически важной роли в развитии ВП. Наличие персистирующих возбудителей приводит к увеличению заболеваемости ВП по всей популяции (ПАР) на 4,75% (-1,12% - 10,72%) для *M. pneumoniae*, 0,23% (-5,48% - 5,94%) для *S. pneumoniae*, 5,59% (-0,67% - 11,86%) для *CMV*, 1,08% (-4,69% - 6,86%) для *HSV III*. Отрицательные значения нижней границы 95% ДИ, что по аналогии с АТР указывает на возможную несостоятельность гипотезы о влиянии персистирующих инфекций на формирование ВП. Подтверждением недостоверности ПАР является высокое значение стандартной ошибки для *S. pneumoniae*, в 10 раз превышающее среднее значение и для *HSV III*, более чем в 2 раза превышающее среднее значение. Как следствие, такие показатели включают в свой диапазон отрицательные значения и 0. Эти положения позволяют нам исключить *S. pneumoniae* и *HSV III* из факторов риска развития ВП. Для *M. pneumoniae* и *CMV* нижние показатели 95% ДИ имеют отрицательные значения. Но наибольший отрезок коридоров колебаний ПАР приходится на положительные значения

Полученные данные позволяют говорить об усугубляющей роли *M. pneumoniae* и *CMV* в развитии ВП. Однако инфицирование *M. pneumoniae*, *CMV* не является достаточной причиной возникновения ВП у детей.

Инфицирование *S. pneumoniae* и *HSV I/II* не играет статистически важной роли в развитии ВП.

Инфицирование *S. pneumoniae* и *HSV I/II* не играет статистически важной роли в развитии ОБ. Большие отрезки отрицательных значений 95% ДИ АТР, ОР, ПАР для *S. pneumoniae* и для *HSV I/II* указывают на невозможность включения данных возбудителей в значимые факторы риска ОБ. Полученные данные свидетельствуют об усугубляющей роли *CMV* и относительно усугубляющей роли *M. pneumoniae* в развитии ОБ. Однако инфицирование *M. pneumoniae*, *CMV* не является достаточной причиной возникновения ОБ у детей.

10. Определение клинической эффективности лекарственного препарата на основании математико-статистического анализа

Оценка эффективности при добавлении монтелукаста в схему противоастматической терапии строилась на определении абсолютной эффективности в группе пациентов, получающих монтелукаст в составе противоастматической терапии, и в группе пациентов, не получающих монтелукаст в составе противоастматической терапии, атрибутивной эффективности, относительной эффективности, популяционной атрибутивной эффективности, а также на определении стандартных ошибок для каждого вида эффективности и доверительного интервала.

Первым этапом в анализе статистической зависимости между эффективностью и наличием монтелукаста в составе противоастматической терапии у детей являлось построение таблицы сопряженности (табл. 5). Расположение строк и столбцов этой таблицы находится в определенном порядке для того, чтобы показатели, рассчитанные из нее в дальнейшем, имели смысл и могли быть корректно интерпретированы. Первая строка отводится для группы, пациенты которой получали монтелукаст в составе противоастматической терапии. Вторая строка содержит пациентов, не получавших монтелукаст в составе противоастматической терапии. Таким образом, первой является группа, представляющая исследовательский интерес, т.е. та группа, пациенты которой получали анализируемый лекарственный препарат. Во второй строке находятся результаты, полученные в той группе пациентов, с которой проводилось сравнение эффективности. Аналогично, в первом столбце записывается, сколько раз исследуемое событие (наступление положительного клинического эффекта) было зарегистрировано в первой и второй группах, во втором столбце – сколько раз оно отсутствовало.

После построения таблицы сопряженности формируется гипотеза: введение монтелукаста в схему противоастматической терапии увеличивает

количество положительных клинических исходов при БА, не требуя замены ЛП, увеличения дозировки используемых и добавления новых ЛП. Первым является расчет абсолютной эффективности (АЭ) - доля клинических эффектов от общего объема группы. В нашем случае это сводится к расчету частоты наступления положительных клинических эффектов в группах пациентов, получавших и не получавших монтелукаст в составе противоастматической терапии. По формуле (23) находим частоту наступления положительных клинических исходов в экспонированной группе (пациенты, получавшие монтелукаст в составе противоастматической терапии) 91,85%.

$$AЭ_n = \frac{c}{B} \quad (23)$$

То есть, в исследуемой экспонируемой выборке 0,9185 ее часть или 91,85% имеют положительные клинические эффекты. Аналогично, по формуле (24) рассчитываем частоту наступления положительных клинических эффектов в неэкспонированной группе (пациенты, не получавшие монтелукаст) 74,85% (абсолютная эффективность в неэкспонированной группе).

$$S_{AЭ} = \sqrt{\frac{AЭ \times (1 - AЭ)}{n}} \quad (24)$$

В результате были получены точечные оценки наступления положительных клинических исходов от назначения схем противоастматической терапии, включающих и не включающих монтелукаст (экспонируемая и неэкспонируемая группы пациентов). Эти частоты были рассчитаны на основе не всей популяции, а лишь ее репрезентативной части, приблизительно отражающей свойства популяции. Такие точечные оценки подвержены статистической ошибке. Поэтому, далее рассчитывали стандартную ошибку полученных АЭ, т.е. статистическую ошибку каждой из

частот, которая дает представление о точности оценки. Стандартная ошибка АЭ экспонируемой группы, рассчитана по формуле (25) и составила 0,017.

$$ДИ_{AЭ} = AЭ \pm t \times S \quad (25)$$

где n - объем экспонируемой или неэкспонируемой группы, т.е. А или В.

Аналогично была рассчитана стандартная ошибка АЭ неэкспонируемой группы. Она составила 0,024.

Т.к. полученные частоты могут измениться при расчете на другой выборке, то было определено насколько существенны будут эти изменения, и какие минимальные интервалы значений покрывают реальные точные значения искомых частот. То есть, какой минимальный интервал содержит реальное значение искомой частоты с вероятностью 95%. Такой интервал является в статистике 95% доверительным интервалом (95% ДИ). С практической точки зрения 95% доверительный интервал означает, что 95% всех потенциальных выборок дадут значения частот, попадающих в полученные интервалы, и лишь в 5% случаев значения частот выйдут за найденные пределы. Чаще всего в исследованиях используется 95% или 99% ДИ (формула 26).

$$AmЭ = AЭэ - AЭн = \frac{a}{c} - \frac{c}{B} \quad (26)$$

где t – критическое значение для уровня статистической значимости. Для 95% ДИ $t=1,96$,

S – стандартная ошибка АЭ.

Т.о., 95% ДИ принимает минимальное значение 88,59% и максимальное 95,12% со средним (АЭэ) 91,85% и стандартной ошибкой 1,70%.

Таким образом, можно констатировать, что использование противоастматической терапии, включая монтелукаст, в исследуемой группе обуславливает 91,85%±1,70% случаев наступления положительных клинических исходов.

Аналогичным образом, по формуле (4) рассчитано 95% ДИ абсолютной эффективности в неэкспонируемой группе 0,7485±0,0236 или 74,85%±2,36%

Таким образом, были получены абсолютные показатели эффективности в группах, получавших и не получавших монтелукаст. Отсюда следует, что увеличение количество наступления положительных клинических исходов при включении в схему монтелукаста увеличивается. Далее было определено, насколько существенный вклад вносит добавление монтелукаста в данное увеличение? Для этого была рассчитана атрибутивная эффективность (АтЭ), которая характеризует ту часть эффективности (долю), которая связана с исследуемым ЛП и объясняется им. АтЭ была рассчитана по формуле (27) и составил 0,1700 или 17,00%.

$$S_{AmЭ} = \sqrt{F \times (1 - F) \times \left(\frac{1}{A} + \frac{1}{B} \right)} \quad (27)$$

То есть, в группе, включавшей монтелукаст, вероятность наступления положительного клинического исхода составила 91,85%, в группе без монтелукаста – 74,85%, АтЭ в этом случае равна 17,00%, то есть монтелукаст увеличивает вероятность наступления события (положительного клинического исхода) на 17,00%.

Для определения стандартной ошибки полученной разницы предварительно была рассчитана объединенная оценка доли по формуле 16, которая составила 0,8240.

Далее была определена стандартная ошибка АтЭ по формуле (28), которая составила 0,0311 или 3,11%.

$$ОЭ = \frac{АЭэ}{АЭн} = \frac{a/A}{c/B} \quad (28)$$

Затем был определен 95% ДИ атрибутивной эффективности: 10,91% - 23,09%.

Таким образом, присутствие монтелукаста в схеме противоастматической терапии у детей, госпитализированных с БА обуславливает увеличение наступления положительного клинического эффекта на 17,00%±6,09%. Таким образом, возможные (истинные) значения,

попадающие в 95% ДИ свидетельствуют о $AЭэ > AЭн$, то есть, монтелукаст увеличивает вероятность наступления положительного клинического эффекта от противоастматической терапии.

В случае, если $AЭэ = AЭн$, то есть, ЛП не изменяет вероятность наступления события, в данном случае положительного клинического эффекта. Если же $AЭэ < AЭн$, то ЛП уменьшает вероятность наступления события.

Методология оценки клинической эффективности основывается на статистических показателях (средняя, ошибка средней, ДИ), которые, в свою очередь, базируются на теории вероятности.

Аналогично рассмотренным особенностям анализа ДИ атрибутивной эффективности, необходимо рассмотрение ДИ всех полученных значений эффективности. Например, в рассчитанных ранее $AЭэ$ и $AЭн$ доверительные интервалы (88,59% и 95,12%, 70,23% и 79,48%.соответственно) не включают в себя 0 или отрицательные значения и, поэтому, могут считаться статистически значимыми.

С помощью $AtЭ$ показано, что использование монтелукаста ведет к увеличению вероятности наступления положительного клинического эффекта в среднем на 17,00%.

С помощью расчета относительной эффективности (ОЭ) была показана сила связи между влиянием ЛП на лечение и исходом, то есть, во сколько раз увеличивается клиническая эффективность схем противоастматической терапии при использовании монтелукаста. Если клиническая эффективность выше в экспонируемой группе, то отношение $AЭэ/AЭн$ будет больше 1, если ниже – то отношение $AЭэ/AЭн$ будет меньше 1, если же клиническая эффективность в двух группах одинаковая, то их соотношение будет равно 1. Таким образом, согласно гипотезе, с математической точки зрения, было доказано, что клиническая эффективность в двух группах больше 1.

По формуле (29) определена ОЭ: 1,23.

$$S_{OЭ} = \sqrt{\frac{1 - AЭэ}{a} + \frac{1 - AЭн}{c}} \quad (29)$$

То есть, использование монтелукаста в противоастматической терапии привело к увеличению клинической эффективности в 1,23 раза. Однако, поскольку речь идет о расчетах, производимых на выборке, то была показана статистическая значимость полученного результата. Для этого по формуле (30) была рассчитана стандартная ошибка ОЭ: 0,0364.

$$ДИ_{OP} = OP \pm Exp\left(\ln \frac{APэ}{APн} \times t \times S\right) \quad (30)$$

По формуле (31) рассчитан 95% ДИ относительной эффективности.

$$ПАЭ = \frac{C}{Q} - \frac{c}{B} \quad (31)$$

Полученные значения ОЭ составили с 95% ДИ от 0,21 до 2,24. По аналогии с АтЭ, возможные (истинные) значения, попадающие в 95% ДИ могут свидетельствовать о:

1. ОЭ > 1, то есть, использование ЛП увеличивает клиническую эффективность;
2. ОЭ = 1, то есть, использование ЛП не изменяет клиническую эффективность;
3. ОЭ < 1, то есть, использование ЛП уменьшает клиническую эффективность.

Рассчитанные значения 95% ДИ не позволяют говорить о состоятельности гипотезы об однозначном увеличении клинической эффективности с математико-статистической точки зрения.

Популяционная атрибутивная эффективность (ПАЭ) - абсолютная разница показателей во всей популяции и в неэкспонированной группе. ПАЭ аналогична АтЭ, но в отличие от последней характеризует популяционную составляющую эффективности и, поэтому, зависит от того, насколько широко распространено использование монтелукаста в данной популяции.

По формуле (32) было определено значение ПАЭ: 0,0755 или 7,55%.

$$S_{ПАЭ} = \sqrt{F \times (1 - F) \times \left(\frac{1}{Q} + \frac{1}{B} \right)} \quad (32)$$

То есть, использование монтелукаста в составе противоастматической терапии приводит к увеличению клинической эффективности во всей популяции на 7,55%.

Для расчета стандартной ошибки полученной разницы предварительно рассчитана объединенная оценка доли. Она была рассчитана ранее согласно формулы (16) и составила 0,8240.

По формуле (33) была определена стандартная ошибка ПАЭ: 0,0258 или 2,58%.

$$ДИ_{ПАЭ} = ПАЭ \pm t \times S \quad (33)$$

По 0, формуле (34) был рассчитан 95% ДИ ПАЭ. Он составил 2,49%-12,61%.

(34)

Таким образом, назначение монтелукаста увеличивает клиническую эффективность от противоастматической терапии в популяции при 95% ДИ от 2,49% до 12,61%.

11. Тестовые задания

Выберите один правильный ответ

1. СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ – ЭТО:

- 1) научная организация регистрации информации;
- 2) оценка и регистрация признаков изучаемой совокупности;
- 3) работа по сбору массовых первичных данных;
- 4) обширная программа статистических исследований.

2. СТАТИСТИЧЕСКАЯ ГРУППИРОВКА - ЭТО:

- 1) объединение данных в группы по времени регистрации;
- 2) расчленение изучаемой совокупности на группы по существенным признакам;
- 3) образование групп зарегистрированной информации по мере ее поступления.

3. ГРУППИРОВОЧНЫЕ ПРИЗНАКИ, КОТОРЫМИ ОДНИ ЕДИНИЦЫ СОВОКУПНОСТИ ОБЛАДАЮТ, А ДРУГИЕ - НЕТ, КЛАССИФИЦИРУЮТСЯ КАК:

- 1) факторные;
- 2) атрибутивные;
- 3) альтернативные.

4. К КАКИМ ГРУППИРОВОЧНЫМ ПРИЗНАКАМ ОТНОСЯТСЯ: ОБРАЗОВАНИЕ СОТРУДНИКОВ, ПРОФЕССИЯ БУХГАЛТЕРА, СЕМЕЙНОЕ ПОЛОЖЕНИЕ:

- 1) к атрибутивным;
- 2) к количественным.

5. СТАТИСТИЧЕСКИЙ ПОКАЗАТЕЛЬ - ЭТО

- 1) размер изучаемого явления в натуральных единицах измерения
- 2) количественная характеристика свойств в единстве с их качественной определенностью
- 3) результат измерения свойств изучаемого объекта

6. В КАКИХ ЕДИНИЦАХ БУДЕТ ВЫРАЖАТЬСЯ ОТНОСИТЕЛЬНЫЙ ПОКАЗАТЕЛЬ, ЕСЛИ БАЗА СРАВНЕНИЯ ПРИНИМАЕТСЯ ЗА ЕДИНИЦУ?

- 1) в процентах
- 2) в натуральных
- 3) в коэффициентах

7. СУММА ВСЕХ УДЕЛЬНЫХ ВЕСОВ ПОКАЗАТЕЛЯ СТРУКТУРЫ

- 1) строго равна 1
- 2) больше или равна 1
- 3) меньше или равна 1

8. ЧТО ПОНИМАЕТСЯ В СТАТИСТИКЕ ПОД ТЕРМИНОМ «ВАРИАЦИЯ ПОКАЗАТЕЛЯ»?

- 1) изменение величины показателя
- 2) изменение названия показателя
- 3) изменение размерности показателя

9. ПОКАЗАТЕЛЬ ДИСПЕРСИИ - ЭТО:

- 1) квадрат среднего отклонения
- 2) средний квадрат отклонений
- 3) отклонение среднего квадрата

10. ВЫБОРОЧНЫЙ МЕТОД В СТАТИСТИЧЕСКИХ ИССЛЕДОВАНИЯХ ИСПОЛЬЗУЕТСЯ ДЛЯ:

- 1) экономии времени и снижения затрат на проведение статистического исследования;
- 2) повышения точности прогноза;
- 3) анализа факторов взаимосвязи.

11. МАЛАЯ ВЫБОРКА - ЭТО ВЫБОРКА ОБЪЕМОМ:

- 1) 4-5 единиц изучаемой совокупности;
- 2) до 50 единиц изучаемой совокупности;
- 3) до 30 единиц изучаемой совокупности.

12. ПРОГНОЗИРОВАНИЕ В СТАТИСТИКЕ - ЭТО:

- 1) предсказание предполагаемого события в будущем;
- 2) оценка возможной меры изучаемого явления в будущем.

13. СТАТИСТИЧЕСКАЯ СВЯЗЬ - ЭТО:

- 1) когда зависимость между факторным и результирующим показателями неизвестна;
- 2) когда каждому факторному соответствует свой результирующий показатель;
- 3) когда каждому факторному соответствует несколько разных значений результирующего показателя.

14. ТЕРМИН КОРРЕЛЯЦИЯ В СТАТИСТИКЕ ПОНИМАЮТ КАК:

- 1) связь, зависимость;
- 2) отношение, соотношение;
- 3) функцию, уравнение.

15. ПО НАПРАВЛЕНИЮ СВЯЗЬ КЛАССИФИЦИРУЕТСЯ КАК:

- 1) линейная;
- 2) прямая;
- 3) обратная.

16. ПРИ КАКОМ ЗНАЧЕНИИ КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ СВЯЗЬ МОЖНО СЧИТАТЬ УМЕРЕННОЙ?

- 1) $r = 0,43$;
- 2) $r = 0,71$.

17. ДЛЯ ОПРЕДЕЛЕНИЯ ТЕСНОТЫ СВЯЗИ ДВУХ АЛЬТЕРНАТИВНЫХ ПОКАЗАТЕЛЕЙ ПРИМЕНЯЮТ:

- 1) коэффициенты ассоциации и контингенции;
- 2) коэффициент Спирмена.

12. Ситуационные задачи

Задача 1. При проверке гипотезы о нормальном распределении генеральной совокупности при заданном уровне значимости α определено критическое значение критерия $\chi^2_{кр} = \chi^2_{кр}(\alpha; k) = 14,4$. Чему должно быть равно наблюдаемое значение статистического критерия, чтобы эмпирические и теоретические частоты различались значимо?

Задача 2. Для того чтобы при заданном уровне значимости $\alpha=0,01$ проверить гипотезу о нормальном распределении генеральной совокупности, если тесны эмпирические и теоретические частоты необходимо определить критическую точку. Определите критическую точку.

Эмпирические частоты	1	6	34	18	6
Теоретические частоты	2	14	32	15	2

13. Эталоны ответов для тестовых заданий и ситуационных задач

Эталоны ответов для тестовых заданий

1. 3	9. 2	17. 1
2. 2	10. 1	
3. 3	11. 3	
4. 1	12. 2	
5. 2	13. 2	
6. 3	14. 1	
7. 1	15. 2	
8. 1	16. 1	

Эталон ответа для ситуационной задачи №1

Согласно критерию согласия Пирсона, для проверки гипотезы о нормальном распределении генеральной совокупности строят правостороннюю критическую область, определяемую неравенством вида $\chi^2 > \chi^2_{кр}$, то есть для того чтобы обеспечить попадание критерия χ^2 (хи-квадрат) в критическую область, необходимо обеспечение справедливости неравенства $\chi^2_{набл.} > 14,4$.

Эталон ответа для ситуационной задачи №2

Согласно критерию согласия Пирсона, для проверки гипотезы о нормальном распределении генеральной совокупности строят правостороннюю критическую область, определяемую неравенством вида $\chi^2 > \chi^2_{кр}$. Чтобы обеспечить попадание критерия χ^2 в критическую область, необходимо найти критическую точку $\chi^2_{кр} = \chi^2_{кр}(\alpha; k)$ по таблице критических точек распределения хи-квадрат с числом степеней свободы $k=s-3$, где по условию $s=5$ —число групп (частичных интервалов) выборки.

Т.о., получаем $\chi^2_{кр} = \chi^2_{кр}(0,01; 2)$

14. Список рекомендуемой литературы.

Основная литература:

1. Управление и экономика фармации: учебник // под ред. И.А.Наркевичв. – М.: ГЭОТАР-Медиа, 2019, 928 с.
2. Лоскутова Е.Е. Управление и экономика фармации: Учебник // Под ред. Е.Е. Лоскутовой– М.: Академия, 2008.
3. Багирова В.Л. Управление и экономика фармации: Учебник. М.: Медицина, 2008, 720 с.

Дополнительная литература:

1. Реброва О.Ю. Статистический анализ медицинских данных. М.: МедиаСфера, 2002. - 312 с.
2. Хабриев Р.У., Ягудина Р.И., Правдюк Н.Г. Оценка технологий здравоохранения: Монография. М.: Медицинское информационное агентство. 2013. -405 с
3. Воробьев П. А. Клинико-экономический анализ. М.: Ньюдиамед. 2008. 777с.
4. Авксентьева М. В., Воробьев П. А., Герасимов В. Б., Горохова С. Г., Кобина С. А. Экономическая оценка эффективности лекарственной терапии (фармакоэкономический анализ). М.: Ньюдиамед. 2000 г.
5. [Оценка медицинских технологий // под общей редакцией член-корр. РАМН, Белоусова Ю. Б. М.: ООО"Издательство ОКИ". 2013 г.](#)
6. [Основные понятия в оценке медицинских технологий. // Под общей ред. Колбина А.С., Белоусова Д.Ю., Зырянова С.К. М.: Издательство ОКИ, 2013. 42 с.](#)
7. Максимов С.А., Зинчук С.Ф., Давыдова Е.А., Зинчук В.Г. Риски и их оценка в медико-биологических исследованиях. Кемерово: КемГМА; 2010. 28с.

8. Зубов Н.Н., Умаров С.З., Бунин С.А. Математические методы и модели в фармацевтической науке и практике. СПб.: Изд-во Политехн,