**Лекция1.3.4.\_Систематическая и случайные ошибки в эпидемиологических исследованиях**

Слайд 2.

При планировании любого исследования, равно как и при оценке качества публикации первым встает вопрос о валидности исследования, т.е. его значимости. При этом можно говорить о внутренней валидности – т.е. условной мере соответствия результатов исследования истине и внешней валидности – условной мере применимости результатов данного исследования к другим группам больных, т.е. репрезентативности результатов.

Т.о. внутренняя валидность показывает нам, насколько найденная взаимосвязь отражает процессы и явления, существующие в реальности.

Слайд 3.

Угроза внутренней валидности исследования создается при наличии систематической или случайной ошибок

Слайд 4.

Мы знаем, что любые наблюдения подвержены влиянию случайности. Это абсолютно неизбежно. Случайная ошибка (СлО) – отклонение результата (отдельного) наблюдения в выборке от истинного значения в популяции, обусловленное исключительно случайностью. Именно поэтому результаты исследования одного и того же вопроса, проведенные разными исследователями, всегда будут хотя бы немного отличаться друг от друга.

Например, если предположить, что два препарата одинаково эффективны и каждый из них вызывает улучшение примерно у половины больных, то в единичном исследовании с небольшим числом больных в сравниваемых группах вполне может оказаться (исключительно из-за чистой случайности), что прием препарата А в большем проценте случаев дает положительный эффект, чем препарат Б или наоборот.

Случайная ошибка может быть выявлена при использовании соответствующих дизайну исследования статистических методов, а избежать ее можно, правильно рассчитав необходимый объем исследуемой выборки.

Увеличение размера группы будет уменьшать вероятность случайной ошибки.

Так, если монету подбросить 2 раза, в силу случайности она дважды может упасть кверху орлом. Можем ли мы сделать в этом случае вывод, что монета всегда падает орлом кверху? Конечно же нет. Это произошло случайно. Если подбросить монету 10 раз, возможно, шансы для выпадания орла и решки будут соотноситься как 4:6. 100 раз – 48:52 и т.д. Чем больше будет количество подбрасываний монеты, тем ближе будет результат к истинному.

Слайд 5.

СисО представляет более серьезную проблему. Под СисО понимают систематическое, неслучайное отклонение результатов и выводов от истины… обусловленное особенностями сбора, анализа, интерпретации, публикации или обзора данных, которое может привести к заключениям, систематически отличным от истины. Синонимом термина СисО являются предвзятость, искажение результатов, bias (англ), что в переводе также означает отклонение, предвзятость, пристрастность.

Систематическая ошибка – это серьезная проблема для исследователя. Связано это с тем, что, в отличие от случайных, она не элиминируются при увеличении выборки. Систематические ошибки нельзя исправить, если данные уже собраны. Статистических (или иных) тестов, оценивающих вероятность систематической ошибки не существует. Это ошибка, отсутствие которой доказать наиболее сложно.

При этом СисО практически сводит на нет весь труд по проведению исследования, поскольку может смещать значение меры эффекта воздействия как к нулю, так и от нуля (нулевой гипотезы). Таким образом, из-за СисО мы можем получить альфа-ошибку, т.е. говорить о наличии различий в тех ситуациях, когда таковых в действительности нет и также – бета-ошибку, т.е. мы можем не увидеть различия в тех ситуациях когда они имеют место быть. Размер ошибки может быть различным, от минимальной, до огромной.

Хотелось бы отметить, что СисО редко является следствием истинной предвзятости, нечестности исследователя. Что способствует ее возникновению? Этому может способствовать незнание особенностей воздействия, исхода, их взаимосвязи или методов их оценки. Надо сказать, что каждый дизайн исследования ассоциируется с характерными для него СО. Полностью избежать СисО в исследовании сложно, хотя бы потому, что существует много объективных факторов, способных повлиять на результаты исследования, и не все эти факторы известны данному исследователю и даже современной науке. Задача исследователя сводится к минимизации риска СО.

Слайд 6.

Систематические ошибки возможны на любом этапе исследования. Наиболее ответственным этапом является этап планирования. Но на последующих этапах – сбора, анализа, интерпретации данных и даже этапе публикации возможно появление систематических ошибок. Применительно к этапу публикации,проблемой является то, что журналы охотнее публикуют результаты исследований, выявляющих наличие какой-либо связи, чем ее отсутствие. Это приводит к т.н. публикационному смещению. В силу этого при использовании для метаанализа только опубликованных статей неизбежно возникнет эффект смещения суммарных результатов. В метаанализ не будут включены результаты, которые не были опубликованы, а они могут противоречить опубликованным.

Риск систематических ошибок можно минимизировать, если тщательно планировать исследование, аккуратно его проводить, правильно анализировать и интерпретировать данные.

Слайд 7.

Существует большое количество видов СО, которые характерны для разных видов исследований.

Систематические ошибки условно можно разделить на две большие группы:

Предвзятость отбора

Предвзятость информации (предвзятость измерения).

Единой и устойчивой классификации систематических ошибок не существует (даже в англоязычной литературе). Ряд исследователей относит к видам систематической ошибки также эффект конфаундинга. О феномене конфаундинга мы сегодня говорить не будем. Этому будет посвящена отдельная лекция.

Слайд 8.

В качестве примера, сегодня обсудим два наиболее частых варианта СисО. Первая из них - предвзятость отбора или систематическая ошибка отбора – это ошибка, возникающая при наличии неслучайных отличий между вошедшими в исследование и не вошедшими в него, а также между исследуемой и контрольной группой. В итоге результат, полученный среди участников исследования будет отличается от гипотетического результата, который был бы получен, если бы в него вошли все люди, соответствующие критериям отбора и распределение по группам было бы случайным.

Особенности механизма систематической ошибки отбора зависят от вида исследования. Однако независимо от дизайна исследования, группы должны быть идентичны, за исключением изучаемого параметра. Например, нельзя сравнивать эффект лекарства у амбулаторных пациентов с таковым у госпитализированных. Ошибка отбора будет иметь место в том случае, если при оценке прогноза заболевания будут анализироваться лишь истории пациентов, посещающих врача, т.е. приверженных пациентов и игнорироваться неприверженные и т.д. Например, при организации РКИ с целью оценки эффекта лекарственного вещества, мы должны взять две абсолютно идентичные по возрасту, полу, тяжести заболевания, опыту лечения группы, что достигается использованием инструментов рандомизации. Одна группа будет получать лекарство, другая плацебо. И это должно быть единственное различие между группами.

Слайд 9.

Пример ошибки отбора: На этапе отбора в исследование, пациентам с исходно более тяжелыми проявлениями болезни врач, следуя порыву помочь больному, будет стараться назначить препарат А (новое, очень популярное средство), а не Б (старое средство), тогда как препарат Б будет назначаться больным с более легкими формами заболевания. Вследствие этого, исходные группы будут отличаться по распределению тяжести заболевания, и результаты не будут валидными, не будут соответствовать истине.

Слайд 10.

Еще один пример. В когортных исследованиях по изучению рисков и пользы от употребления алкоголя, была выявлена «J-образная» связь между употреблением алкоголя и смертностью. Лучший исход (по показателям преждевременной смерти) отмечался в группе умеренно пьющих людей. Трезвенники со значительно большей вероятностью погибали в более молодом возрасте, чем обычный человек, потреблявший умеренное количество алкоголя.

Почему был получен такой результат? Можем ли мы сделать вывод, что употребление алкоголя не влияет на продолжительности жизни? Конечно, нет.

При подробном анализе группы т.н. трезвенников оказалось, что

Популяция трезвенников включала не только людей, ведущих здоровый образ жизни, но и людей, вынужденных бросить пить по состоянию здоровья, т.е. тех, кто не может больше пить, тех, кто помимо отказа от алкоголя изменил диету и образ жизни в целом, включает определенные религиозные или этнические группы, а также тех, кто пьет «как рыба», но предпочитает об этом не распространяться. Т.е. В этом случае нечетко были продуманы и обозначены критерии отбора исследуемых в группы, что привело к возникновению систематической ошибки отбора и получению неверного вывода.

Слайд 11.

Следующий вид ощибки - СисО, обусловленная измерением, возникает, когда в сравниваемых группах больных используются разные методы измерения.

Например, даже результаты измерения уровня АД будут существенно отличаться – разница может составить до 15 мм рт. ст - в зависимости от того, где, как, кем будет производиться это измерение, как сидит больной, как расположена его рука и т.д. Поэтому методы измерения и сбора информации должны быть четко оговорены и утверждены до мельчайших деталей на этапе планирования исследования. И технология измерения эффекта должна быть одинаковой для всех пациентов, включенных в исследование. Так, например, если оцениваемым исходом является развитие рака молочной железы, нужно четко определить, каким образом в данном исследовании будет выставляться диагноз РМЖ – будет ли диагноз основан на маммографии, результатах УЗИ или результатах гистологического исследования.

Еще один пример такого рода ошибки: Представим, что мы решили выяснить связь между использованием контрацептивов и развитием флеботромбоза у женщин, госпитализированных по поводу флеботромбоза и по другим причинам. Формируем две группы для ИСК: с флеботромбозом и без, а в качестве источника информации о приеме контрацептивов берем данные из историй болезни. В данном случае будет иметь место ошибка измерения (информации), связанная с тем, что вероятность того, что врач расспросит женщину с флеботромбозом о приеме в прошлом оральных контрацептивов будет значительно выше, чем в случае госпитализации женщины с другим диагнозом. В силу этого информация в историях болезни будет неполной, она не может использоваться для анализа. В противном случае, полученные результаты будут некорректными, не валидными.

Слайд 12.

Еще один пример такого рода ошибки: Представим, что мы решили выяснить связь между использованием контрацептивов и развитием флеботромбоза у женщин, госпитализированных по поводу флеботромбоза и по другим причинам. Формируем две группы для ИСК: с флеботромбозом и без, а в качестве источника информации о приеме контрацептивов берем данные из историй болезни. В данном случае будет иметь место ошибка измерения (информации), связанная с тем, что вероятность того, что врач расспросит женщину с флеботромбозом о приеме в прошлом оральных контрацептивов будет значительно выше, чем в случае госпитализации женщины с другим диагнозом. В силу этого информация в историях болезни будет неполной, она не может использоваться для анализа. В противном случае, полученные результаты будут некорректными, не валидными.

В заключение хотелось бы отметить, что тщательный анализ на предмет наличия систематических ошибок необходим не только перед проведением исследования, но и при чтении публикаций. Даже в очень хороших, высокорейтинговых журналах могут встретиться публикации с неправильными выводами, полученными в результате систематической ошибки. В качестве примера можно привести, например, опубликованную в 1999 году в журнале Lancet статью о связи вакцинации от кори-паротита-краснухи с развитием аутизма у детей. Результатом этой публикации явилось снижение доверия врачей вакцинации, снижение мотивации населения, уменьшение охвата прививками и, в конечном итоге, вспышки и эпидемии кори в самых разных странах. В последствии оказалось , что данные результаты были получены вследствие систематической ошибки и неправильно спланированного исследования. Статья была изъята, а автор ее дисквалифицирован.

Очень важно уметь оценивать правильность выполнения исследования в соответствии с правилами, используемыми для минимизации систематических ошибок для того, чтобы опираться в работе на достоверные факты и не стать жертвой ошибки. Методам контроля систематических ошибок будет посвящена следующая лекция.

Слайд 14.

Начнем с ошибок, типичных для исследования «случай-контроль». Данный тип исследования наиболее подвержен разного рода предвзятостям, о которых необходимо помнить для получения валидных результатов.

Напомню, что в ИСК разделение на группы происходит не по принципу наличия (отсутствия) воздействия, а в зависимости от наличия или отсутствия того или иного исхода. Затем собирают сведения (путем опроса или поиска в медицинских документах) о возможном воздействии на участников исследования в прошлом того или иного фактора и оценивается частота воздействия изучаемого фактора в прошлом для каждой группы. Важной задачей при организации такого рода исследований является составление определения болезни или случая и правильный отбор «случаев» и «контролей». С какими «подводными камнями» здесь можно столкнуться?

Слайд 15.

Наиболее частым типом ошибки в исследованиях типа СК является т.н. ошибка, или предвзятость, классификации. Это ошибка в определении экспозиции и/или исхода. Соответственно, экспонированные участники классифицируются как неэкспонированные и наоборот. Аналогично – для определения возникновения исхода – люди с исходом (например, заболеванием) классифицируются как здоровые и наоборот.

Слайд 16.

Предвзятость выбора участников группы «случай» возможна при использовании тестов с недостаточной чувствительностью или специфичностью, а также при использовании не очень надежных источников информации. Как правило, принятие решения о наличии заболевания или его отсутствии основано на комбинации клинических симптомов, лабораторных, инструментальных и патологических исследований. Чрезвычайно важно то, какие критерии применяются для идентификации того, у кого есть заболевание. Если используются неспецифические критерии, большинство, но не все люди с этим заболеванием будут выявлены. В то же время многие люди которые не имеют заболевания, будут включены ошибочно. Например, если для определения случаев инфаркта миокарда в качестве диагностического критерия использовать боль в груди, большинство, но не все, случаи сердечного приступа будут включены (при этом люди с «немыми, безболевыми формами сердечного приступа будут пропущены), но люди с другими заболеваниями, которые могут вызвать боль в груди (например, язва желудка) будут включены ошибочно.

Слайд 17.

Теперь: что касается контрольной группы. Контрольная группа нужна нам для того, чтобы оценить распределение экспозиции в популяции, из которой произошли «случаи». При этом что важно: участники этой группы должны быть набраны в то же время и из той же популяции, т.е. должны иметь равную со «случаями» возможность быть подвергнутыми воздействию изучаемого фактора. В качестве примера систематической ошибки можно привести исследование о превентивной роли прохождения ПАП-мазка в отношении развития рака шейки матки (РШМ). В качестве «случаев» отбирались пациенты с только что выявленным РШМ, госпитализированные в онкологическое отделение, «контролей» – женщины, проживающие в этом же районе, набранные при «обходах» квартир в рабочее время. В обеих группах собиралась информация о прохождении ПАП-мазка в течение года, предшествовавшего выявлению РШМ (для группы «случай»), и года, предшествовавшего дате включения в исследование (для группы «контроль»).

Слайд 18.

Здесь представлены результаты данного исследования (левая четырехполная таблица) , в финале которого получено ОШ=1, означающее отсутствие какого-либо положительного эффекта от взятия Пап-мазка. Фактически речь идет об отсутствии пользы от Пап-теста в отношении своевременнорго выявления изменений шейки матки на стадии предрака - дисплазии. Однако этим результатам, видимо, не стоит доверять, поскольку данное исследование страдает предвзятостью выбора контроля. Случилось это потому, что исследователи включили в контрольную группу только женщин, которые были дома во время рабочего дня. Довольно высока вероятность того, что эти женщины не трудоустроены или социально неблагополучны или очень пожилые. Следовательно, они реже будут проходить регулярные медицинские осмотры, включая взятие ПАП-теста. Таким образом, нарушается одно из важных правил организации ИСК, гласящее, что контрольная группа должна быть частью той же популяции, из которой происходят пациенты группы «случая», а частота экспозиции (в данном случае взятия ПАП-мазка) должна равняться таковой в общей популяции. Истинный результат, полученный при правильном отборе контролей, представлен справа. Мы видим, что ОШ = 0,4, что позволяет сделать совсем другой вывод – о протективной роли данного теста. Мы видим, что наличие систематической ошибки отбора контролей в данном исследовании привело к нивелированию эффекта воздействия исследуемого фактора.

Слайд19.

Следующий вид ошибки, имеющий отношение к отбору пациентов - так называемая систематическая ошибка Бергсона. Она может возникать, когда субъектом изучения являются госпитализированные пациенты. Связана она с тем, что вероятность наличия исследуемого фактора риска среди госпитализированных может быть значительно выше, чем в целом в популяции. В результате этого - явления, не имеющие связи на популяционном уровне, в группе госпитализированных пациентов могут неожиданно выглядеть статистические связанными. Так, например, если при изучении влияния курения на возникновение рака лёгких в качестве случая отобрать пациентов госпитализированных с этим диагнозом, а в качестве группы сравнения пациентов, госпитализированных с другим диагнозом, связь рака легкого с курением может оказаться недооцененной. Причиной ошибочной оценки в данном примере может стать то обстоятельство, что курение, кроме рака лёгкого, вызывает множество других заболеваний, могущих привести к госпитализации. Поэтому доля курящих в группе сравнения окажется, скорее всего, выше, чем среди популяции в целом.

Слайд 20.

Для минимизации ошибок классификации (отбора в группы)необходимо:

Для отбора пациентов в группы использовать четкие и недвусмысленные критерии. В группу «случай» лучше включать пациентов со «свежим» диагнозом, т.е. анализировать показатели заболеваемости, а не распространенности. Для исключения заболевания в группе «контроль» нужно использовать высокочувствительные тесты. Участники контрольной группы должны быть набраны в то же время и из той же популяции, т.е. должны иметь равную со «случаями» возможность быть подвергнутыми воздействию изучаемого фактора.

Еще один важный принцип заключается в том, что контрольные образцы должны отбираться независимо от статуса воздействия. Другими словами, экспонированные и неэкспонированные члены популяции должны иметь одинаковую вероятность отбора их для исследования

Слайд 21

Необходимо помнить также о т.н. «ошибке самоотбора». Она заключается в том, что как «случаи», так и «контроли» могут выказать различную степень желания участвовать в исследовании в зависимости от наличия экспозиции и/или заболевания

Обычно доля участия выше для экспонированных «случаев»: в исследовании более охотно будут участвовать люди, подозревающие у себя наличие изучаемого заболевания, если для этого есть основания как, например, отягощенный семейный анамнез. Риск возникновения заболевания у таких людей может быть выше, чем у других членов популяции, что также приведет к отличиям от естественного уровня заболеваемости. Этим объясняется негативный эффект данной ошибки на результаты исследования. При этом степень такого отклонения от естественного уровня практически невозможно установить, поскольку данные о тех, кто отказался участвовать в исследовании отсутствует или крайне скудные.

Близко к ошибке самоотбора лежит ошибка, связанная с нежеланием части исследуемой популяции участвовать в исследовании. Например если исследование проводится с применением опросника, может возникнуть систематическая ошибка, связанная с тем, что не все, получившие опросник, возвращают его с ответами, а соотношение исследуемых факторов среди не вернувших анкету может быть совсем иным. В этом случае важно знать, чем отличаются те, кто вернул заполненный опросник от тех, кто этого не сделал.

Слайд 22.

Как избежать ошибки самоотбора?

Универсальное правило для этого – высокие доли участия, т.е. небольшой процент тех, кто не смог участвовать, например, не вернул анкету. Оптимальной считается доля участия в 80% и выше.

Если возможно – необходимо оценить особенности людей, участвующих в исследовании и не участвующих.

Необходимо проводить поиск связи желания участвовать в исследовании по тем или иным причинам.

Например, известно то, что количество людей с тяжелой, средней и легкой степенью заболевания в популяции соотносится как 1:10:100

Представим, что в исследовании, среди «случаев», это соотношение составило 1:2:4. Это значит, что в исследовании участвует больше людей со средне-тяжелыми и тяжелыми формами заболевания, и выборка не является репрезентативной.

Слайд 23.

Следующий тип ошибки, характерный для ретроспективных исследований, включая ИСК, это предвзятость воспоминания. Это вариант предвзятости информации, при котором вероятность указания на воздействие или исход связана с воздействием или исходом.

Представим, что предпринято исследование с целью оценки гипотезы о связи между использованием гербицидов и развитием саркомы мягких тканей.

Контрольных пациентов можно взять из популяции в целом, а можно использовать пациентов с другими типами злокачественных новообразований.

Исследование, которое использовало контроль из популяции в целом, будет иметь тенденцию переоценивать риск, а исследования, использовавшие пациентов с другими типами злокачественных новообразований в качестве контроля, будут склонны недооценивать риск из-за ошибки воспоминания.

Слайд 24.

Еще пример «ошибки воспоминания». В расследовании вспышки кишечной инфекции, участники-случаи куда более детально помнят свой пищевой анамнез, поскольку они уже думали об этом, многократно вспоминали и пытались понять причину заболевания.

Ещё хуже, если «случаи» обсудили свои воспоминания между собой, пришли к общему согласию по поводу фактора передачи и попытаются навязать свою точку зрения, например, остальным пациентам в палате.

Еще один вид такой предвзятости получил название предвзятость материнского воспоминания. Логично ожидать, что матери детей родившихся с врождёнными дефектами развития гораздо чаще сообщают информацию о тех событиях, которые могли повлиять на здоровье их детей. Матери здоровых детей заинтересованы в этих воспоминаниях гораздо в меньшей степени, поэтому информация от тех и других будет значительно отличаться по качеству

Слайд 25.

Среди прочих систематических ошибок необходимо упомянуть предвзятость интервьюера, когда человек, проводящий интервью, знакомый с гипотезой исследования, может невольно оказывать влияние (провоцировать) тот или иной ответ, заводя наводящие вопросы либо подсознательно используя невербальные аспекты коммуникации.

Слайд 26.

Как можно избежать предвзятости информации?

Для этого можно использовать сокрытие (маскировку) гипотезы исследования от опрашивающих и опрашиваемых, тщательно разрабатывать опросники, использовать другие источники данных, помимо опроса.

Слайд 27.

Когортные исследования менее подвержены ошибкам в сравнении с исследованиями типа случай-контроль. Больше подвержены ошибкам ретроспективные когортные исследования ввиду возможности искажения или потери информации о воздействии фактора риска, поэтому они менее ценны, чем проспективные, хотя значительно дешевле и не такие трудоемкие.

Слайд 28.

Интересным видом ошибки является также т.н. эффект здорового работника, который наиболее актуален для изучения воздействия вредных производственных факторов. В целом ряде исследований, в которых делались попытки сравнить смертность среди рабочих, подверженных действию определенных производственных факторов риска с показателями смертности в общей популяции, были получены парадоксальные, на первый взгляд, выводы: часто оказывалось, что показатели смертности среди рабочих подверженных вредным профессиональным факторам, значительно ниже чем в популяции в целом. Причина возникновения систематической ошибки связана с тем что значительное количество членов общей популяции просто не может работать связи с различными заболеваниями в конечном итоге приводящими к смерти

Слайд 29.

Серьезным проблемой является эффект конфаундинга, возможный при всех видах исследований. Конфаундер, или вмешивающийся фактор — это переменная, которая может вызвать или предотвратить изучаемый исход, но не является промежуточной в причинной цепи. Если поправки на вмешивающийся фактор внести невозможно, то его влияние не будет отличимо от влияния изучаемого воздействия (воздействий). Таким образом конфаундинг может представлять угрозу валидности исследования тогда, когда существует какой-то третий фактор, связанный как с воздействием, так и исходом, что может приводить к искажению результатов.

Слайд 30.

Проблеме конфаундеров посвящена отдельная лекция. Позвольте лишь коротко, тезисно, перечислить способы их контроля. К ним относятся, к примеру,: рандомизация, рестрикция, подбор контролей, стратификация, регрессионный анализ, а также использование дизайна перекрестного исследования

Слайд 31.

Т.о., СисО ошибку можно минимизировать, правильно выбрав метод исследования и способ формирования выборки, стандартизировав инструменты измерения и статистического анализа, использовав методы коррекции вмешивающих факторов (конфаундеров) и пр.

Возможность ошибок во время отбора пациентов в исследование следует учитывать еще на этапе планирования, а ошибки, связанные с факторами, которые вмешиваются, необходимо учитывать при анализе данных после завершения исследования. В случае выявления систематических ошибок всегда проводят оценку их клинической значимости для результатов исследования.Слайд 18.