

МИНИСТЕРСТВО ЗДРАВООХРАНЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
САМАРСКИЙ ГОСУДАРСТВЕННЫЙ МЕДИЦИНСКИЙ  
УНИВЕРСИТЕТ



*Г.П.КОТЕЛЬНИКОВ*  
*А.С.ШПИГЕЛЬ*

**ДОКАЗАТЕЛЬНАЯ МЕДИЦИНА**  
***НАУЧНО-ОБОСНОВАННАЯ***  
***МЕДИЦИНСКАЯ ПРАКТИКА***

*Монография*

САМАРА 2000

**Доказательная медицина. Научно-обоснованная медицинская практика:**  
*Монография / Г.П. Котельников, А.С. Шпигель.* - Самара; СамГМУ, 2000. - с.

ISBN 5-8017-0047-1

Посвящена быстро развивающейся новой дисциплине - клинической эпидемиологии. Рассматриваются основные принципы этой науки и возможность использования их в практической работе для принятия решений при оценке эффективности лечебного и профилактического вмешательств. Обсуждаются требования в отношении описания структуры и представления результатов клинических испытаний. Показана возможность практической реализации доказательной медицины на примере разработки и внедрения формулярной системы для обеспечения затратно-эффективного рационального назначения и использования лекарственных средств.

Монография может представлять интерес для практических врачей всех специальностей, научных работников, преподавателей медицинских вузов, студентов и организаторов здравоохранения.

**Рецензенты:** *Павловичев С.А.* - докт. мед. наук, профессор, *Крюков Н.Н.* - докт. мед. наук, профессор.

ISBN 5-8017-0047-1

© Самарский государственный  
медицинский университет, 2000

© Г.П.Котельников, А.С.Шпигель, 2000

Редактор Т.И. Заболоцкая  
Корректор Т.И. Заболоцкая  
Компьютерная верстка О.А. Карасевой  
Дизайн обложки И.Н. Левшича

Сдан в набор 5.10.99. Подписан в печать 20.12.99.  
Формат 60x84/16. Бумага офсетная. Печать офсетная.  
Усл. печ. л. 7,0. Тираж 500 экз. Заказ .

Самарский государственный медицинский университет.  
443099, г. Самара, ул. Чапаевская, 89.

Типография ООО Научно-внедренческой фирмы  
"Сенсоры. Модули. Системы"

Лицензия ПЛД 67-33 от 03.02.97 г.  
443001, Самара, ул. Молодогвардейская, 196.

## ВВЕДЕНИЕ В НАУЧНО-ОБОСНОВАННУЮ МЕДИЦИНСКУЮ ПРАКТИКУ

### 1. Клиническая эпидемиология - новая область знаний

На рубеже 80-90-х годов, благодаря трудам группы канадских ученых - D. Sackett, В. Haynes, G. Guyatt, P. Tugwell из университета МакМастера Онтарио, в медицине сформировалась новая область знаний - *клиническая эпидемиология*, рассматривающая врачебное искусство с позиции строгих научных принципов [1].

Предметом изучения клинической эпидемиологии является разработка научных основ врачебной практики - свода правил для принятия клинических решений.

**Главнейший принцип клинической эпидемиологии - каждое клиническое решение должно основываться на строго доказанных научных фактах.** Этот постулат получил название *«evidence-based medicine»*, в буквальном переводе - *«медицина, основанная на фактах»* либо, что более точно отражает значение термина, *«научно-обоснованная медицинская практика»*, или *«научно-доказательная медицина»*.

Широкое внедрение научно-обоснованной медицинской практики во многом обусловлено экономическим аспектом. **От того, насколько убедительны научные данные в отношении клинической и стоимостной эффективности, зависит правильность размещения финансовых средств.** Даже в высокоразвитых странах ресурсы, выделяемые обществом на здравоохранение, ограничены. **Естественно, что наиболее эффективно направлять эти ресурсы на развитие методов профилактики, диагностики и лечения, практическая польза которых подтверждена исследованиями, удовлетворяющими критериям научно-обоснованной медицинской практики.**

Научно-обоснованная медицинская практика - это попытка изменить сложившуюся в течение веков систему авторитарных отношений в медицине и не ставить в основу принятия решений укорен-

нившуюся традицию или мнение авторитета. *Индивидуальный врачебный опыт и мнение даже весьма авторитетных экспертов рассматриваются как не имеющие достаточной научной основы.*

Точка зрения экспертов часто разноречива. Иногда суждение, однажды высказанное авторитетным экспертом, кочует из одного руководства в другое, несмотря на очевидные противоречащие факты. В качестве примера С.Е. Бащинский (1996) [2] приводит рекомендацию Lie K. et all. (1974) [3] по профилактическому введению *лидокаина* при остром инфаркте миокарда. В 1974 году появилась публикация, согласно которой профилактическое введение *лидокаина* в первые двое суток острого инфаркта миокарда значительно уменьшало риск фибрилляции желудочков, хотя и не влияло на смертность. На основании этих данных авторы знаменитого руководства «Болезни сердца» под редакцией Е. Браунвальда стали рекомендовать рутинное профилактическое введение *лидокаина*. Это мнение стало практически общепринятым и воспроизводилось в большинстве обзоров и монографий вплоть до начала 90-х годов. И это несмотря на то, что к 1990 году были опубликованы результаты 12 рандомизированных контролируемых исследований и практически во всех профилактическое введение *лидокаина* при инфаркте миокарда приводило к увеличению смертности [4]. Лишь в четвертом издании «Болезней сердца», вышедшем в 1992 году, авторы признали тактику рутинного введения *лидокаина* ошибочной и даже вредной.

Не отрицая полностью огромной важности личного опыта и значительного вклада исследователей в решение той или иной проблемы, научно-обоснованная медицинская практика ориентирована на следующие положения:

- в большинстве клинических ситуаций диагноз, прогноз и результаты лечения отдельного больного неопределенны и поэтому должны выражаться через вероятности;
- в клинические наблюдения заложены предвзятость и систематические ошибки, поскольку сделаны они экспериментатором-врачом на объекте наблюдения - *человеке*;
- любые исследования, включая клинические, подвержены влиянию случайности;

- клиницисты должны полагаться на наблюдения, основанные на твердых научных принципах, включающих способы уменьшения предвзятости и оценку роли случайности.

Систематическое отклонение полученных показателей от истинных значений может возникать на любой стадии исследования вследствие дефектов его планирования, организации и проведения. Такой процесс называется *смещением*, или *систематической ошибкой*. Искажение может быть обусловлено и *случайностью*, которая приводит к отклонению конечного результата в любую сторону от истинного. *Смещение* и *случайность* не исключают друг друга.

Следующий стержневой принцип научно-обоснованной медицинской практики связан с критическим анализом информации. Значимость каждого факта тем больше, чем строже научная методика исследования, в ходе которого факт получен.

«Золотым стандартом» считают рандомизированные контролируемые исследования, когда ***пациентов распределяют по группам случайным образом (рандомизация)***, при этом группы не должны различаться по параметрам, влияющим на исход заболевания.

Ежегодно публикуются результаты сотен рандомизированных контролируемых исследований, которые меняют устоявшиеся стандарты врачевания.

***Ни один новый метод лечения, профилактики или диагностики не может быть признан без обязательной тщательной проверки в ходе рандомизированных контролируемых исследований.***

Особое внимание следует обращать не только на ***бесполезность***, но и на ***отсутствие научно-обоснованных доказательств*** пользы применения различных методов лечения и профилактики. В настоящее время это положение приобретает особую актуальность, так как появилась отчетливая тенденция к преобладанию клинических исследований, финансируемых фармацевтической промышленностью и производителями различных медицинских товаров и услуг.

Так, в нашей стране довольно широкое распространение получил метод гипербарической оксигенации. При анализе 446 клинических работ по применению этого метода выяснилось, что ни в одном клиническом исследовании не проводилась рандомизация, не применялось плацебо [5]. Только в 5,4% из них была группа кон-

троля. Приятно, что к этим пяти процентам можно отнести работы самарских ученых (Ратнер Г.Л., Давыдкин Н.Ф., Котельников Г.П. и др.). Другой пример - использование низкоинтенсивного лазерного излучения при лечении ряда внутренних болезней. Только в 10% из 561 публикации был проведен сравнительный анализ с показателями, полученными в контрольной группе, в 1,8% случаев использовался плацебо-контроль. При этом остается неясным, использовалась ли хотя бы в одном исследовании рандомизация [6]. Приведенные примеры показывают *не бесполезность* этих методов, а *отсутствие научно-обоснованных доказательств* их пользы. К явлениям подобного рода можно отнести и метод иридодиагностики, и электронную «кремлевскую» таблетку, и рекомендации к широкому применению пищевых добавок, комплексных гомеопатических лекарственных средств, некоторых видов минеральной воды.

***Любого критически мыслящего человека должна насторожить тенденция к широкому распространению научно не доказанных методов лечения и диагностики.***

## **2. Оценка качества клинической информации и ее интерпретация**

В статье «Какие публикации заслуживают доверия практического врача?» В.В. Двойрин [7] подробно обсуждает методы оценки качества клинической информации и умение правильно ее интерпретировать.

Если новый метод позволяет получить высокий процент излечения больных, страдающих определенным заболеванием, до того неизлечимым, то доказать его эффективность можно путем оценки результатов лишь в одной группе больных без сопоставления с показателями в контрольной группе. Однако такие новые методы лечения, обеспечивающие очевидный клинический эффект, появляются чрезвычайно редко. По этой причине неконтролируемые клинические испытания применяются в основном при изучении новых лекарственных препаратов на этапах установления их максимальной дозы, побочных и токсических проявлений, активнос-

ти и спектра действия, а также в тех случаях, когда, по предварительным данным, ожидается резкое повышение эффективности лечения.

Как правило, новая методика лечения позволяет лишь на несколько процентов повысить клинический эффект, что затрудняет объективное его выявление. В связи с этим ориентируются на контролируемое клиническое испытание - проспективное исследование, где сопоставляемые группы получают различные виды лечения: больные контрольной группы - стандартное (*обычно лучшее, по современным представлениям*, так называемый “золотой стандарт”), а пациенты опытной группы — новое лечение.

Например, "золотым стандартом" при лечении неосложненных инфекций нижних отделов мочевыводящих путей у амбулаторных больных можно считать применение современного уросептика *норфлоксацина* (нолицина<sup>®</sup>). Данный препарат имеет высокую микробиологическую активность и способность преодолевать резистентность к традиционным препаратам (ампициллин, котримаксазол). Его отличает высокая клиническая эффективность и хорошая переносимость.

В.В. Двойрин справедливо отмечает, что одно из обязательных условий надежности исследования - *сопоставимость* сравниваемых групп. Этот важнейший принцип часто нарушается. Например, показатели, полученные при обследовании пациентов, несопоставимы, если в одну из групп включают больных, леченных новым методом, а в другую - отказавшихся от него.

Согласие пациента подвергнуться лечению новым методом зависит от типа его нервной деятельности, степени выраженности заболевания, общего состояния, возраста и других факторов, которые могут иметь прогностическое значение. Иногда сопоставляются результаты нового метода лечения, проводимого в одной клинике, с эффективностью стандартного способа лечения, используемого в других лечебных учреждениях. В этом случае неоднородность групп еще больше увеличивается за счет различий в технической оснащенности и квалификации кадров, использования разных вспомогательных лечебных методов.

Примерно такими же недостатками отличается метод так называемого «исторического контроля», когда сопоставляются результаты лечения, полученные до внедрения нового метода, с эффективностью современного лечения, т.е. сравнивают показатели специально подобранной для применения нового метода лечения группы больных со случайно взятой, неподобранной группой «прошлого». При этом не учитываются такие обстоятельства, что в последние годы стало возможным успешно лечить больных, ранее считавшихся неизлечимыми, улучшились диагностические методы, изменились уход, вспомогательная терапия и оптимальные сроки наблюдения.

Важнейшее условие, обеспечивающее надежность и достоверность контролируемого исследования, — это *однородность* групп больных по всем признакам, которые влияют на исход заболевания. Достичь возможности сопоставимости групп наблюдения в наиболее полном объеме можно только при использовании *рандомизации* - метода случайного распределения пациентов на группы. Однако такое распределение не может быть достигнуто беспорядочным отбором без предварительно составленного научно-обоснованного плана, который должен обеспечить каждому больному, включенному в исследование, равную вероятность получения альтернативного лечения.

***Истинная рандомизация предполагает обязательное соблюдение двух условий: во-первых, непредсказуемый характер распределения больных на группы (исследователь не может предугадать, в какую группу попадет следующий больной), во-вторых, исследователь, проводящий набор пациентов, не должен знать, в какие группы они попадают (“слепой отбор”).***

В последние годы, благодаря широкой пропаганде рандомизированных испытаний в зарубежной и отечественной литературе, термин “рандомизация” стал использоваться авторами публикаций без достаточных оснований, в качестве “магического заклинания” для придания исследованию научного веса и во избежание последующей критики. В результате появляются малонадежные исследования, не только компрометирующие саму идею рандомизации, но и засоряющие литературу.



По сведениям K.F. Schultz et al. [8], только 9 — 15% статей в медицинских журналах содержат адекватное описание процесса рандомизации. *Упоминание о процессе рандомизации не означает правильного ее проведения; гораздо чаще, чем можно было бы ожидать, рандомизация осуществляется неправильно.*

Существует множество методов «случайного» распределения препаратов между пациентами — подбрасывание монеты (орел/решка), вытаскивание карты из колоды (красная/черная масть), по датам рождения (четная/нечетная). Ни один из этих методов не дает адекватного принципа формирования исследуемых групп. В настоящее время единственным признанным методом рандомизации является *использование компьютерных кодов.*

Специальная компьютерная программа на основании случайно полученных чисел генерирует последовательность распределения пациентов по группам. Необходимо помнить, что несопоставимость групп наблюдения может возникнуть и после рандомизации, если не все больные будут в дальнейшем участвовать в исследовании, например, из-за ошибок диагностики, побочных эффектов лечения, серьезных проявлений сопутствующих заболеваний, отказа продолжать лечение или перехода в другую лечебную группу. В публикации должна быть предусмотрена возможность оценки таких вариантов. Один из способов свести к минимуму возможность ошибок — проведение анализа в группах, исходя из допущения, что все больные получили лечение, предписанное при рандомизации. Такой подход уменьшает вероятность ложноположительного результата, то есть выявления эффекта лечения, когда он на самом деле отсутствует. Полагают, что если более 80% из включенных в группу больных наблюдались до конца исследования, то результаты могут быть достаточно надежными.

Для повышения эффективности рандомизации применяется предварительная **стратификация** (stratum — слой, пласт). Смысл стратификационной рандомизации заключается в том, что распределение вариантов лечения осуществляется в однородных группах больных, сформированных по ведущим прогностическим признакам. Чем меньше будут отличаться между собой больные в пределах такой группы и чем больше различия между группами при пра-

вильном выборе признака стратификации, тем надежнее результаты исследования. Вывод о существенности различий в эффекте лечения или профилактики можно делать лишь при однородности сопоставляемых групп по важнейшим прогностическим факторам.

К сожалению, во многих исследованиях под «существенностью» или «несущественностью» различий в распределении таких факторов понимают лишь наличие или отсутствие статистически достоверных различий. Между тем, очевидно, что влияние прогностического фактора на результат сопоставления определяется не столько частотой его встречаемости в отдельных лечебных группах, сколько «прогностической силой».

«Сильный» в прогностическом отношении фактор может оказать большее влияние на результат сопоставления эффективности двух методов лечения (даже при статистически недостоверном различии его частоты в сравниваемых группах), чем значительно различающийся по частоте, но «слабый» в прогностическом отношении фактор.

Довольно часто полученные данные содержат множество переменных, которые по природе своей никак не связаны между собой. Если их чисто механически подвергнуть обработке для выявления взаимосвязи между переменными, то **по теории вероятности каждая двадцатая корреляция окажется статистически достоверной с уровнем статистической значимости  $p < 0,05$ !** Факт проведения множественных сравнений может быть не упомянут, а в окончательный отчет попадают случайно выявленные взаимосвязи, которым будет придано большое значение. Такое сочетание случайно обнаруженных взаимосвязей с предвзятым представлением только «интересных» данных встречается в научных статьях очень часто.

Неправильный подход к оценке статистической достоверности информации часто приводит к игнорированию клинически важных сведений и принятию во внимание клинически бесполезных данных. Значительно увеличив число наблюдений, исследователи могут доказать наличие даже столь небольших различий, которые не имеют клинического значения. Поэтому важно понимать, что **статистическая достоверность — это не синоним клинической значимости результатов исследования.**

При окончательном принятии решения о возможности использования рекомендуемого метода лечения врач должен отдавать себе отчет, насколько выводы исследования согласуются с современными биомедицинскими представлениями, а также быть уверенным, что назначает своим больным новое лечение не потому, что, по мнению исследователя, оно должно быть эффективным, а потому, что оно в действительности имеет преимущество перед традиционными методами.

При оценке результатов, полученных от применения нового метода лечения, анализу должны быть подвергнуты все виды и степень выраженности эффектов, случаи безуспешного лечения, а также разнообразные побочные явления.

**Основными критериями оценки эффективности должны быть клинически важные результаты или исходы лечения**, например, выздоровление, продолжительность и качество жизни, частота возникновения отдаленных осложнений, степень трудовой, медицинской и социальной реабилитации, продолжительность жизни без рецидивов.

*Нередко в качестве критериев оценки эффективности вмешательства используют так называемые “суррогатные” исходы.*

Под суррогатным исходом в клинических испытаниях понимают лабораторный или выявляемый при физикальном исследовании показатель, заменяющий клинически значимый результат лечения. При этом предполагается, что изменения этого показателя в ходе лечения должны отразиться и на клинически значимом исходе. Например, *при исследовании эффективности профилактики остеопороза суррогатным исходом будет изменение плотности костей, клинически значимым — снижение частоты переломов; при изучении эффекта воздействия гипохолестеринемических препаратов суррогатный исход будет заключаться в снижении уровня холестерина крови, а клинически значимый — в снижении смертности и заболеваемости инфарктом миокарда.* К сожалению, суррогатные исходы далеко не всегда коррелируют с клинически значимыми и об этом надо помнить при оценке результатов исследований.

### 3. Понятия: истинный клинический исход и косвенные критерии оценки

В окончательных клинических испытаниях в качестве основного критерия оценки должен использоваться *истинный клинический исход* - клиническое проявление, которое имеет существенное для больного значение и которого он хотел бы избежать. Например, смерть, а также потеря зрения, необходимость применения искусственной вентиляции легких и иные явления, существенно снижающие качество жизни. Исследования, в которых изучаются такого рода клинические исходы, требуют значительных материальных и временных затрат. В результате в последнее время уделяется большое внимание альтернативным способам решения проблемы или косвенным критериям оценки.

По определению R.J. Temple [9], «... **косвенный критерий оценки в клиническом испытании - лабораторный показатель или симптом, который заменяет клинически значимый исход, прямо характеризующий самочувствие больного, его функциональное состояние и выживание. Изменения косвенного критерия оценки, вызываемые лечением, должны отражать изменения клинически значимого критерия (исхода)**».

Fleming T.R., De Mets D.L. [10] аргументированно показывают возможность допущения ошибок при использовании косвенных критериев оценки в клинических испытаниях. Среди различных объяснений данного явления упоминается возможность того, что патологический процесс может влиять на клинический исход, развиваясь несколькими патогенетическими путями, часть из которых не связана с косвенным критерием; воздействие вмешательства на эти пути отличается от его влияния на косвенный критерий. Еще более вероятно то, что вмешательство может влиять на клинический исход с помощью механизмов, полностью неизвестных и непредсказуемых. На примере различных заболеваний авторы убедительно показывают, как косвенные критерии оценки привели к ошибкам относительно реального воздействия лечения на здоровье больных.

Широко распространенная ошибка заключается в том, что исход, представляющий собой результат корреляции (т.е. коррелирующий с истинным клиническим исходом), можно использовать в качестве достоверного косвенного критерия оценки (т.е. заменять им истинный клинический исход). Однако такая замена обоснована только в том случае, если воздействие изучаемого вмешательства на косвенный критерий позволяет предсказать его влияние на клинический исход, а это более жесткое требование, чем наличие корреляции.

В качестве примеров косвенных критериев оценки можно назвать подавление желудочковых аритмий, снижение уровня холестерина или артериального давления в кардиологических исследованиях, уменьшение размеров опухоли в исследованиях, посвященных лечению рака. *Косвенные критерии оценки очень редко, а то и вовсе не отражают важные клинические исходы в клинических испытаниях.*

Классическим примером ненадежности косвенных критериев оценки является использование такого показателя, как снижение частоты желудочковой экстрасистолии вместо снижения смертности от сердечно-сосудистой патологии. Наличие желудочковой аритмии почти в четыре раза повышает риск смерти от сердечно-сосудистых осложнений, в частности, риск внезапной смерти [11, 12]. Естественно, было высказано предположение, что подавление желудочковых аритмий после инфаркта миокарда уменьшит уровень смертности. Три новых препарата - *энкаинид*, *флекаинид*, *мортицизин* - оказались эффективны для восстановления ритма, поэтому их широко применяли при лечении больных с тяжелыми или опасными для жизни нарушениями сердечного ритма. Эти лекарственные средства принимали более 200 тыс. человек в год. Затем в США были проведены работы [13] по изучению влияния аритмий на уровень смертности. В исследовании оценивалось, как эти три препарата влияют на выживаемость у больных с инфарктом миокарда при наличии как минимум 10 желудочковых экстрасистол в минуту. Результаты оказались ошеломляющими. Испытание *энкаинида* и *флекаинида* было прекращено на ранних этапах после того, как у больных, принимавших эти препараты, было зарегист-

рировано 33 случая внезапной смерти, в то время как в группах, получавших плацебо, таких случаев было отмечено 9. Всего умерло 56 больных, получавших энкаинид и флекаинид, и 22 пациента, принимавших плацебо. После получения окончательных данных оказалось, что число случаев внезапной смерти составляло соответственно 43 и 16, а общее число летальных случаев - соответственно 63 и 26. Последующие результаты свидетельствовали об увеличении уровня смертности и у больных, получавших *морицизин* [13].

Следующий пример относится к профилактике возникновения аритмий. Для поддержания синусового ритма после кардиоверсии у больных с мерцанием предсердий использовался *хинидин* [14]. Мета-анализ результатов 6 испытаний показал, что прием хинидина способствовал сохранению синусового ритма в течение 1 года у 50% больных по сравнению с 25% пациентов, не получавших его. Однако уровень смертности на фоне приема хинидина увеличился с 0,8 до 2,9%! Предотвращение рецидивов мерцательной аритмии - ценное качество препарата, но оно явно проигрывает на фоне повышения уровня смертности. Подобный недостаток воздействия был выявлен и у *лидокаина*, мета-анализ показал, что **уменьшение риска возникновения желудочковой тахикардии на одну треть сопровождается повышением риска смерти в такой же пропорции** [15, 16].

Одним из первых инотропных препаратов, предложенных для лечения застойной сердечной недостаточности, был *милринон*. Выполненные исследования показали, что он повышал сердечный выброс и толерантность к физической нагрузке. *Милринон* является инотропным средством (как и *дигоксин*), увеличивает силу сокращения сердечной мышцы. После проведения рандомизированного двойного слепого плацебо-контролируемого исследования было продемонстрировано увеличение уровня общей смертности у больных, принимавших милринон, по сравнению с получавшими плацебо соответственно 29 и 23% ( $p = 0,04$ ) [17].

Таким образом, хотя повышение сердечного выброса, фракции изгнания и толерантности к физической нагрузке коррелирует с более длительным выживанием пациентов при застойной сердечной недостаточности, улучшение этих показателей не является на-

дежным прогностическим критерием влияния лечения на уровень смертности. Конечно, повышение толерантности к физической нагрузке само по себе является важным клиническим исходом и могло бы рассматриваться как желательный эффект. Однако вряд ли врач или непосредственно сам больной смогут принять верное решение, зная о воздействии лечения только на симптоматику и не зная его влияния на смертность. Вместе с тем, следует отметить, что воздействие препарата на уровень смертности можно оценить только при длительных исследованиях.

Еще один классический пример - снижение уровня липидов. Давно известно, что концентрация липидов в крови (общего холестерина, его фракций и триглицеридов) позволяет достоверно прогнозировать уровень смертности от сердечно-сосудистых заболеваний. Однако некоторые исследователи оспаривают существование взаимосвязи между уменьшением содержания липидов и снижением уровня общей смертности [17-19]. С семидесятых годов для снижения уровня холестерина применяются такие препараты как *клофибрат* и *ниацин*. Однако, как показало крупное семилетнее исследование [20], ни один из них не уменьшил общий уровень смертности.

Наряду с этим, недавнее клиническое испытание, проведенное в Скандинавии, показало, что у больных со стенокардией или инфарктом миокарда в анамнезе *симвастатин* снижает уровень холестерина на 25%, а общую смертность - на 30% [21]. Это первое крупное испытание, которое продемонстрировало положительное влияние гиполипидемического препарата на общую смертность. Независимо от того, связано ли достигнутое с помощью симвастина снижение уровня смертности с уменьшением концентрации холестерина в крови, оценка эффективности лечения только на основании этого показателя явно недостаточна.

Без учета таких клинических критериев оценки, как общая смертность, *фибраты* могли бы широко применяться благодаря их гиполипидемическому действию.

***Одним из факторов риска смерти от сердечно-сосудистых заболеваний*** является артериальная гипертония. Снижение артериального давления на каждый миллиметр ртутного столба приводит к

снижению уровня смертности на 5% и частоты возникновения инсультов на 10%. В одном из крупных клинических исследований по лечению гипертонии [22] выявлено снижение общей смертности при мягкой гипертонии на фоне ступенчатого лечения, начинающегося с диуретиков. Результаты недавнего исследования [23] подтвердили положительное воздействие гипотензивной терапии в виде назначения низких доз диуретиков на выживаемость, частоту несмертельного инфаркта миокарда и инсульта.

Анализ современных подходов к лечению гипертонии [24] показывает, что около 50% всех гипотензивных средств занимают два их класса: ингибиторы ангиотензин-превращающего фермента и антагонисты кальция. Эти препараты по своей стоимости более чем в 30 раз дороже диуретиков. Однако ни в одном из проведенных рандомизированных испытаний не выяснилось, уменьшают ли лекарственные средства какого-либо из этих двух классов смертность от сердечно-сосудистых заболеваний или заболеваемость при первичном профилактическом лечении больных с артериальной гипертонией. Более того, недавно проведенное популяционное исследование типа «случай-контроль» позволило высказать предположение о том, что применение антагонистов кальция может повышать риск возникновения инфаркта миокарда при артериальной гипертонии [25]. Проводились рандомизированные испытания этих препаратов при застойной сердечной недостаточности и инфаркта миокарда. Мета-анализ влияния антагонистов кальция на выживаемость в этих испытаниях выявил возможное неблагоприятное воздействие данных лекарственных средств [26]. Выказано предположение [24], что использование антагонистов кальция, *клинически полностью не исследованных, для лечения артериальной гипертонии*, может стоить только в США более 2,5 млрд долларов в год.

Хотя влияние на артериальное давление кажется надежным косвенным критерием оценки эффективности применения низких доз диуретиков, его использование может привести к ошибочным выводам при изучении новых гипотензивных средств. Благоприятное гипотензивное влияние антагонистов кальция реализуется и через такие механизмы действия, которые пока неизвестны и непредсказуемы.



Весьма показательный пример - развитие *остеопороза* в менопаузе. В этот период у женщин снижается масса костной ткани и развивается остеопороз, что неизбежно повышает риск переломов бедренных и других костей [27, 28]. Для увеличения массы костной ткани и уменьшения частоты переломов достаточно часто назначают *эстрогены* или *препараты кальция*. Широко используется *фторид натрия*, стимулирующий образование костной ткани, хотя его применение с этой целью и не одобрено.

Проведение плацебо-контролируемого рандомизированного испытания, в которое были включены 202 женщины с остеопорозом и переломами позвонков, при длительности наблюдения 4 года показало, что лечение повышало минеральную плотность костной ткани в поясничном отделе позвоночника на 35% ( $p < 0,001$ ). Однако в группе, получавшей препарат фтора, чаще, чем в группе, принимавшей плацебо, отмечались новые переломы позвонков (соответственно 163 и 136 переломов) и переломы иной локализации (соответственно 72 и 24 перелома;  $p = 0,01$ ). В.L. Riggs et al. [27] сделали заключение о том, что *лечение препаратом фтора повышает минеральную плотность, но кости при этом становятся более ломкими!*

Есть и другие примеры неудачного использования косвенных критериев оценки - испытания вакцин, когда в качестве косвенного критерия учитывают наличие нейтрализующих антител или клеточный иммунный ответ [29].

Тщательное подтверждение достоверности косвенного критерия оценки осуществляется редко. В окончательных клинических испытаниях по изучению влияния вмешательства на клинические исходы частота ложноположительных и ложноотрицательных результатов должна быть низкой, обычно от 2,5 до 10%.

Следовательно, чтобы достоверно замещать собой клинический исход, косвенный критерий оценки должен обеспечивать высокий уровень точности при прогнозировании влияния вмешательства на истинный клинический исход. Прогнозы с точностью до 50% - то же самое, что прогнозы с помощью подбрасывания монетки.

Методы определения достоверности косвенных критериев оценки обсуждаются в работах L.S. Freedman et al. [30]. Для подтвержде-

дения обоснованности использования косвенных критериев требуется глубокое понимание причинно-следственных путей развития заболевания, ожидаемых и непредвиденных механизмов воздействия вмешательства. Однако практически эти требования трудновыполнимы.

Косвенные критерии следует использовать там, где они лучше себя зарекомендовали - при оценке новых методов лечения в предварительных испытаниях, когда изучается биологическая активность лекарственных средств. Результаты таких испытаний, в свою очередь, помогут решить, следует ли проводить более крупные и длительные клинические испытания. Информация о косвенных критериях оценки в окончательных испытаниях важна для понимания механизмов действия вмешательства, однако **главной целью исследований должно быть получение прямых доказательств безопасности вмешательства и его влияния на истинные клинические исходы.**

#### **4. Современные методические требования для авторов публикаций, посвященных оценке эффективности вмешательства**

Озабоченность мировой научной общественности качеством представления данных рандомизированных контролируемых исследований выразилась в принятии рядом ведущих медицинских журналов **единых методических требований для авторов публикаций**. С января 1997 г. журналы JAMA, Lancet и BMJ обязали авторов следовать этим рекомендациям. Ожидается, что все крупные журналы в скором времени присоединятся к ним. В России редакция “Международного журнала медицинской практики” полностью опубликовала текст этих рекомендаций и призвала отечественных ученых-медиков использовать их в своих научных отчетах [31].

Одно из важнейших условий объективности данных - *воспроизводимость*. Клиническое исследование трудно отнести к «чистому эксперименту». В них всегда заложена опасность систематических ошибок (предвзятости), избежать которых можно лишь следуя четким научным принципам.

Авторы должны представить основные результаты исследования таким образом, чтобы была возможность оценить эффектив-

ность вмешательства не только качественно, но и количественно. Это позволит читателям определить, достаточна ли клиническая значимость эффекта (не путать со статистической значимостью!), как соотносится полезный эффект с возможным риском и побочными действиями вмешательства, целесообразно ли новое вмешательство с экономической точки зрения. Эти соображения и должны в итоге быть основой для принятия решений о внедрении результатов исследования в практику здравоохранения.

### **Параметры для представления эффекта вмешательства**

Выделяют 3 основных параметра для представления эффекта вмешательства:

- **относительное снижение частоты неблагоприятных исходов;**
- **абсолютное снижение частоты неблагоприятных исходов;**
- **число больных, которых необходимо лечить в течение определенного времени, чтобы предотвратить один неблагоприятный исход.**

При любом способе представления данных недостаточно указать на наличие статистически достоверного различия по критерию  $p$ ; необходимо рассчитать 95% доверительный интервал (ДИ), в пределах которого может колебаться величина эффекта. ДИ - диапазон колебаний истинных значений в популяции. Величины, полученные в исследованиях на выборке больных, отличаются от истинных величин в популяции вследствие влияния случайности. Так, 95% ДИ означает, что истинное значение величины с вероятностью в 95% лежит в пределах рассчитанного интервала. Доверительные интервалы помогают сориентироваться, соответствует ли данный диапазон значений представлениям читателя о клинической значимости эффекта и каких результатов можно ожидать, применив описанную методику на сходной группе больных.

Формула расчета ДИ зависит от типа данных, для которых он рассчитывается. Методики расчета ДИ для средней величины, а также для пропорции, разницы средних величин и разницы двух пропорций, отношения двух пропорций, отношения рисков и шансов между собой отличаются.

**Исход** - клинически значимое явление (выздоровление, смерть, осложнения) у пациентов в опытной и контрольной группах, лабораторный показатель или признак, который служит объектом интереса исследователя. При проведении клинических испытаний исходы служат критериями оценки эффективности лечебного или профилактического воздействия, информативности диагностического метода.

Таблицу, в которой приводятся возможные исходы, исследования по изучению эффективности лечебного воздействия называют *таблицей сопряженности* (табл. 1). С помощью этой таблицы рассчитывают рекомендованные ключевые показатели.

Таблица 1

Г р у п п а	Изучаемый эффект (исход)		
	Есть	Нет	Всего
Изучаемая (I)	A	B	A + B
Контрольная (II)	C	D	C + D

Частота исходов в группе лечения (A/A+B) — ЧИЛ.

Частота исходов в контрольной группе (C/C +D) — ЧИК.

***Случаи, когда исследуемый препарат или метод лечения снижает вероятность развития неблагоприятного исхода***

**СОР (снижение относительного риска)** — относительное уменьшение частоты неблагоприятных исходов в группе лечения по сравнению с контрольной группой, рассчитываемое как  $|\text{ЧИЛ} - \text{ЧИК}| / \text{ЧИК}$ ; приводится вместе с 95% ДИ.

**САР (снижение абсолютного риска)** — абсолютная арифметическая разница в частоте неблагоприятных исходов между группами лечения и контроля. Рассчитывается как  $|\text{ЧИЛ} - \text{ЧИК}|$ .

**ЧБНЛ** — число больных, которых необходимо лечить определенным методом в течение определенного времени, чтобы предотвратить неблагоприятный исход у одного больного. Рассчитывается как  $1/\text{САР}$  и приводится вместе с 95% ДИ.

*Случаи, когда исследуемый препарат (метод лечения) повышает  
вероятность развития благоприятного исхода*

**ПОП (повышение относительной пользы)** — относительное увеличение частоты благоприятных исходов в группе лечения по сравнению с контрольной группой, рассчитываемое как  $|\text{ЧИЛ} - \text{ЧИК}| / \text{ЧИК}$ ; приводится вместе с 95% ДИ.

**ПАП (повышение абсолютной пользы)** — абсолютная арифметическая разница в частоте благоприятных исходов между группами лечения и контроля. Рассчитывается как  $|\text{ЧИЛ} - \text{ЧИК}|$ .

**ЧБНЛ** — число больных, которых необходимо лечить определенным методом в течение определенного времени, чтобы достичь определенного *благоприятного* эффекта у одного больного. Рассчитывается как  $1/\text{ПАП}$  и приводится вместе с 95% ДИ.

*Случаи, когда исследуемый препарат (метод лечения) повышает  
вероятность развития неблагоприятного исхода*

**ПОР (повышение относительного риска)** — относительное увеличение частоты *неблагоприятных* исходов в группе лечения по сравнению с контрольной группой. Рассчитывается как  $|\text{ЧИЛ} - \text{ЧИК}| / \text{ЧИК}$ ; приводится вместе с 95% ДИ (ПОР также используется для оценки влияния факторов риска на развитие заболевания.)

**ПАР (повышение абсолютного риска)** — абсолютная арифметическая разница в частоте неблагоприятных исходов между группами вмешательства и контроля. Рассчитывается как  $|\text{ЧИЛ} - \text{ЧИК}|$ . ПАР также используется для оценки влияния факторов риска на развитие заболевания.

**ИПВ (индекс потенциального вреда)** — число больных, которые должны получить экспериментальное лечение, чтобы у одного дополнительного пациента развился неблагоприятный исход, — по сравнению с больными из контрольной группы. Рассчитывается как  $1/\text{ПАР}$  и приводится вместе с 95% ДИ.

Чаще всего в исследованиях идет речь о сравнении результатов двух (или более) разных воздействий или результатов воздействия и отсутствия такового (плацебо). Плацебо-контролируемые иссле-

дования оправданы в случаях, когда не существует альтернативного способа лечения, эффективность которого не вызывает сомнений. Так, например, недопустимо оценивать эффект новых антикоагулянтов при остром инфаркте миокарда по сравнению с плацебо, поскольку убедительно доказана полезность аспирина.

Считаем целесообразным привести в качестве примера результаты совместных с Н.В. Вакуровой исследований о возможности лечения и профилактики силикоза гомеопатическими лекарственными средствами [32].

Эффект лечения оценен у 89 лиц. Клинические исследования являлись рандомизированными, проспективными, контролируемыми. Процедура рандомизации была разработана до начала исследования. Перед проведением рандомизации больные были стратифицированы по параметрам, включая профмаршрут (работа на одном предприятии), санитарно-гигиенические условия труда.

Наблюдения проводились после окончания первоначального и двух последующих курсов лечения в стационаре (через 1 и 2 года). В качестве контрольных лекарственных средств (ЛС) использовали плацебо или общепринятое лечение силикоза. Эффективность лечения определялась числом исследуемых, у которых произошло улучшение клинико-функциональных характеристик.

Показана эффективность гомеопатических препаратов: улучшились субъективные и объективные клинические характеристики, отмечена положительная динамика показателей функции внешнего дыхания, бронхофиброскопии и гистологического анализа слизистой бронхов.

Повышение относительной пользы (ПОП) - относительное увеличение частоты благоприятных исходов в группе лечения по сравнению с контрольной составил 93%. Общепринято, что значения этого показателя, превышающие 50%, всегда соответствуют клинически значимому эффекту.

Повышение абсолютной пользы (ПАП) — абсолютная арифметическая разница в частоте благоприятных исходов между группами лечения и контроля составила 25%.

Число больных, которых необходимо лечить изучаемым методом, чтобы достичь определенного благоприятного эффекта (ЧБНЛ = 1/ПАП) у одного дополнительного больного, равно 4.

Помимо лечебного изучалось профилактическое действие гомеопатической фармакотерапии. Протективный эффект обосновывался возможностью влияния изучаемых лекарственных средств на естественное течение заболевания. В двух группах обследуемых, с подозрением на силикоз, требующим динамического наблюдения, одна из которых получала гомеопатические препараты ( $n = 32$ ), а другая лекарственных средств не получала ( $n = 30$ ) производился расчет неблагоприятных исходов (случаев постановки диагноза при повторном обследовании).

Рассчитывался такой показатель, как снижение относительного риска (СОР) — относительное уменьшение частоты неблагоприятных исходов в группе лечения по сравнению с контрольной группой. Он характеризовался отношением разницы между частотами исходов в группах, получавших профилактическое лечение и не получавших, к частоте в группе, не получавших препараты.

$СОР = [ЧИЛ - ЧИК] / ЧИК = 0,125 - 0,27 / 0,27 = 0,54$  или 54%.

САР (снижение абсолютного риска) — абсолютная арифметическая разница в частоте неблагоприятных исходов между группами изучаемой и контроля. Она рассчитывалась как  $[ЧИЛ - ЧИК]$  и равнялась 0,143 или 14,3%.

ЧБНЛ — число больных, которых необходимо лечить гомеопатическим методом, чтобы предотвратить неблагоприятный исход у одного больного. Рассчитывался как  $1/САР$  и составлял 7. То есть, число лиц, которым необходимо провести протективное назначение гомеопатических лекарственных средств, чтобы получить эффект у одного дополнительного пациента, равно 7.

В другой работе (совместно с О.Б. Журовой) проводилась оценка эффективности лечения и профилактики гомеопатическими лекарственными средствами бронхиальной астмы у детей в процессе проспективного, рандомизированного, сравнительного с "золотым стандартом", контролируемого клинического исследования на основе изучения изменений субъективных и объективных клинических признаков, показателей функции внешнего дыхания и эозинофильно-катионного протеина.

Эффективность лечения анализировалась в период обострения бронхиальной астмы.

Обследуемые (135 детей) рандомизированно были разделены на три группы. Родители были информированы о возможности приема плацебо и лечения их детей гомеопатическими лекарственными средствами и дали на это согласие.

В первой группе (46 детей) больным назначали общепринятое лечение (*фликсотид с вентолином* ингаляционным способом плюс плацебо - капли водно-спиртового раствора и гранулы). Пациенты второй группы (45 детей) получали гомеопатические лекарственные средства в виде капель и гранул плюс ингаляции плацебо. Больным третьей группы (44 ребенка) назначали только плацебо (ингаляции, водно-спиртовой раствор и гранулы).

При оценке метода лечения анализу подвергались все виды и степень выраженности эффектов, случаи безуспешного лечения, а также разнообразные побочные явления. В качестве основного критерия оценки эффективности или истинного исхода лечения, по результатам которого составлялась таблица сопряженности, выбран параметр "*приступы удушья*".

Для сравнительной оценки эффективности различных видов лечения рассчитывался такой показатель, как ЧБНЛ.

Эффективность гомеопатического метода лечения по сравнению с плацебо представлена следующими результатами: ЧИЛ =  $18/45 = 40\%$ ; ЧИК =  $3/44 = 6\%$ ; ПАП =  $40\% - 6\% = 34\%$ ; ЧБНЛ =  $1/34 = 3$ .

При сравнении эффективности общепринятого лечения с плацебо получены следующие показатели: ЧИЛ =  $44/46 = 96\%$ ; ЧИК =  $3/44 = 6\%$ ; ПАП =  $96\% - 6\% = 90\%$ ; ЧБНЛ =  $1/90 = 1$ .

Иными словами, гомеопатический метод в период обострения бронхиальной астмы по сравнению с общепринятым "золотым стандартом" (*фликсотид плюс вентолин*) не достаточно эффективен у больных со среднетяжелым и тяжелым течением заболевания. Таким образом, для лечения и профилактики силикоза использование комплексных гомеопатических лекарственных средств оправдано, а для лечения бронхиальной астмы в период обострения у детей - ограничено.



## 5. Методы и критерии отбора информации

Умение извлечь из многочисленных научных источников полезную для практики информацию формирует врача, обладающего серьезным профессиональным преимуществом перед коллегами. *Критический анализ информации - важнейший принцип научно-обоснованной медицинской практики.*

Существуют различные методы отбора из потока медицинской литературы информации в соответствии с принципам доказательной, научно-обоснованной медицинской практики [33]. Критерии отбора зависят от их тематики.

- ***Критерии отбора статей, посвященных лечению или профилактическому вмешательству:***

- рандомизированное включение пациентов в группы сравнения;
- прослеживание на протяжении всего периода исследования, по крайней мере, 80% первоначально включенных пациентов;
- использование конечных точек (исходов), имеющих важное клиническое значение;

- ***Критерии отбора статей о диагностических методах:***

- четкое определение групп сравнения, минимум одна из которых не должна иметь рассматриваемого заболевания;
- существование объективного («золотого») стандартного теста или другого общепринятого в клинической практике стандарта, с результатами которого сопоставляются данные изучаемого метода;
- отсутствие у исследователей, интерпретирующих результаты тестов, информации о данных, полученных с помощью других диагностических методов;
- адекватность анализа результатов плану исследования.

- ***Критерии отбора статей по изучению прогноза заболевания:***

- организация исходной когорты, состоящей из людей, у которых в момент включения в нее отсутствуют признаки заболевания, служащие конечными точками (исходами) исследования;

- прослеживание, по крайней мере, 80% первоначально включенных пациентов на протяжении всего периода исследования (до запланированных конечных точек наблюдения - исходов);
- адекватность анализа результатов плану исследования.

- **Критерии отбора статей по изучению причинно-следственных связей, этиологии и факторов риска:**

- четкое определение групп риска и контроля, подобранных для сравнения;
- оценка слепым методом исходов и факторов риска для каждой из групп.

- **Критерии отбора обзоров литературы:**

- четко сформулированная цель обзора;
- подробное описание источников и методов поиска литературы;
- четко сформулированные критерии отбора статей для обзора.

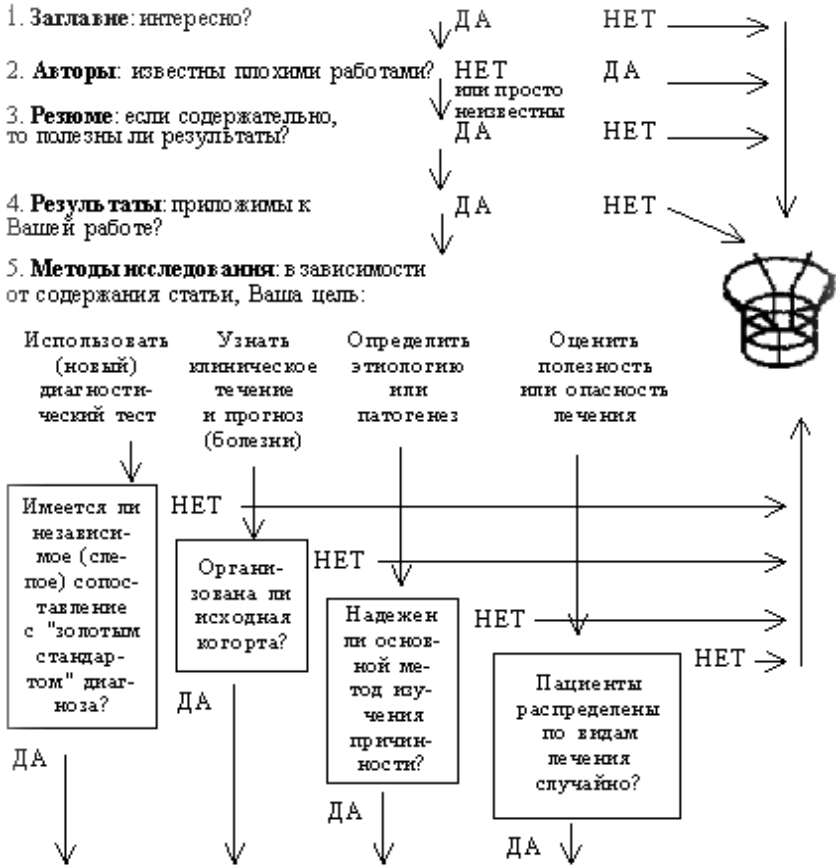
Становится общепринятым, что для современного врача умение критической оценки результатов научных исследований столь же важны и необходимы, как, например, умение аускультировать больного.

В связи с этим считаем целесообразным подробно привести положения и рекомендации по внедрению доказательной медицины, предлагаемые В.В. Власовым в серии замечательных статей, опубликованных в «Международном журнале медицинской практики» [33-36]. Одна из возможных стратегий для поиска нужных материалов иллюстрируется на рис. 1.

В первую очередь и наиболее критично автор предлагает читать раздел «Методы исследования», чтобы сразу исключить статьи, не отвечающие стандартам качества. Получив утвердительный ответ, можно приступить к чтению. Отрицательные ответы означают, что статья не соответствует основному критерию качества и ее незачем читать.

Научные исследования можно разделить на две категории: одни проводятся для выдвижения гипотез, другие - для их проверки. **Для проверки гипотез подходят только рандомизированные контролируемые исследования.** Следует различать *пилотное* исследование,

выдвигающее гипотезы, от исследования, претендующего на окончательное решение проблемы (но на самом деле не являющегося таковым!).



6. Особое внимание следует обращать на раздел "Методы исследования".

Рис.1. Схема возможной стратегии для поиска необходимых статей

Информацию, используемую для принятия клинических решений, можно разделить на *первичную* (данные оригинальных иссле-

дований, опубликованные в *рецензируемых* научных журналах) и *вторичную* (обзорные и редакционные статьи, учебники, мнения экспертов), *прямую* (полученную в ходе *клинических работ*) и *косвенную* (полученную в *эксперименте*), *сильную и слабую* (в зависимости от дизайна - способа или схемы проведения исследования). В качестве основы для принятия клинических решений научно-обоснованная медицинская практика отдает *приоритет первичной, прямой и сильной информации*.

При критической оценке публикации очень важно определить, в какой степени ее результаты можно распространить на конкретную клиническую ситуацию (так называемая «внешняя обоснованность», или «обобщаемость» исследования). Нередко читатели склонны результаты экспериментальных исследований, полученных на *животных*, переносить на *человека*, а результаты обследования *узко* ограниченной группы больных - на пациентов, отличающихся по возрасту или степени тяжести заболевания или же распространять данные об эффективности *конкретного* препарата на  *всю*  лекарственную группу.

Не менее важная задача состоит в том, чтобы понять, сформулирована ли гипотеза *до начала* исследования или на основании анализа полученных результатов.

Авторы часто склонны переоценивать результаты своей работы. Надо уметь критически оценить степень научной доказательности и решать, соответствует ли дизайн исследования поставленным задачам (так называемая «внутренняя обоснованность»). «Внутренняя обоснованность» во многом зависит от соответствия дизайна исследований поставленным задачам.

## **6. Взаимосвязь дизайна и структуры клинических испытаний**

Степень научной доказательности во многом определяется соответствием дизайна и структуры исследований.

**Дизайн** – это схема, шаблон, общий план исследования, его организационный каркас. Ни один из видов дизайна не имеет преимуществ перед другими. Все зависит от задач конкретного исследования. Правильный выбор дизайна определяет успех испытания.

Для исследований с участием двух групп пациентов наиболее распространены *параллельный* (рис. 2) и *перекрестный* (рис. 3, 4) дизайны.



Рис. 2. Параллельный дизайн

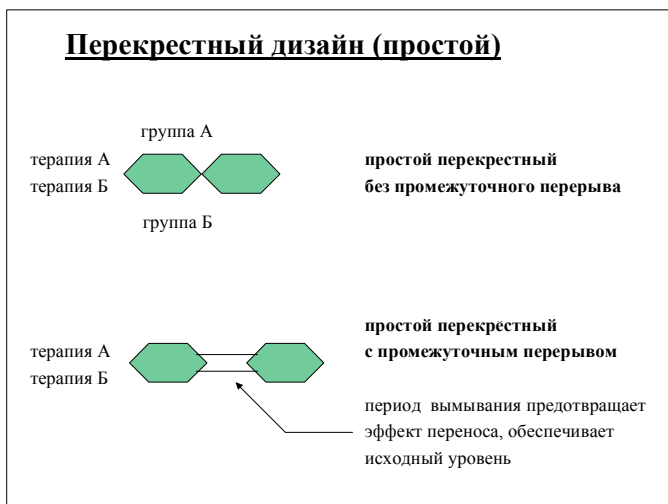


Рис. 3. Простой перекрестный дизайн

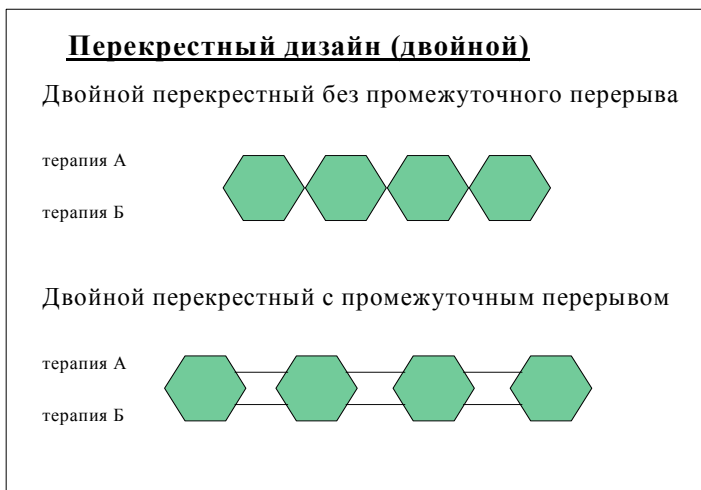


Рис. 4. Двойной перекрестный дизайн

В параллельном исследовании одной группе пациентов от начала и до конца исследования назначают одну терапию, а второй группе – другую. При перекрестном исследовании каждой группе прописывают в разное время оба вида терапии, например, первая группа пациентов вначале получает препарат А, затем препарат Б, а вторая группа – вначале препарат Б, затем препарат А.

*Преимущества параллельного дизайна перед перекрестным:*

1. Параллельное исследование можно закончить быстрее, так как в каждой группе будет всего один период лечения.
2. Качество данных параллельного исследования более «устойчиво» к нарушениям протокола, например, к пропускам пациентами визитов, их выбыванию из испытания.
3. Перекрестный дизайн может быть использован только при обследовании пациентов со стабильным длительным течением заболевания, так как их состояние должно быть примерно одинаковым перед началом проведения обоих периодов лечения.
4. При параллельном исследовании отсутствует влияние предшествующей терапии, когда лечение первым препаратом влияет на результаты лечения вторым.

*Недостатки параллельного дизайна по сравнению с перекрестным:*

1. Для проведения параллельного исследования требуется большее число пациентов.
2. Параллельный дизайн предполагает большую вариабельность данных, так как разные пациенты получают различную терапию.

Для уменьшения влияния предшествующей терапии в перекрестных исследованиях между периодами лечения различными препаратами обычно существует отмывочный период, в течение которого пациенты не получают никакой терапии. Их состояние приближается к исходному. Кроме того, удается избежать возможных лекарственных взаимодействий между исследуемыми препаратами. Иногда отмывочный период бывает в начале исследования, чтобы максимально уменьшить эффекты предыдущего лечения, и реже, в конце исследования.

Отмывочный период в конце исследования вводится для того, чтобы оценить состояние пациента через некоторое время после окончания приема препарата или избежать взаимодействия между ним и обычным последующим лечением.

Кроме того, существует понятие «*вводный период*». В течение вводного периода в начале клинического испытания пациенты могут:

- 1) не получать никакого лечения (в этом случае понятие «вводный период» соответствует понятию «отмывочный период»);
- 2) находиться на диете;
- 3) принимать плацебо (в конце вводного периода исследователь проверяет правильность и аккуратность приема препарата пациентом и на основании этого делает вывод, подходит ли он для продолжения участия в исследовании);
- 4) получать активную терапию (при завершении вводного периода в зависимости от результатов терапии, например, принимается решение о рандомизации пациента в одну из трех групп, одна из которых должна принимать ту же дозу исследуемого препарата, другая – более высокую либо комбинацию с другим препаратом, а третья – более низкую).

Отмывочные и вводные периоды, особенно при перекрестном дизайне, значительно удлиняют сроки проведения исследования, но они необходимы. В современной борьбе за рынок быстрота выполнения параллельных исследований является большим преимуществом. Поэтому *большинство проводимых в настоящее время клинических испытаний имеет параллельный дизайн*. Иногда в одном исследовании могут встречаться самые разнообразные комбинации параллельного и перекрестного дизайнов. Кроме того, существует множество других дизайнов, но они встречаются относительно редко и мы не будем на них останавливаться.

Важно различать *несравнительные* и *сравнительные* исследования. В несравнительных клинических испытаниях результаты исследуемого лечения ни с чем не сравнивают. При этом используют либо методы описательной статистики, констатирующие наблюдения (например, «в конце лечения исследуемым препаратом нормализация АД произошла у X пациентов из общего числа Y, включенных в исследование, что составляет Z% от Y»), либо анализируют динамику определенного критерия в одной группе пациентов (например, «в начале лечения исследуемым препаратом среднее значение диастолического АД составляло X мм рт ст., в конце лечения – Y мм рт ст.; снижение АД было достоверным с вероятностью  $p < 0,05$ ). В сравнительных клинических испытаниях эффект действия препарата сопоставляют с *другим активным веществом* или *плацебо*.

*Контролируемые исследования*. В широком смысле это исследования, проводимые в точном соответствии с тщательно спланированным протоколом и под контролем *монитора*, этического комитета и официальных инстанций. Большее значение имеет другое, «узкое» понимание, в соответствии с которым исследование называют контролируемым, когда изучаемый препарат сравнивают с контролем (лечением с уже известной эффективностью и переносимостью).

Например, сравнение влияния двух, ранее не исследованных доз нового препарата на пациентов в двух параллельных группах, нельзя отнести к контролируемым исследованиям, так как неизвестны эффективность и переносимость обоих методов лечения, а



сравнение воздействия более высокой дозы препарата с дозой, влияние которой уже хорошо исследовано в этой группе пациентов, можно.

Кроме того, бывают *плацебоконтролируемые* исследования одной, двух групп пациентов и более.

В исследовании с одной группой пациентов все участники получают одну и ту же терапию. Если установлены критерии для изменения терапии при определенных условиях, то в конце обследования одна группа пациентов может оказаться разделенной на несколько подгрупп.

Различные варианты организации исследований в зависимости от постановки научно-клинических задач. Совокупность их особенностей принято называть *структурой исследования*.

Для каждой структуры сформированы требования к подготовке исследования, специфические стандарты качества, стандарты обработки и представления данных. *От структуры зависит главное - доказательность результатов*. Обусловлено это тем, что классическое научное исследование предполагает устранение влияния всех случайных (посторонних) факторов, необходимость тождественности всех возможных особенностей обследуемых, кроме одного - изучаемого фактора. Только в этом случае можно предположить, что разница между обследованными людьми вызвана именно этим фактором.

В XX веке классическая модель исследования - «изучение единственной разницы» - перестала быть единственной, так как появились новые методы анализа данных. Первым и главным методом был факторный анализ, созданный Р. Фишером. Сегодня в большинстве биомедицинских исследований рассматривается вопрос: не может ли связь явления В с особенностью А объясняться влиянием постороннего фактора С?

При планировании исследования заранее учитывают возможные посторонние факторы и затем при анализе данных «вычитают» их влияние. Это позволяет ответить на вопрос, является ли самостоятельной связь В с А или она полностью может объясняться действием фактора С.

**В современной биологической и медицинской литературе встречается множество разнообразных структур исследований.**

**Основные варианты структуры исследования [33, 37-41]:**

Поперечные.

Описания спектра проявлений болезни.

Диагноза/стадии болезни.

Нормальных вариаций.

Тяжести болезни.

Болезненных процессов.

Аналитические.

Наблюдательные.

Описания случаев.

Продольные.

Перспективные.

Естественного развития, прогноза.

Причинных факторов и инцидента.

Неконтролируемого вмешательства.

С преднамеренным вмешательством.

Параллельные.

Последовательные.

Перекрестные.

С самоконтролем.

С внешним контролем.

Ретроспективные.

Наблюдательные.

*Поперечными, или одномоментными (cross-sectional)* называют исследования, в которых обследование каждого пациента выполняется *однократно*. В итоге описывается картина болезни у одного больного или группы пациентов, уточняются симптоматика, отношение отдельных симптомов с основным диагнозом и тяжестью болезни. Истинно поперечные исследования встречаются редко, поскольку врачи стремятся в пределах возможного построить временные ряды явлений, стадий заболевания. В прошлом наблюдение за развитием болезней было основным методом их изучения. Этот метод позволял получить верные представления главным образом о скоротечных заболеваниях. В настоящее время основным

объектом медицинских исследований становятся хронические, длительно протекающие болезни. Поэтому простое наблюдение может привести к ошибке. В связи с этим ориентируются на продольные исследования.

*Продольными (longitudinal, лонгитудинальными)* называют исследования с выделением группы людей, за которыми в течение некоторого времени наблюдают и *повторно* (хотя бы один раз) оценивают их состояние. Наблюдение не обязательно должно быть длительным. Для анализа естественного развития острого заболевания достаточно нескольких дней, для других известных заболеваний на это могут потребоваться месяцы, годы или десятилетия.

Если исследование выполнено только с помощью анализа материалов, взятых из историй болезни или иных документов, то это называется *ретроспективным сбором* данных. Например, по архивным документам выделяют группу лиц, прошедших в клинике некоторое обследование 10 лет назад, и выясняют их состояние на сегодняшний день. Очевидный недостаток - невозможность изменить методику первичного обследования, подбор групп, а достоинство - это оперативность и дешевизна.

Ретроспективно, в прошлом, отмечают у больных наличие какого-то общего фактора, например, курения. Если в контрольной группе этот фактор в прошлом встречался реже, то можно предположить, что заболевание связано с ним. Такая организация соответствует *исследованию типа «случай-контроль» (ИСК)*. В основном этот вид исследования был распространен в 50-е годы.

К сожалению, никогда нет уверенности, что наблюдаемая в этих исследованиях группа отражает общие для всех больных характеристики. Кроме того, сведения о прошлом не бывают абсолютно верными. Поэтому *результаты ИСК обычно не рассматриваются как исчерпывающе доказательные*.

Тем не менее, поскольку такая организация исследований относительно дешева и проста, методология такого типа исследований применяется.

**Перспективные рандомизированные контролируемые исследования служат «золотым стандартом» для оценки эффективности лечебных и профилактических вмешательств.**

**Проспективным (prospective)** называют исследование, когда группа больных специально формируется и целенаправленно отслеживается.

Принципы организации проспективных исследований подробно описаны [38] и непрерывно совершенствуются. Главная особенность таких работ - полная определенность наблюдаемой группы, выявляемых состояний и методов обследования **до начала исследования**. Проспективные исследования дорогостоящи, однако они самые плодотворные и доказательные.

*Проспективные исследования предполагают наличие опытной и контрольной групп, пациентов распределяют по группам случайным образом (рандомизация), при этом группы не должны различаться по параметрам, влияющим на исход заболевания.*

Состояние всех пациентов прослеживается в течение определенного, часто довольно длительного времени (проспективное исследование), по истечении которого сравнивается частота наступления клинически важных конечных точек (выздоровление, смерть, осложнения) в опытной и контрольной группах.

Различают наблюдения, при проведении которых *исследуется естественное развитие заболеваний*. При этом результаты обследования соотносятся с предшествовавшими им исходными особенностями изучаемых параметров. Такие работы называют *исследованиями причинных факторов*. Например, заболеваемость ишемической болезнью сердца зависит от артериального давления и концентрации холестерина в крови.

Проспективные исследования могут быть и с *преднамеренным вмешательством* в естественный ход событий. К таким исследованиям относят многие испытания средств лечения и профилактики.

Варианты их проведения — параллельное исследование двух групп, последовательное применение испытуемого лечения и контрольного (перекрестно или с самоконтролем), использование в качестве контроля другой, внешней группы пациентов.

Контролем могут быть:

1. Плацебо.
2. Другое активное лечение.
3. Группа, не получающая никакого лечения.

4. Другая доза того же препарата.

5. Группа, получающая «обычное лечение». Оно строго *не оговаривается протоколом*, и в этом состоит *отличие* от группы «другое активное лечение», где сравнительная терапия *четко определена протоколом*.

6. Сравнение с анамнестическими данными этих же пациентов.

7. Сравнение с анамнестическими данными других пациентов.

Отношение к анамнестическому (историческому) контролю в настоящее время скептическое. Spilker В. приводит пример, когда в 44 из 56 работ с историческим контролем изучаемый препарат имел преимущество перед сравнительной терапией; в то же время, лишь 10 из 50 рандомизированных проспективных исследований это подтвердили. Основная причина – невозможность полностью избежать субъективности при подборе исторической группы сравнения, что приводит к ложным выводам.

## 7. Принципы планирования и составления программ клинических испытаний

Наиболее богатый критический опыт накоплен при разработке методов лечения с использованием новых лекарственных средств. Обусловлено это тем, что в современном мире лекарства сопутствуют человеку в течение всей его жизни, поэтому врачи и пациенты должны быть уверены в их качестве, надежности и безопасности.

**Создание нового препарата невозможно без научных исследований с участием человека в качестве субъекта-испытуемого.** Это связано с тем, что получить полную информацию о лекарстве с учетом только экспериментов на животных практически нереально. У животных отличается фармакокинетика препаратов, сложно моделировать многие болезни, а адекватно исследовать реакцию на психотропный препарат вообще невозможно.

При разработке и изучении новых лекарственных средств приходится решать две главные проблемы: с одной стороны, необходимо *получить достоверные данные об эффективности и безопасности лекарственного средства* у человека, с другой, при проведе-

нии клинических исследований *нужно соблюдать права личности, не подвергать участников медицинского эксперимента излишнему риску*, следуя заповеди “не навреди”.

Еще в древности ученые-медики пытались анализировать и сравнивать действие различных лекарственных средств и противопоставлять научно обоснованные данные невежественным заблуждениям. Тогда же возникли и проблемы проведения клинических испытаний как нравственного, так и научного характера. В III веке до н.э. Птолемей настаивал на том, чтобы при испытании лекарств на осужденных преступниках было получено их согласие. Долгое время соблюдение этических норм оставалось на совести врачей и не было определено официальными законами. Даже в нынешнем столетии имело место немало негуманных исследований. Достаточно вспомнить опыты нацистских медиков над узниками концлагерей, испытания, проведенные на заключенных в США, исследования над психически больными в некоторых странах Европы.

В 1932 году общественная служба по венерическим заболеваниям США в штате Джорджия начала исследование естественного течения нелеченного сифилиса у 400 чернокожих больных из бедных слоев общества. Никто из них не знал о своем участии в эксперименте. Ужасный опыт над людьми длился до 1972 года, хотя к этому времени уже имелись эффективные средства лечения данного заболевания.

Таким образом, до недавнего времени права людей, участвующих в медицинских исследованиях в качестве субъектов, грубо нарушались. Больные зачастую не знали о том, что являются участниками клинических испытаний. Их жизнь и безопасность подвергались излишнему риску. В современном цивилизованном мире, где правам человека уделяется столь большое внимание, сложившаяся ситуация не могла долго существовать.

Толчком к разработке всеобъемлющих правил проведения клинических исследований послужила катастрофа, вызванная вследствие применения талидомида в 1959–1961 гг. В ФРГ, а затем в некоторых других странах возникла вспышка фокомелии (“тюленьи конечности”) – врожденной аномалии развития, при которой отмечались дефекты трубчатых костей конечностей и формировались

нормальные или рудиментарные кисти и стопы, расположенные близко к туловищу, как у тюленей. Одновременно встречались и другие пороки развития. Спонтанно фокомелия развивается крайне редко. Ретроспективные исследования продемонстрировали связь между приемом талидомида во время беременности и рождением детей с аномалиями развития, но эти данные оказались недостаточно убедительными.

***Были организованы проспективные наблюдения за беременными, принимавшими талидомид.***

Случилось худшее – новое средство стало причиной самой страшной трагедии в истории современной лекарственной терапии. Многие дети, рожденные с фокомелией, умерли, но многие живут с деформациями конечностей, глаз, ушных раковин, сердца.

В ФРГ было зарегистрировано 10 тыс. детей с деформациями конечностей, из которых 5 тыс. выжили. Всего в мире насчитывается около 10 тыс. таких больных. Катастрофа с использованием талидомида показала, что *производство, применение и клинические испытания лекарственных средств должны находиться под строгим государственным контролем.*

Изучение эффективности и безопасности новых лекарств должно проходить стандартизировано и беспристрастно и основываться на объективных наблюдениях, а не на личном мнении отдельных, пусть даже ведущих в своей области специалистов. Но сама процедура клинических испытаний должна носить гуманный характер. *Права и интересы пациентов должны быть надежно защищены.*

В соответствии с Правилами «Качественной клинической практики» (GCP) клиническое испытание проходит в правовом пространстве, которое определяют, с одной стороны, официальные инстанции (в России - это Министерство здравоохранения, Фармакологический государственный комитет и Государственный институт доклинической и клинической экспертизы лекарств), с другой, независимый этический комитет.

Четыре стороны участвуют в любом современном клиническом испытании нового лекарственного средства: спонсор исследования (обычно это фармацевтическая компания), монитор (или контракт-

тно-исследовательская организация), врач-исследователь и пациент. Успешное выполнение проекта предполагает сотрудничество спонсора, исследователя и монитора во благо конкретного пациента и общества в целом.

***Совсем недавно в России введены в действие Отраслевой стандарт “Протоколы ведения больных. Общие требования” (приказ МЗ РФ от 03.08.99 г. № 303) и, что особенно важно, приказ № 253 от 30.06.99 г. “О совершенствовании клинических исследований лекарственных средств”.***

Последний приказ разработан с целью обеспечения государственной политики в области эффективности и безопасности лекарственных средств и защиты прав пациентов, вовлеченных в клиническое исследование. В нем четко определено, что клинические исследования проводятся для получения с помощью применения научных методов оценок и доказательств эффективности и безопасности лекарственных средств, данных об ожидаемых побочных эффектах от применения лекарственных средств и эффектов взаимодействия с другими лекарственными средствами.

Этим документом установлен единый порядок *проведения аккредитации учреждений здравоохранения*, в том числе клинических баз научно-исследовательских институтов и высших учебных заведений, а также других лечебно-профилактических, *проводящих клинические исследования*.

Утверждено Положение о порядке аккредитации и лицензирования учреждений здравоохранения, имеющих право проводить клинические исследования лекарственных средств. Организация работы по лицензированию возложена на Департамент научно-исследовательских и образовательных медицинских учреждений Минздрава России\*.

---

\* Пока рукопись находилась в наборе, вышел Приказ № 422 от 24.11.99 г., в котором функции по выдаче учреждениям здравоохранения лицензий на проведение клинических исследований лекарственных средств возложены на Департамент государственного контроля качества, эффективности, безопасности лекарственных средств и медицинской техники. Эти функции предусмотрены Постановлением Правительства от 12.11.99 г. № 1246 "О внесении изменений и дополнений в Положение о Министерстве здравоохранения Российской Федерации".



Применению препарата у человека предшествует огромное число токсикологических, фармакокинетических и фармакодинамических исследований на животных. Поэтому вышесказанное о необходимости тщательного планирования программы клинических испытаний в полной мере относится и к доклиническим исследованиям. Не меньшее, а может быть и большее значение имеет грамотное выполнение самого клинического испытания, следование требованиям протокола, корректность и честность при сборе и анализе данных.

### 7.1. Фазы клинических испытаний

Фазы клинических испытаний четко разграничены. Обычно лекарственное средство проходит четыре фазы клинических испытаний; вторая фаза делится на фазы IIa и IIb.

**Фаза I.** Первый опыт применения нового активного вещества у человека. Чаще всего, исследования начинаются у добровольцев (взрослые здоровые мужчины). Главная цель исследований – решить, стоит ли продолжать работу над новым препаратом и, если удастся, установить дозы, которые впоследствии будут использованы во время II фазы испытаний.

В процессе выполнения I фазы исследователи получают предварительные данные о безопасности препарата и делают первое описание его фармакокинетики и фармакодинамики у человека.

Исследования I фазы – это широкий спектр медицинских экспериментов. Они обычно продолжаются и тогда, когда уже начата II, а иногда и III фаза испытаний (обычно все фармакокинетические исследования проводятся уже в I фазе).

#### **В ходе испытаний I фазы исследуют:**

1. Безопасность, переносимость, фармакокинетику (ФК) и фармакодинамику (ФД) одной дозы (в том числе определение максимально переносимой дозы).
2. Безопасность, переносимость, ФК и ФД множественных доз.
3. Биодоступность.
4. Пропорциональную ФК и ФД одной дозы и множественных доз при различных путях введения.

5. Метаболизм лекарственного средства и его взаимосвязь с массой тела.

6. Влияние возраста, пола, характера принимаемой пищи, функции печени и почек на ФК и ФД одной дозы и множественных доз.

7. Лекарственные взаимодействия.

**Исследования I фазы имеют общие особенности:**

1. Участие небольшого числа добровольцев; в среднем от 4-х до 24-х человек (до 80 в течение выполнения всей I фазы).

2. Каждое из исследований проводится в одном центре.

3. Каждое исследование длится несколько дней, максимум несколько недель.

4. Тщательно контролируются медицинским персоналом (обычно добровольцы находятся под наблюдением 24 ч в сутки).

Иногда повышенная токсичность препарата (например, для лечения рака или СПИДа) делает проведение таких исследований у здоровых добровольцев неэтичным. В этих случаях их выполняют с участием только пациентов, страдающих соответствующим заболеванием. Обычно эти нетерапевтические исследования проходят в специализированных учреждениях.

**Фаза IIa.** Как правило, это первый опыт применения лекарственного средства у пациентов с заболеванием, для лечения которого предполагается использовать данный препарат.

В ходе IIa фазы необходимо убедиться в активности исследуемого вещества, оценить краткосрочную безопасность, установить популяцию пациентов, режим дозирования, выяснить зависимость эффекта от дозы, определить критерии оценки эффективности.

Испытания проводятся на ограниченном числе пациентов (100-300), за которыми осуществляется тщательное наблюдение, иногда в стационаре.

**Фаза IIb.** Более обширные исследования у пациентов с заболеванием, которое является основным предполагаемым показанием к назначению лекарственного средства (для лечения, диагностики или профилактики).

Главная цель – доказать эффективность и безопасность нового препарата. Результаты базовых исследований служат основой для

планирования исследований III фазы и в значительной степени влияют на решение о регистрации препарата. Многие считают исследования II фазы наиболее важным моментом в создании нового лекарства.

**Фаза III.** Многоцентровые испытания с участием больших (и, по возможности, разнообразных) групп пациентов (в среднем, 1000-3000 человек). В последнее время появился термин «мегаисследования» (megatrials), в которых могут принимать участие свыше 10 тыс. пациентов. Исследования III фазы проводят для получения дополнительных данных о безопасности и эффективности различных форм препарата.

В ходе выполнения III фазы изучают характер наиболее частых нежелательных реакций, клинически значимые лекарственные взаимодействия, влияние возраста, сопутствующих состояний. Обычно клинические испытания этой фазы являются двойными слепыми контролируруемыми рандомизированными исследованиями, условия проведения которых максимально приближены к обычным условиям применения препарата.

Данные, полученные в клинических испытаниях III фазы, являются основой для создания инструкции по применению лекарства и важным фактором для принятия официальными инстанциями решения о его регистрации и возможности медицинского использования.

В этой фазе выделяют период, к которому относят исследования, проводимые от времени подачи материалов для регистрации препарата в официальные инстанции до момента регистрации и получения разрешения на медицинское применение. Они выполняются для того, чтобы получить некоторые дополнительные сведения о препарате, оценить качество жизни, положение будущего препарата на фармацевтическом рынке.

**Фаза IV.** Исследования проводятся после начала продажи препарата с целью получить более подробную информацию о безопасности и эффективности различных лекарственных форм и доз, длительного применения в различных группах пациентов и при различных факторах риска и, таким образом, более полно оценить стратегию применения лекарственного средства.

В исследованиях принимает участие большое число пациентов, что позволяет выявить ранее неизвестные и редко встречающиеся нежелательные реакции. Существует понятие *постмаркетингового наблюдения*. Эти неэкспериментальные исследования-наблюдения иногда относят к V фазе клинических испытаний.

После регистрации препарата клинические испытания, целью которых является изучение новых, незарегистрированных показаний, методов применения или комбинаций, рассматриваются как испытания нового лекарственного средства, т.е. считаются исследованиями ранних фаз.

## **7.2. Использование плацебо и рандомизации в слепых исследованиях - возможность выявить снижение преднамеренных и непреднамеренные искажений**

*Плацебо* - это фармацевтический продукт, не содержащий активного начала. При проведении сравнительных исследований плацебо по форме, цвету, вкусу, запаху, методике назначения полностью имитирует исследуемый препарат. Результаты, полученные в группе пациентов, которым было назначено плацебо, являются как бы фоном, обусловленным естественным течением болезни в ходе клинического испытания без исследуемой терапии. Результаты, полученные в группе активного лечения, оценивают на этом фоне.

### ***Причины, по которым в исследования включают группу плацебо:***

1. Контроль психологических аспектов участия в клиническом испытании. Дело в том, что «внутренняя среда» исследования, в которой находится пациент, отличается от условий применения препарата в обычной практике.

2. Подписание информированного согласия и осознание своего участия в научном исследовании «самого современного препарата», повышенное внимание медицинского персонала, значительное количество дополнительных обследований, необходимость часто посещать лечебное учреждение - все это влияет на ответ организма в процессе проводимого лечения.

3. Правильная оценка частоты нежелательных явлений. К нежелательным явлениям относят любое неблагоприятное с медицин-

ской точки зрения событие, происходящее в период клинического испытания, поэтому какое-то число таких событий не связано с приемом исследуемого препарата. Предполагается, что эти «фоновые» события наблюдаются с одинаковой частотой в обеих группах, что позволяет выделить у пациентов действительно лекарственные реакции.

4. Возможность корректно интерпретировать данные и сделать правильные выводы об эффективности и безопасности лекарственного средства.

**Слепые исследования.** В каждом клиническом испытании принимает участие большое число сторон: пациент, исследователь, вспомогательный персонал, монитор, статистики, группа обработки и анализа данных. Любой из участников, невольно или преднамеренно, может исказить данные и повлиять тем самым на результаты исследования. Например, пациент, знаящий, что он получает активный препарат, может более настойчиво искать у себя признаки улучшения и давать «позитивную» интерпретацию настоящим или мнимым изменениям в своем состоянии.

Врач-исследователь, заведомо убежденный в преимуществах одного из сравниваемых видов лечения, может субъективно трактовать те или иные наблюдения. Поэтому в сравнительных исследованиях как можно меньше участвующих сторон должны знать, какое лечение получает каждый конкретный пациент. Тем самым значительно снижается возможность преднамеренных искажений, а непреднамеренные распределяются между группами примерно в равной пропорции и нивелируют друг друга.

В зависимости от широты круга, находящегося в неведении относительно терапии у конкретного пациента, выделяют следующие исследования:

1. *Открытые* - врач и пациент знают, какая терапия назначена.
2. *Простые слепые* - пациент не знает, какое лечение ему назначено, а исследователь располагает такой информацией.
3. *Двойные слепые* - ни врач, ни пациент не знают назначенной терапии.

4. *Тройные слепые* - ни врач, ни пациент, ни те, кто организует исследование и обрабатывает данные не знают назначенной терапии у конкретного пациента.

Существует термин «*полные слепые исследования*», когда в неведении относительно лечения находятся все, кто взаимодействует напрямую с пациентом, исследователем и данными, например, рентгенологи, патологоанатомы, статистики. При этом все стороны не имеют информации о типе лечения до завершения анализа полученных результатов. Коды, контролирующие распределение пациентов между группами, подготавливают лица, не связанные с данным исследованием. В ходе двойных и тройных слепых исследований в каждом исследовательском центре находятся закрытые коды, которые в случае медицинской необходимости (например, при возникновении некоторых серьезных нежелательных явлений) могут быть вскрыты для того, чтобы определить, какую терапию получает пациент.

Желательно, чтобы как можно большее число исследований могли быть двойными или тройными слепыми. Даже при оценке эффективности физиотерапевтических методов или иглоукалывания возможно применение слепых методов.

Особенно слепое клиническое испытание желательно проводить в случаях, когда велика роль сопутствующего лечения. Например, в ходе проведения клинического испытания антибиотиков при раневых инфекциях необходимо, чтобы методика обработки и перевязки ран проводилась одинаково всем больным. В данном случае двойное «ослепление» абсолютно востребовано.

Довольно часто для проведения слепого клинического испытания приходится применять творческий подход. Например, если стандартный лекарственный препарат вводится ингаляционным способом, а новый — посредством инъекций, то стандартное лечение можно дополнить инъекциями плацебо, а новое лечение — ингаляциями плацебо. Тогда больной и медицинский персонал не смогут различать два разных способа лечения.

К исключениям относятся фармакокинетические исследования I фазы, ранние исследования II фазы, длительные исследования II и III фаз, обычно являющиеся продолжением двойных слепых ис-

следований, а также исследования, которые не могут быть двойными слепыми или тройными слепыми по этическим соображениям.

Кроме термина «слепые» существует имеющий то же значение термин «маскированные» исследования. В соответствии с традициями он применяется, например, при проведении клинических испытаний офтальмологических препаратов.

Естественно, что процесс распределения пациентов между группами должен иметь элементы случайности, то есть быть рандомизированным, что позволяет добиться эквивалентности групп между собой и максимально уменьшить необъективность при получении данных. Группы должны быть эквивалентны по продолжительности и тяжести заболевания, полу, возрасту, сопутствующим состояниям и терапии – тогда клиническое испытание будет сбалансированным. В противном случае сравниваемые между собой виды терапии будут применяться в различных по составу группах, т.е. как бы у разных популяций пациентов. Это не позволит сравнить эффекты от разных методов лечения и сделать правильные выводы. Например, если в одну группу попадут пациенты с более тяжелым течением заболевания, то эффективность лечения может быть ниже.

Для клинического испытания методов лечения разработаны весьма полные стандарты [20, 42-44]. Однако следует отметить, что в настоящее время проведение клинических испытаний лекарственных и нелекарственных методов лечения является не только наукой, но и особым медицинским бизнесом.

По ряду обстоятельств, в том числе и по этой причине, значительная часть статей, представляющих результаты испытания методов лечения, как совершенно справедливо отмечает В.В. Власов [36], содержит непригодную для использования информацию.

Полноценная статья должна включать удовлетворительное описание 10 наиболее важных принципов в отношении описания структуры и представления результатов исследования. Эти своеобразные критерии достоверности относятся к любому вмешательству — *лечебному или профилактическому.*

## 8. Методологические стандарты испытаний методов лечения (критерии достоверности)

### **Структура исследования:**

1. Наличие контрольной группы.
2. Ясные критерии отбора (включения и исключения) больных.
3. Включение больных в исследование до рандомизации.
4. Случайный метод распределения больных по группам (рандомизация).
5. “Слепое” лечение.
6. “Слепая” оценка результатов лечения.

### **Представление результатов:**

7. Информация об осложнениях и побочных эффектах лечения.
8. Информация о числе больных, выбывших в ходе клинических испытаний.
9. Адекватный статистический анализ, а также ссылки на использованные статьи и программы.
10. Информация о размере выявленного эффекта и статистической силе исследования.

**Самый простой критерий качества исследования — это его структура.**

В исследовании должны участвовать “опытная” и “контрольная” группы, соответственно получающие изучаемое лекарственное средство и плацебо или иное лечение (обычно наиболее приемлемое на момент клинического испытания). Если никакого контроля нет, оценить эффект изучаемого лечения невозможно.

**Где клинические испытания неправильно организованы - там, как правило, чаще приводятся “блестящие” результаты.**

К сожалению, такого рода неполноценных статей множество, и их тем больше, чем чаще применяются экзотические средства лечения (магнитные поля, лазерное облучение крови, использование пищевых добавок и минеральной воды).

Если критерии *исключения* пациентов из испытания достаточно часто указываются в публикациях, то критерии *включения* — редко. Вместе с тем, необходимо знать, по какому принципу отбирали больных и где проводили лечение. Существенным является и



то, включались ли, например, в испытание все больные, получавшие лечение, или только лечившиеся в университетской клинике.

Другим распространенным недостатком является то, что не указываются причины, по которым больные выбывают в процессе выполнения испытаний.

Можно ожидать самых разнообразных искажений в следующих ситуациях:

- а) выбывшие больные достоверно отличаются по каким-то признакам от наблюдавшихся до конца испытания;
- б) если выбыло более 20% пациентов.

У больных, не получающих предписанного лечения, состояние здоровья и прогноз обычно хуже. Поэтому в идеале должны быть представлены результаты как у лиц, фактически получивших предписанное лечение, так и у всех больных, которым лечение было запланировано при рандомизации [46]. Очень важно в организации клинического испытания — определить сроки, когда следует прекратить лечение.

**Необходимо оценить применяемые методы статистического анализа.** Реже всего приводится описание статистической силы исследования, т.е. возможности выявить с определенной степенью вероятности небольшие различия между сравниваемыми группами. Статистическая сила снижается при использовании малочисленных групп в “опыте” и “контроле”. Это особенно важно при получении отрицательного результата исследования.

Оценить достаточность числа больных для отрицательного результата исследования можно с помощью приводимых таблиц. В ячейке, соответствующей частоте событий в группе лечения и в контрольной группе, представлено число больных в каждой группе, необходимое для выявления снижения частоты на 25% (табл. 2) или 50% (табл. 3). Если в рассматриваемой статье число пациентов окажется меньшим, значит эффект мог быть не обнаружен только из-за малого количества обследуемых.

Таблица 2

**Оценка достаточности числа больных в группах сравнения: величина выборки в каждой из двух групп, необходимая для отклонения гипотезы о 25% разности частоты определенного исхода в группах по A.S. Detsky, D.L. Sackett [46] и В.В. Власову [36]**

Частота исхода в контрольной группе	Частота исхода в группе лечения									
	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05
0,95	14	68								
0,9	11	38	105							
			7							
0,8		18	78							
0,7		11	28	165						
0,6			15	43	402					
0,5				20	65	160				
						7				
0,4					11	25	98			
0,3						14	30	152		
0,2						13	34	261		
0,1							13	34	589	
0,05								17	74	1245

*Примечания: Результат эксперимента отражается в частоте определенного исхода (выздоровления, смерти, исчезновения клинических симптомов) в группе лечения и в контрольной группе. При оценке результатов различие между данными групп пациентов недостоверно. Таблица показывает минимально необходимое число больных в каждой группе, чтобы быть уверенным в том, что клинически значимое различие в 25% ( $p < 0,05$ ; одностронний тест) не может быть случайно отвергнуто. Таблица предназначена только для оценки результатов исследования, а не для планирования исследований.*

Таблица 3

**Оценка достаточности числа больных в группах сравнения: величина выборки в каждой из двух групп, необходимая для отклонения гипотезы о 50% разности частоты определенного исхода в группах [36, 46]**

Частота исхода в контрольной группе	Частота исхода в группе лечения												
	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,08	0,06	0,04	0,02		
0,98	14	50	580										
			3										
0,95	12	37	921										
0,9	14	26	236										
0,8		15	63										
0,7			29	268									
0,6				17	68								
0,5				11	30	260							
0,4					16	62							
0,3						10	26	205					
0,2							20	145					
0,1								17	102	254	2017		
											7		
0,08								14	66	131	453		
0,06								12	44	76	179	1313	
0,04								10	31	47	87	274	
0,02									22	30	47	97	561

Результат эксперимента отражается в частоте определенного исхода (выздоровления, смерти, исчезновения клинических симптомов) в группе лечения и в контрольной группе. При оценке результатов различие между группами недостоверно. Таблица показывает необходимое минимальное число больных в каждой группе, чтобы быть уверенным в том, что клинически значимое различие в 50% ( $p < 0,05$ ; односторонний тест) не может быть случайно отвергнуто. Таблица предназначена только для оценки результатов исследования, а не для планирования исследований!

При статистическом анализе оценка различий между группами должна быть скорректирована с учетом факторов, потенциально влияющих на исход (прогноз).

Исчерпывающее описание всех 10 позиций редко встречается даже в статьях, опубликованных в лучших журналах. Это не означает, что высокое качество статей недостижимо, просто редакторы сформулировали достаточно ясные и **общие для всех журналов требования** относительно недавно [42]. Поэтому при рассмотрении статьи нужно подходить к обнаруживаемым недостаткам клинического испытания взвешенно и попытаться ответить на 6 основных вопросов.

*Было ли распределение больных по группам рандомизированным?*  
Не следует доверять статьям об исследованиях, где распределение больных по группам было нерандомизированным (неслучайным) или метод распределения был неудовлетворительным, или вообще нет его описания.

В.В. Власов [36] с ссылкой на D.H. Spodick [47] приводит ряд причин отказа от рандомизации в клинических исследованиях. Главная из них - недостаточная осведомленность врачей. Но есть еще ряд “поведенческих” причин:

- существующий метод лечения может поддерживаться (или отвергаться) авторитетными учеными или сложившейся традицией;
- отношения врача и больного формируют желание помогать, а не испытывать новый вид лечения;
- нежелание подвергать свой авторитет, свой способ действий испытаниям;
- боязнь того, что при случайном распределении лечения больной не получит оптимальной помощи или лекарство может оказаться вредным (это заблуждение, так как именно этими опасениями и определяется необходимость проведения рандомизированных испытаний);
- неблагоприятное влияние результатов предыдущих, плохо выполненных испытаний, формирующих ошибочные исходные ожидания.

Следует четко представлять, что именно *из-за плохо выполненных первичных клинических испытаний* множество больных получало лечение, которое позднее, при тщательно проведенной проверке, оказалось неэффективным (например, перевязка внутренней грудной артерии, профилактика атеросклероза клофибратом, лечебная бронхография, лечение язвы желудка замораживанием). Список лечебных и профилактических методов, которые при проведении качественных клинических испытаний оказались бесполезными, можно продолжать на многих страницах.

Обилие исследований, не выдерживающих критики, согласуется с законом: *объем некачественных данных всегда имеет тенденцию к превалированию над доброкачественными сведениями.*

При использовании результатов несовершенных исследований существует способ взвесить вероятность ошибки (анализ чувствительности к смещениям). Он не слишком сложен [48], но выходит за пределы нашего обсуждения.

*Ориентировано ли исследование на основной критерий оценки эффективности лечения?*

Подмена “конечных” результатов (истинных критериев оценки — клинических исходов) “промежуточными” (косвенными критериями оценки вроде снижения уровня глюкозы или холестерина в крови, артериального давления) весьма распространена.

Как бы ни снижался уровень смертности от сердечно-сосудистых заболеваний, но если показатель общей смертности повышается, значит, основной результат не достигнут. ***Если исследование с отрицательным результатом отвечает критериям качества, то оно может быть полезным уже тем, что позволит не вводить в практику новый неэффективный метод лечения.***

*Значимы ли результаты исследования? Достоверны ли они статистически?*

*Статистически значимо* то, что действительно происходит с высокой вероятностью. *Клинически значимо* то, что своими размерами убеждает врача в необходимости изменить свою практику в пользу нового образа действий.

Малый размер эффекта означает, что лишь очень небольшая часть больных (например, 1%) получает пользу от такого лечения

[49]. Размер эффекта может быть выражен различным образом. **В настоящее время популярна мера эффективности вмешательства — число больных, которых надо пролечить, чтобы получить эффект у одного дополнительного больного (ЧБНЛ, number needed to treat — NNT)**. Это, пожалуй, самая очевидная мера определения эффективности лечения при его сравнении с другим методом лечения.

Необходимо дополнительно оценить полученные результаты и с точки зрения самого больного. Лечение может быть травматичным, неприятным, болезненным, неэстетичным, противоречащим религиозным установкам, а это иногда может быть решающим основанием для отказа от него.

*Доступен ли метод в практике?*

Приведенные критерии качества относятся не только к одному виду клинических испытаний — собственно испытаниям методов лечения, но и касаются многочисленных исследований, посвященных изучению побочных эффектов, оптимизации лечения, комбинированного лечения и прочих уточняющих обстоятельств.

Все исследования эффектов лечебного вмешательства должны проводиться по схеме контролируемого клинического испытания, независимо от того, считаются ли эти эффекты главными или второстепенными, считается ли лечение основным или дополнительным. Отход от стандарта качества исследования неминуемо приводит к получению некачественных данных.

## 9. Оценка публикаций, посвященных методам диагностики

Диагностический тест (прием, метод) может быть использован в разных целях.

◆ Если исследование проводят людям, считающим себя здоровыми, с целью выявить среди них тех, кто болен, хотя они и не подозревают об этом, такое исследование называется *скринингом* (от англ. *screen* — *просеивать*). Пример — массовые флюорографические обследования.

◆ Если исследование выполняют всем пациентам, оказавшимся в больнице или поликлинике в связи с любым заболеванием, то это называется *поиском сопутствующих заболеваний*.

◆ Если исследование проводят пациентам, обратившимся с жалобами, для выяснения причины их страданий, уточнения диагноза, то это — собственно *диагностическое исследование*, самый важный случай. Примеры — электрокардиография при жалобах на боль в области сердца.

◆ Если исследование проводят людям, получающим лечение, чтобы оценить его успешность, то это — тест на эффективность лечения (*мониторинг*).

Для оценки значимости необходимо получить ответ на следующие вопросы:

1. Есть ли сопоставление с “золотым стандартом”?
2. Проведено ли это сопоставление слепым методом?
3. Четко ли описан способ формирования изученной группы пациентов?
4. Имели ли обследованные пациенты достаточно широкий спектр заболевания (легкие и тяжелые случаи, лечившиеся и нелеченные, с поражениями разной формы, с сопутствующими заболеваниями)?
5. Оценена ли воспроизводимость результатов теста?
6. Оценена ли надежность интерпретации результатов слепым методом?
7. Рационально ли определена “норма” — отсутствие болезни?
8. Если тест предлагается как дополнение к пакету других тестов, то показано ли, что применение этого теста улучшает общий результат диагностики?
9. Обоснована ли “точка разделения” здоровых и больных?
10. Достаточно ли подробно описана тактика применения теста, чтобы его можно было правильно использовать?
11. Оценена ли общая полезность теста?

Разрабатывая новый метод, авторы обычно имеют в виду, что он обеспечивает определенные преимущества перед теми, которые уже применяются. Например, диагностика становится прижизненной, более ранней, более щадящей, более дешевой, в то время как метод “золотого стандарта” точен, но его использование ограничено.

Обязательная задача исследования диагностического метода — оценить, насколько хорошо диагноз, поставленный с помощью использования нового метода, соответствует “золотому стандарту” (референсному методу), самому надежному из соответствующих методов.

Следует четко уметь определять - удовлетворителен ли референсный метод? Действительно ли он соответствует наиболее надежному методу диагностики, который рассматривается медицинской практикой как стандартный? Иногда взятый авторами референсный метод сам по себе неудовлетворителен. В этом случае не следует тратить время, чтобы разобраться, соответствует ли ему новый тест.

Традиционный метод оценки результатов теста - построение так называемого латинского квадрата (четырёхпольной таблицы), представлен табл. 4.

Таблица 4

**Четырёхпольной таблица сопряженности -  
так называемый латинский квадрат**

Обследуемые	Показатели при обследовании		Всего
	здоровых	больных	
	Отрицательный результат	Положительный результат	
Здоровые	Истинно-отрицательные случаи (ИО) (здоров) <b>A</b>	Ложно-положительные случаи (ЛП) (болен) <b>C</b>	<b>A + C</b>
Больные	Ложно-отрицательные случаи (ЛО) (здоров) <b>B</b>	Истинно-положительные случаи (ИП) (болен) <b>D</b>	<b>B + D</b>
Всего	<b>A + B</b>	<b>C + D</b>	



Варианты совпадения и несовпадения референсных данных и результатов диагностического теста:

- истинно-отрицательные случаи (ИО), когда истинное отсутствие заболевания совпало с отрицательным результатом теста (здоров);
- ложноотрицательные случаи (ЛО), когда у больных людей получен отрицательный результат (здоров);
- ложноположительные случаи (ЛП), когда истинное отсутствие заболевания совпадает с положительным результатом теста (болен);
- истинно положительные случаи (ИП), когда истинная болезнь совпадает с положительным результатом (болен).

Данные, подставляемые в табл. 4, позволяют вычислять так называемые операционные характеристики теста: *чувствительность*  $D/(B+D)$ , *специфичность*  $A/(A+C)$ , *прогностическую ценность положительного* (ПЦП)  $=D/(C+D)$  и *отрицательного* (ПЦО)  $=A/(A+B)$  *результатов теста.*

В исследовании обязательно должна быть приведена такая таблица или сведения, достаточные для ее создания. В противном случае невозможно узнать, полезен ли рассматриваемый диагностический тест.

Чувствительность и специфичность являются *стабильными характеристиками диагностического теста.* Это означает, что они не зависят от распространенности болезни в исследуемой группе. Если чувствительность равна 0,75 (75%), то независимо от того, сколько пациентов с этой болезнью находится в исследуемой группе, 3/4 из них будут выявлены с помощью теста. Если специфичность составляет 0,9, то среди лиц, не имеющих данного заболевания, положительный результат (ложный!) будет получен лишь у 10%.

При выполнении диагностического теста врача в основном интересует, насколько высока вероятность болезни у лиц с положительным результатом и насколько она низка у пациентов с отрицательным результатом. Данный показатель называется *прогностической ценностью, положительных (ПЦП) и отрицательных (ПЦО) результатов.* Этими значениями нельзя ограничиваться,

так как их величина зависит от распространенности заболевания в изучаемой группе.

Если исследователи обращают особое внимание на высокую надежность постановки диагноза (прогностической ценности положительного результата), надо знать, каковы чувствительность и специфичность. Возможно, привлекательный результат — просто следствие того, что в исследуемой группе распространенность заболевания очень высока, допустим, благодаря предварительному отбору пациентов.

В.В. Власов в своей работе об исследованиях, посвященным методам диагностики, приводит яркие и очень простые примеры, систематизируя требования, предъявляемые к диагностическому тесту [34].

Электрокардиография, незаменимая при обследовании больных с возможными заболеваниями сердца, практически оказывается бесполезной в профилактических обследованиях населения. Нагрузочное ЭКГ-тестирование — весьма эффективный метод для обследования мужчин среднего возраста, предъявляющих жалобы на боли в груди. Однако этот метод не информативен при профилактических обследованиях молодых мужчин и женщин в условиях низкой распространенности ишемической болезни сердца.

В свою очередь, у пожилого мужчины, предъявляющего жалобы, соответствующие типичной стенокардии, нагрузочное ЭКГ-тестирование практически бесполезно, так как оно не может ни заметно увеличить вероятность ИБС при положительном его результате, ни существенно снизить вероятность болезни при отрицательном.

Следовательно, нагрузочное ЭКГ-тестирование (впрочем, как и любой другой диагностический тест) наиболее полезно проводить в условиях неопределенности диагноза.

На наш взгляд, в некоторых случаях при интерпретации результатов исследования желательно различать такие категории, как *параметр* и *признак*. К *параметрам* относят показатели, характеризующие свойство процесса, явления или системы в абсолютных или относительных величинах, *признакам* — альтернативные характеристики, имеющие только два значения: наличие или отсутствие.

Параметр переходит в разряд признаков, когда его оценка проводится по отсутствию или наличию повышения (или понижения) относительно условной границы. Одним из авторов [51] в собственных исследованиях еще в 1990 г. при изучении нейрогормональной дисрегуляции при воздействии профессиональных вредностей физической и химической природы был использован такой подход.

Проводился анализ частоты выхода показателей за пределы доверительного интервала контрольной группы с учетом средней ошибки ( $m$ ), характеризующей точность определения средней величины, с целью установления операционных характеристик тестов. Для оценки индивидуальных результатов изменений границы референсных интервалов устанавливали с учетом среднеквадратического отклонения -  $1,5 \delta$ . За границу или точку разделения принимали верхний (в случае повышения) или нижний (в случае понижения) доверительный интервал средних значений концентраций гормонов у пациентов контрольной группы  $M + tm$ , где значение  $t$  соответствовало величине выборки при 5% вероятности ошибки ( $p < 0,05$ ). Такой прием дал возможность определить эффективность разрабатываемых тестов путем расчета их операционных характеристик – чувствительности и специфичности.

*Чувствительность* (sensitivity, Se) определялась как доля лиц с патологией, у которых отмечался позитивный результат (выход за дискриминантную линию):

$$Se = (PS/S) \times 100\%,$$

где PS – истинно положительные результаты;

S – количество человек в группе с патологией.

*Специфичность* (specificity, SP) – это вероятность отрицательного результата у лиц без патологии (контрольная группа):

$$SP = (NH/H) \times 100\%,$$

где NH – истинно отрицательные результаты,

H – количество лиц в контрольной группе.

Наиболее информативным считали тест, который при данной прогностичности положительного результата (PVP) имел большую прогностичность отрицательного (PVN) или при данной прогностичности отрицательного – большую прогностичность положи-

тельного. В свою очередь, прогностичность положительного результата (PVP) определяли как частоту совпадения признака с патологией:

$$PVP = (PS/P) \times 100\%,$$

где PS – истинно положительные результаты;

P – общее число положительных результатов.

Прогностичность отрицательного результата соответствовала частоте его совпадения с отсутствием заболевания:

$$PVN = (NH/N) \times 100\%,$$

где NH – истинно отрицательные результаты,

N – общее число отрицательных результатов.

Десять лет назад приходилось объяснять, что термины «чувствительность» и «специфичность» использованы именно как операционные характеристики эффективности тестов, а не как характеристики специфичности воздействия неблагоприятных производственных факторов.

В настоящее время к диагностическому тесту предъявляют еще более строгие дополнительные требования [34]. При *правильной* организации исследования оценка результатов изучаемого теста должна быть выполнена “вслепую” — так, чтобы специалист не знал результатов других анализов у этого пациента, прежде всего — референсного теста. Его результаты тоже должны оцениваться вслепую, независимо от результатов изучаемого теста. В противном случае происходят ошибки типа “смещения к предполагаемому диагнозу”. Обычно они возникают непреднамеренно: зная о предположительном диагнозе, специалист склонен к предвзятости при оценке полученных данных.

*Необходимо, чтобы было четкое описание способа формирования исследуемой группы пациентов.* Должно быть указано, по какому принципу пациенты подбирались в конкретную группу, каковы критерии включения и исключения, какова исходная совокупность — посетители обычной поликлиники или больные специализированного стационара. От этого зависит не только распространенность болезни в исследуемой группе, но и ее тяжесть, соотношение разных форм болезни и многое другое.

Если группа пациентов “контрольная” (не имеет данной болезни) подбиралась отдельно, то ясно должно быть сказано, как это делалось. Иначе может случиться, что отличия “больных” от “здоровых” обусловлены тем, что “здоровые” — это молодые активные студенты, а “больные” — пациенты пожилого возраста, находящиеся в стационаре.

*Адекватно ли представлены в контрольной группе сопутствующие и сходные заболевания, а в группе больных по референсному критерию — разные формы основного заболевания?*

Возможно, полученные исследователями характеристики изучаемого теста относятся лишь к одной, тяжелой форме заболевания или к его поздней стадии, или только к случаям, когда отсутствуют другие сходные заболевания. Поскольку в обычной практике встречаются все формы болезни и любые сопутствующие заболевания, такой тест может оказаться практически бесполезным.

*Обязательно следует оценивать воспроизводимость диагностического теста.* Адекватный тест должен быть воспроизводимым и точным. Следует помнить, что воспроизводимость результатов инструментальных и, тем более, неинструментальных клинических исследований далеко не так высока, как обычно представляется. Большинство врачей переоценивают надежность объективной диагностики.

*Как правило, в клинической практике тест применяется в совокупности с другими. Однако полезность изучаемого теста должна быть продемонстрирована при его самостоятельном применении.*

В некоторых исследованиях при очевидной неэффективности диагностики с помощью изучаемого теста предлагается использовать его в качестве “дополнения к другим методам исследования”. Это ложная рекомендация. Тест, неэффективный сам по себе, никакой информативности не добавит к другим. Необходимо оценить, насколько дополнительное применение этого теста уточняет диагностику в целом. Только такая оценка эффективности комплекса тестов доказательна.

*Надежность интерпретации результатов теста должна быть оценена слепым методом.* Это означает, что результат оценивается повторно тем же специалистом (сходимость) или разными (воспро-

изводимость) в условиях, когда эти специалисты не имеют никаких других данных о пациенте.

*Особый вопрос связан с понятием “норма”, как отсутствие болезни.* При испытании диагностических тестов обычно сравнивают показатели больных не с аналогичными у “здоровых”, а с теми, у кого, согласно *референсному* диагностическому методу, нет изучаемой болезни. В некоторых случаях исследователь сопоставляет результат теста у “больных” с результатом в “норме”. В этом случае необходимо формулировать, что имеется в виду под “нормой”.

Концепция “нормы” используется очень широко и она весьма расплывчата. Почти любое толкование “нормы” имеет свои недостатки. Поэтому в лабораторном деле отказались от термина “норма” в пользу более конкретных “референсных величин”. Если подобрать группу людей, отвечающих некоторым стандартным требованиям клинического здоровья, то такую группу можно назвать *референсной*. Применение теста к людям, составляющим такую группу, в некоторых стандартных условиях и при стандартной методике дает совокупность референсных величин.

В статистическом смысле эти референсные величины распределены таким образом, что их можно описать с помощью известных математических способов. Например, многие фактические распределения удовлетворительно описываются с помощью средней величины и среднего квадратического отклонения, т.е. они, соответствуют закономерностям “нормального” или Гауссовского распределения. В случае, если фактическое распределение не соответствует какому-либо теоретическому или если исследователь принципиально не желает искать способ описания, он может найти границы, включающие, например, 95% средних величин (или 90, 99%) методом перцентилей. Такой интервал, содержащий большинство референсных величин, называют референсным интервалом. **Референсный интервал — это не норма “вообще”, а характеристика определенной референсной группы и условий исследования.** Выход за его пределы не считается *ненормальным*, а лишь нехарактерным для референсной группы. Важно подчеркнуть, что при любом подходе к определению границ “нормы” применение этого понятия на практике приводит к тому, что почти у всех здоровых людей обнару-

живаются какие-нибудь отклонения. Чем больше разнообразных тестов, тем больше отклонений. В итоге затрачиваются значительные средства на дополнительное обследование. Для разных болезней следует устанавливать свой диагностический порог (точку разделения). Он может находиться как внутри “нормы”, так и за ее пределами.

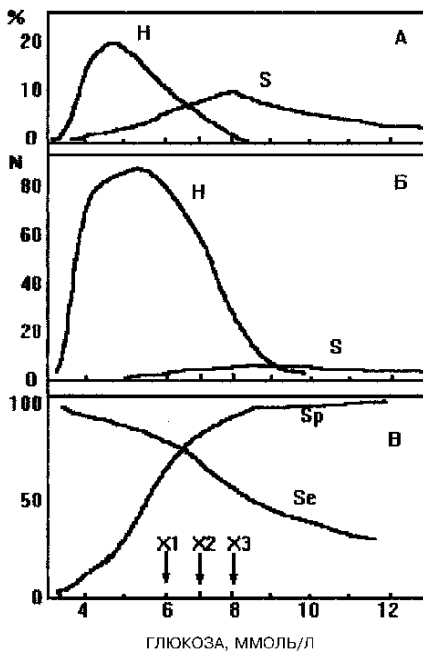


Рис. 5. Распределение концентрации глюкозы у здоровых людей (H) и больных диабетом (S) при исследовании венозной крови натощак [34, 52]:

- А - распределение обследованных, %;
- Б - распределение, в абсолютных цифрах;
- В - изменения чувствительности (Se) и специфичности (Sp) при сдвиге точки разделения (x1, x2, x3)

Обоснованность выбора “точки разделения” здоровых и больных имеет очень важное значение. Точкой разделения называют величину физиологического показателя, которая служит границей, разделяющей здоровых и больных. Данное положение хорошо демон-

стрирует рис. 5, на котором показано, что у здоровых людей и у больных диабетом концентрации глюкозы в крови значительно варьирует и широко перекрываются. Выбрав точку разделения ( $x_2$ ), можно оценить чувствительность, специфичность и прогностическую ценность положительного и отрицательного результатов.

Если точку разделения сместить в сторону более высоких концентраций ( $x_3$ ), то снизится чувствительность, но зато возрастут специфичность и прогностическая ценность положительного результата. Если точку разделения сдвинуть влево ( $x_1$ ), то повысятся чувствительность и прогностическая ценность отрицательного результата, но зато снизится специфичность и прогностическая ценность положительного результата.

Чтобы описать, как изменяется результат исследования в зависимости от выбора точки разделения, используют так называемый ROC-анализ (Receiver Operating Characteristic analysis). ROC-кривая или характеристическая кривая — это кривая зависимости чувствительности от вероятности ложноположительных результатов, т.е. величины 1 минус специфичность (рис. 6).

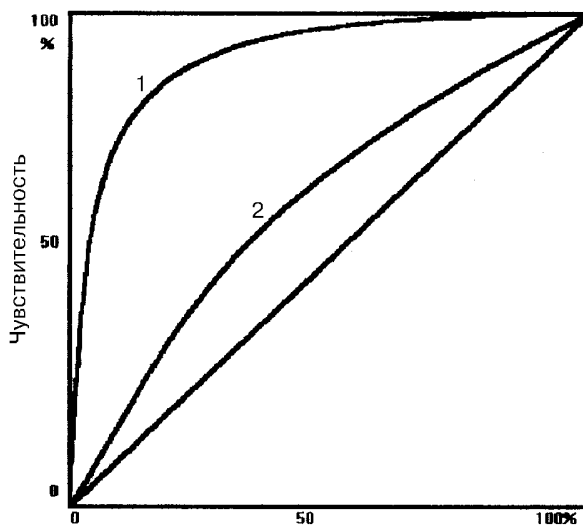


Рис. 6. Характеристические кривые для высокоинформативного (1) и низкоинформативного (2) методов диагностики



Чем ниже на графике опускается точка разделения по шкале концентрации глюкозы, тем полнее выявляются больные. Одновременно все большее число здоровых людей оцениваются как больные.

Информативность диагностического теста определяется тем, насколько высоко лежит его характеристическая кривая. Выбор точки разделения не меняет положения кривой. Происходит лишь перемещение по кривой. Наиболее простой и понятный способ ее описания — измерение площади под кривой.

*Чем ближе площадь под кривой к площади квадрата, тем эффективнее диагностический тест (кривая 1 на рис. 6).*

*Чем ближе кривая к диагонали, тем ниже эффективность диагностического метода (кривая 2 на рис. 6).*

*Диагональ соответствует полной неразличимости здоровых и больных изучаемым методом.*

*Точки, лежащие ниже диагонали, соответствуют “обратной” диагностике, т.е. ситуации, когда здоровые чаще называются больными, чем истинные больные.*

Построение характеристической кривой — наиболее современный и полный метод описания эффективности диагностического теста. Когда имеется характеристическая кривая, можно обсуждать вопрос о том, какую точку разделения выбрать.

Следует отметить, что простые способы не дают хороших результатов. Например, ошибочно стремление к минимуму числа ложных результатов. При таком подходе уравниваются ложноположительные и ложноотрицательные результаты.

Между тем, цена диагностических ошибок почти никогда не бывает одинаковой. Ложноположительный результат может угрожать, например, лишь дополнительным визитом к врачу, а ложноотрицательный — невыявлением опасной, но излечимой болезни.

Последствия ошибок и правильных решений зависят от конкретной ситуации. Поэтому выбор точки разделения через взвешивание последствий — очень ответственное дело. К сожалению, сегодня не существует универсальных рецептов для этого. Необходимо тщательно изучить, насколько обоснован выбор точки разделения в каждом случае.

Для внедрения в практику диагностического теста следует очень тщательно и точно описать подробности применения теста. Например, условия подготовки пациента, хранения и транспортировки проб. Неясность в подобных “простых” мелочах может сделать выполнение теста невозможным.

Наконец следует оценить общую полезность применения теста. Критерий эффективности диагностики — возможность выяснения истинного диагноза. Критерий полезности теста — улучшение исхода лечения.

Желательно иметь ответы на следующие вопросы:

*Каковы преимущества пациентов, которым установлен диагноз?*

*Существует ли при данном заболевании эффективное лечение?*

*Какова судьба пациентов с ложноположительным диагнозом?*

*Каковы последствия для пациентов с ложноотрицательным диагнозом?*

Несмотря на то, что в ведущих журналах рецензенты применяют к статьям описанные критерии, в большинстве публикуемых статей нет исчерпывающих ответов на все вопросы. Обусловлено это тем, что достаточные сведения о диагностическом методе накапливаются в ходе множества исследований. Кроме того, некоторые подробности могут быть сообщены в отдельных публикациях. Так, результаты оценки воспроизводимости метода иногда публикуются отдельно от результатов его оценки в сопоставлении с референсным методом. Однако в любом случае при анализе результатов научных исследований необходимо исходить из соответствующих критериев. Только так можно правильно оценивать и вносить в свою практику лучшие, проверенные методы диагностики.

## **10. Трудности внедрения научно-обоснованной медицинской практики и пути их преодоления**

На пути внедрения научно-обоснованной медицинской практики встречается немало трудностей. G.Gyatt et al. [43, 44] отмечают наиболее важные.

♦ Большинство практических врачей не владеют принципами критической оценки публикаций, их пугает сложность овладения такими навыками.

♦ Люди предпочитают быстрые и простые ответы. Медицинские руководства в виде «поваренных книг» с однозначными и простыми рецептами имеют свою привлекательность.

♦ Научно-обоснованная медицинская практика требует дополнительных затрат времени и сил, что может восприниматься врачом как *отвлечение* от основной деятельности.

♦ Существует много клинических ситуаций, для решения которых не хватает научно-обоснованных фактов. Нет достаточной мотивации, чтобы менять свои привычки.

Совершенно нереально требовать от врача общей практики каждый раз самостоятельно отыскивать и критически оценивать факты, необходимые для решения клинической ситуации, особенно в нашей стране, где ощущается острый недостаток в современной медицинской информации. Необходимо учитывать и то обстоятельство, что 80% медицинских публикаций в мире издается на английском языке. И наконец, ни один практический врач просто физически не в состоянии осваивать колоссальный объем новой информации.

Оптимальное решение проблемы - привлечь экспертов, которые, основываясь на принципах научно-доказательной медицины, подготовят информационный продукт для практических врачей в виде клинических рекомендаций, систематических обзоров, литературных дайджестов по важнейшим проблемам медицины. Такая деятельность развивается по нескольким направлениям.

*Разработка научно-доказательных клинических рекомендаций по наиболее важным медицинским проблемам.* Данный подход получил широкое распространение за рубежом за последнюю четверть века. В 90-е годы принцип подготовки клинических рекомендаций претерпел существенные изменения: *от метода экспертных оценок и консенсуса перешли к современным методам отбора и критической оценки научных фактов.* Обычно инициаторами клинических рекомендаций выступают профессиональные врачебные ассоциации или правительственные организации, образующие экспертные группы, в задачу которых входит изучение всей доступной литературы по определенному вопросу, ее критическая оценка с применением соответствующих современных принципов и форму-

лирование четких рекомендаций. Будучи внедренными в практику, хорошо составленные клинические рекомендации обеспечивают, помимо прочего, колоссальный экономический эффект.

*Формирование базы данных систематических обзоров рандомизированных контролируемых исследований.* Это направление активно развивается в рамках международного содружества ученых, поставивших своей целью выявлять, систематизировать и обобщать результаты всех когда-либо опубликованных рандомизированных контролируемых исследований. Работа эта далеко не закончена, база данных постоянно пополняется за счет вновь опубликованных рандомизированных контролируемых исследований. В ней появляются новые тематические разделы. С помощью специальной методики ученые суммируют данные, полученные в ходе разных изысканий по одной проблеме. В результате такого синтеза информации удается объективно, на основе статистических выкладок, оценить степень полезности различных лечебных, диагностических и профилактических вмешательств. Строгий научный подход к отбору и синтезу информации отличает систематические обзоры от обычных литературных обзоров, публикуемых в медицинских журналах. Последние, как правило, страдают субъективизмом, поскольку авторы не ставят своей задачей анализ и критический разбор всех исследований по проблеме, а скорее, наоборот, подбирают литературные источники таким образом, чтобы подтвердить или опровергнуть определенную точку зрения. По этой причине *научно-доказательная медицина не рассматривает обычные литературные обзоры в качестве надежных источников информации.* К настоящему времени базы данных уже распространяются на различных электронных носителях (дискетах, компактных оптических дисках), а в ближайшем будущем - и через Интернет.

*Создание журнальных клубов и реферативных дайджестов.* Подавляющее число рандомизированных контролируемых исследований, которые могут представлять интерес для практических врачей, публикуется в достаточно ограниченном круге крупнейших международных медицинских журналов, имеющих высокие тиражи и лидирующих в Science Citation Index. Как показал опыт ACP Journal Club, только 2% статей из 50 крупнейших журналов удов-

летворяют этим требованиям. Опросы читателей научных журналов показывают, что в 80% случаев прочитываются лишь их краткие резюме и редакционные комментарии по поводу этих статей. Поэтому важно уметь строго оценивать *структурированные резюме*, которые содержат те же разделы, что и оригинальные статьи - цели исследования, методы, важнейшие результаты и практические выводы.

## 11. Практическая реализация доказательной медицины - разработка формулярной системы (совместно с С.И. Кузнецовым)

Принципы доказательной медицины широко внедряются для принятия взвешенных решений о выборе того или иного лекарственного средства.

Проблема рационального использования лекарственных средств (ЛС) требует комплексного научного обоснования основных принципов ее реализации и принятия решений в методологическом и законодательном плане. В конечном счете, ее решение должно обеспечить затратно-эффективное рациональное назначение и использование лекарственных средств на основе сравнения их реального применения с установленными стандартами. В настоящее время такой процесс называют «разработкой формуляра». *Под «формуляром» понимают особый список лекарственных средств, который носит ограничительный характер и позволяет использовать только те лекарства, которые входят в этот список.* Формуляры разрабатываются отдельно для каждого типа лечебного учреждения, а на их базе создается формуляр региона (города, области). Этим достигается значительное сокращение номенклатуры используемых лекарств и повышается их терапевтическая отдача, облегчаются закупки.

Такой список составляется путем тщательного анализа заболеваемости (по каждой отдельной нозологии), уровней потребления и стоимости курса лечения каждого заболевания. В дальнейшем такой список лекарств (формуляр) может быть издан в виде справочника-формуляра, обычно карманного размера, который дает

врачам и провизорам данного лечебного учреждения базовую информацию о применении, показаниях и противопоказаниях, побочных эффектах, взаимодействии с другими лекарствами.

На уровне лечебного учреждения, подразумевается создание формулярной комиссии, которая отвечает за разработку ограниченного списка лекарственных средств, утвержденных для использования и закупок данной больницей.

В настоящее время не менее 70% лекарств на мировом рынке являются дубликатами или не жизненно важными средствами. Некоторые лекарства имеют неблагоприятное соотношение риск/терапевтический эффект. Многие новые лекарства имеют показания, не существенные для основных потребностей населения. И почти всегда они значительно дороже существующих лекарств.

***Правильно внедренная формулярная система может привести к следующим позитивным результатам:***

1. Исключение небезопасных и неэффективных лекарств.
2. Уменьшение количества наименований закупаемых лекарств и снижение общих расходов на закупки, что позволяет теми же средствами закупать большее количество безопасных и эффективных лекарственных средств.
3. Сокращение продолжительности лечения и пребывания пациентов в стационаре достигаемое исключением небезопасных и неэффективных лекарств.
4. Возможность, основываясь на наличии конечного списка используемых лекарств, проводить целевые программы повышения квалификации и собирать более полную информацию о лекарственных средствах.

В итоге, формуляр лекарственных средств является источником создания и обновления стандартов качества лечения в конкретных лечебно-профилактических учреждениях.

**Формулярный список не аналогичен перечню жизненно важных лекарственных средств. Обычно он носит ограничительный характер, в то время как перечень жизненно важных ЛС не ограничивает их использование только внесенными в него, а имеет, скорее, рекомендательный характер. Разработанные в отдельных лечебных уч-**

**реждения формулярные списки лекарственных средств становятся основой для создания регионального формуляра ЛС.**

Формулярная система облегчает закупки лекарственных средств, уменьшает расходы и оптимизирует лекарственную терапию.

Данный раздел монографии предлагает последовательный поэтапный подход к разработке и внедрению формулярной системы в лечебном учреждении. С небольшими модификациями такой подход может быть применен в любом лечебном учреждении, закупающем, хранящем и использующем лекарственные средства. Подобным же образом может быть создан и единый формуляр для группы лечебных учреждений или административных территорий - районов или областей.

Совместно с С.И. Кузнецовым (Главное управление здравоохранения Администрации Самарской области) были проанализированы годовые закупки лекарств за 1997 г. стационарами области. Анализ показал, что во всех больницах, за редким исключением, 95% расходов на медикаменты приходится на 100-120 лекарственных средств. Это косвенно подтверждает необходимость и возможность внедрения формулярной системы на территории области.

Желание внедрить формулярную систему может исходить как от медицинского персонала лечебного учреждения, так и от его администрации. Независимо от того, кто проявляет инициативу, успех внедрения формулярной системы зависит от поддержки всех работников лечебного учреждения.

Прежде всего, необходимо собрать информацию об использовании лекарств в больнице:

- ◆ годовой бюджет на лекарственные закупки;
- ◆ процентное соотношение бюджета на лекарственные закупки и общего годового бюджета больницы;
- ◆ общее количество используемых больницей лекарственных средств;
- ◆ стоимость за предыдущий год неиспользованных в связи с истечением срока годности лекарств;
- ◆ список десяти главных по стоимости используемых лекарств;
- ◆ список зафиксированных побочных эффектов;

- ♦ данные о смертности в результате использования лекарств;
- ♦ примеры дублирования лекарств.

На уровне лечебного учреждения, врачи должны осознать идею использования ограниченного числа лекарственных средств и быть готовыми изменить свои привычные модели назначения лекарств. Лицам, отвечающим за закупку ЛС для больницы, в свою очередь, необходимо переориентироваться на закупки только по формулярному списку. Рекомендуется создать формулярно-терапевтический комитет (ФТК).

Главной задачей формулярного терапевтического комитета является разработка и внедрение профессиональной стратегии отбора лекарственных средств, оценки их закупок, безопасного использования и информационного обеспечения в своем лечебном учреждении.

Необходимо определиться и в потребности образовательных программах повышения квалификации по вопросам применения лекарственных препаратов. Комитет также выполняет следующие функции:

- ♦ разрабатывает критерии оценки лекарственных препаратов при рассмотрении возможности их включения в формуляр больницы;
- ♦ является консультационной службой для медицинского персонала и администрации во всех вопросах, связанных с использованием лекарственных препаратов;
- ♦ способствует обеспечению оптимальной лекарственной терапии для всех пациентов путем разработки стандартов качества лечения;
- ♦ объективно оценивает клинические данные по новым лекарственным препаратам, предложенным для применения;
- ♦ предотвращает закупки лекарств-дубликатов;
- ♦ разрабатывает основной список лекарственных средств для закупок и применения в больнице (формуляр);
- ♦ рекомендует дополнения или исключения из списка лекарственных средств, утвержденных для применения в больнице;



- ♦ организует и планирует необходимые программы обучения для медицинского персонала по всем вопросам применения лекарственных средств;
- ♦ анализирует сообщения о возникающих побочных реакциях при применении лекарственных препаратов;
- ♦ проводит постоянные программы оценки лекарственной терапии;
- ♦ разрабатывает и издает формулярный справочник.

Решение о создании и составе ФТК принимается главным врачом больницы совместно с заместителем по лечебной части. Как правило, в состав ФТК входят:

Председатель - обычно зам. по лечебной работе больницы.

Секретарь - заведующий аптечным отделением, если оно существует, или провизор, или клинический фармаколог.

Заведующие основными больничными отделениями.

Авторитетные врачи и специалисты.

Ответственный за лекарственные закупки, если нет больничной аптеки (всего 7-11 человек с правом голоса).

При рассмотрении отдельных лекарственных препаратов, комитет приглашает на заседания необходимых специалистов.

***Внедрение формулярной системы в больницы затрудняется, если у членов ФТК отсутствует подготовка по клинической фармакологии.***

В связи с этим, может потребоваться направление некоторых членов формулярного комитета на курсы повышения квалификации по клинической фармакологии.

ФТК решает следующие вопросы:

1. Критерии отбора лекарственных препаратов для формуляра.
2. Внесение и исключение лекарств из формуляра.
3. Требования к заполнению листов назначений.
4. Использование не формулярных лекарств.
5. Мониторинг побочных эффектов.
6. Оценка использования лекарств.
7. Применение лекарств, находящихся на клинических испытаниях.

Отношения сотрудников больницы и представителей фармацевтических компаний и дистрибьюторов.

При рассмотрении лекарств для включения в формуляр должны учитываться следующие обстоятельства:

- обоснованная потребность в данном препарате;
- отсутствие в больничном формуляре другого лекарства, удовлетворяющего ту же потребность;
- доступность информации о проведенных клинических испытаниях данного лекарства;
- не допускается использование комбинированных лекарств кроме случаев, когда они имеют явное терапевтическое преимущество перед монопрепаратом;
- стоимость лекарства должна быть оправдана его эффективностью;
- доступность для закупок у поставщиков;
- заведующие отделениями должны получать информацию о том, какое лекарственное средство в настоящее время рассматривается для включения в формуляр или исключения из него с тем, чтобы они могли высказать свое мнение;
- формулярное лекарство (лекарства), уступающее новому, исключается из формуляра, если оно не используется по другим показаниям.

Одной из самых важных концепций во внедрении формулярной системы является требование отбора лекарств и внесения их в формуляр под **генерическим наименованием**. Каждое лекарство имеет химическое наименование и международное непатентованное наименование, или генерическое. Генерическое наименование является официальным, независимым от производителя. Коммерческое, или фирменное наименование (например, Nolicin®) дается производителем или дистрибьютором для того, чтобы потребитель отличал их продукцию от подобной других компаний. Использование генерических наименований в формулярных списках лекарств дает следующие преимущества:

- генерические наименования более информативны, чем фирменные, и отражают принадлежность к определенному классу и химической структуре;

- назначение генерических препаратов позволяет больничному провизору проводить генерическую замену, тогда как фирменная номенклатура обязывает отпускать только указанное в рецепте средство;

- генерические наименования позволяют закупать препараты у различных поставщиков и, как правило, дешевле фирменных;

- использование фирменных названий считается обоснованным, если биодоступность и биоэквивалентность лекарственных форм, выпускаемых различными фирмами варьируется столь существенно, что может привести к изменению терапевтического действия.

Это относится в первую очередь к сердечным гликозидам, противосудорожным, антиаритмическим средствам, гормонам, антикоагулянтам и другим препаратам с узким терапевтическим индексом.

Другой обоснованной причиной внесения фирменных наименований в формуляр является использование препаратов пролонгированного действия. В этом случае важно точно обозначить их фирменное наименование. Формулярный терапевтический комитет должен выбрать классификацию лекарственных средств. Лекарственные препараты можно классифицировать по следующим принципам:

- *по терапевтическому применению* - например, антиангинальные препараты, противоопухолевые средства, гипотензивные средства, антибиотики;

- *по фармакологическому действию* - например, антагонисты кальция, диуретики, вазодилататоры, антикоагулянты;

- *по химическому строению* - например, сердечные гликозиды, алкалоиды, стероиды, бензодиазепины, фторхинолоны, цефалоспорины;

- *по нозологическому принципу*, что наиболее удобно для клинициста - например, средства для лечения бронхиальной астмы, стенокардии, артериальной гипертензии, язвенной болезни. Единой классификации, принятой всеми странами в мире, не существует. В некоторых странах, например в Великобритании, Национальный формуляр организован по заболеваниям или по органам.

В формулярах США чаще применяют классификацию лекарственных средств по фармакотерапевтическому действию. Скандинавские страны используют комбинированную систему классификации по анатомическим, терапевтическим и химическим критериям.

*Для больниц рекомендуются классификации, основанные на терапевтическом применении препаратов, доступным для понимания медицинских работников среднего звена, провизоров, фармацевтов и немедицинского персонала, занимающегося лекарственными закупками. В настоящее время в мире все чаще обращаются к классификационной схеме, использованной ВОЗ при составлении примерного перечня жизненно важных лекарственных средств.*

Принятие данной классификации, таким образом, значительно упрощает сравнение цен и заказ лекарств для национальных программ. В любом случае каждое лечебно-профилактическое учреждение само выбирает наиболее приемлемый тип классификации. Для начала работы по отбору лекарственных средств в больничный формуляр необходимо собрать и проанализировать следующую информацию:

### ***1. Данные о структуре заболеваемости***

Больничный формуляр лекарственных средств должен отвечать структуре заболеваемости обслуживаемого населения. Поэтому работа по отбору лекарств в формуляр предваряется сбором и анализом статистических данных о преобладающих заболеваниях и категориях пациентов. Результатом такого анализа может стать, например, список из 50 основных нозологий или причин госпитализации с количеством случаев по каждой отдельной нозологии за определенный период времени. Качество информации будет зависеть от точности собираемых статистических данных.

### ***2. Наличие информации о лекарственных средствах***

Нельзя правильно отобрать лекарственные средства для своего формуляра без надежной объективной информации, помогающей исключить из пользования лекарства сомнительной эффективности. Лечебные учреждения, принимающие решение о внедрении у себя формулярной системы, должны непредвзято оценить всю имеющуюся в их распоряжении информационную литературу по ле-

карственным средствам и пользоваться только объективными источниками.

**3. Перечень лекарственных средств,  
закупленных и использованных лечебным учреждением  
в течение предыдущего года**

Аптечное отделение больницы или отдел закупок должны предоставить полный список лекарственных средств, закупленных лечебным учреждением в течение прошедшего календарного или финансового года, или за *последние* 12 месяцев. Предпочтительно иметь данные именно за год, так как это помогает избежать искажений информации в связи с сезонными колебаниями, но при необходимости можно использовать данные и за более короткий промежуток времени. Для проведения анализа необходимо следующее: наименование ЛС, дозировка, лекарственная форма, закупочная стоимость и использованное за исследуемый период количество.

Затем необходимо проанализировать структуру заболеваемости для того, чтобы понять, удовлетворяются ли терапевтические нужды пациентов и правильно ли закупаются и используются лекарственные средства.

Анализ проводится следующим образом:

- собрать данные о заболеваемости (в идеале, 50 основных нозологий) в порядке убывания числа случаев по каждой нозологии. Эта информация будет решающей при выборе приоритетов в процессе анализа групп лекарственных средств;

- подсчитать процент каждого из 50 заболеваний по отношению к общему числу заболеваний. Не следует рассматривать заболевания, не требующие лекарственной терапии. Например, миопия может занимать до 20% случаев у пациентов с нервными расстройствами или расстройствами органов чувств, но необходимое лечение может заключаться лишь в простой проверке зрения и подборе очков.

Используя сведения о лекарственных закупках, необходимо выбрать лекарства (и данные об их стоимости), которые были использованы для лечения 50 главных заболеваний, а затем подсчитать процент их стоимости по отношению к стоимости всех лекарственных закупок за исследуемый период. Сравнение двух полу-

ченных списков даст представление о соответствии расходов на лекарственную терапию потребностям лечения. После такого анализа необходимо установить, на какие конкретно лекарственные средства тратится большая часть бюджета больницы. Проводится, так называемый ABC и VEN анализ. *ABC анализом называется метод распределения лекарственных средств по трем группам в соответствии с их годовым потреблением (стоимость единицы отпуска умноженная по годовое использование).*

**Класс А:** 10-20% препаратов, на которые расходуется 70-80% бюджета больницы на лекарственные средства.

**Класс В:** средний уровень потребления.

**Класс С:** большинство препаратов с низкой частотой использования, на которые в сумме расходуется не более 25% лекарственного бюджета.

ABC-анализ позволяет получить точную объективную картину расходования средств лекарственного бюджета лечебного учреждения. Параллельно с ABC-анализом проводится VEN-анализ, который позволяет устанавливать приоритеты отбора и закупок лекарственных препаратов в соответствии их классификацией на жизненно важные (Vital), необходимые (Essential), и второстепенные (Non-essential):

Жизненно важные (Vital)	Лекарства, важные для спасения жизни, имеющие опасный для жизни синдром отмены, постоянно необходимые для поддержания жизни
Необходимые (Essential)	Лекарства, эффективные при лечении менее опасных, но серьезных заболеваний
Второстепенные (Non-essential)	Лекарства для лечения легких заболеваний, лекарства сомнительной эффективности, дорогостоящие лекарства с симптоматическими показаниями

Использование обеих методик дает необходимые данные для принятия решений по исключению лекарств из формуляра больницы, включению в него, и сведения о тех лекарствах, которые используются либо недостаточно, либо избыточно.

Затем необходимо провести анализ терапевтических классов лекарственных средств и отобрать лекарства для формулярного списка. Это наиболее важный шаг в процессе разработки формуляра, так как именно на этом этапе создается список лекарственных средств для применения в больнице. Анализ терапевтических классов оказывает положительное воздействие как на терапевтическую, так и на экономическую сторону лечебного процесса. Терапевтические аспекты включают в себя повышение качества оказываемой пациенту помощи благодаря отказу от применения небезопасных и/или малоэффективных препаратов. В дополнение к этому, результаты данного шага повлияют и на стоимость лечения, что позволит создать основу для более экономного расходования бюджета больницы. Важно тщательно спланировать и провести анализ терапевтических классов лекарственных средств, а также выделить достаточно времени для его проведения. На данном этапе должно быть принято решение, по какому пути пойдет разработка формулярного списка.

Есть две возможности: ФТК может принять решение считать все используемые на настоящий момент в больнице лекарственные средства первоначальным формуляром. В процессе анализа лекарств, часть из них будет исключена из списка или будут внесены новые. Это наиболее распространенный подход. Формуляр может создаваться по мере последовательного рассмотрения каждого отдельного терапевтического класса лекарственных средств.

По завершении анализа первого терапевтического класса формуляр будет содержать только препараты, входящие в этот класс. И сразу же можно будет начать закупку и назначение именно этих выбранных препаратов. Остальные лекарственные средства вносятся в формуляр по мере рассмотрения последующих терапевтических классов. Полученная в результате проведения ABC/VEN анализа информация может быть использована для составления календарного плана рассмотрения терапевтических классов лекарств. Например, если получены данные о том, что на антибиотики приходится 30% расходов лекарственного бюджета больницы, ФТК может принять решение об анализе этой группы лекарств в первую очередь.

Приоритетными для первоочередного анализа будут и группы лекарств, создававшие в прошлом проблемы. Может оказаться сложным сразу же составить полное расписание анализа лекарственных препаратов прежде, чем будет проведено рассмотрение нескольких приоритетных классов. После этого легче определить, сколько реально требуется времени для анализа. Окончательное решение будет определяться количеством анализируемых лекарств и возможностями Комитета подготовить материал, необходимый для анализа. Основные вопросы, ответы на которые послужат основанием для включения лекарственного препарата в формуляр следующие:

- ◆ Существует ли обоснованная потребность в данном препарате?
- ◆ Удовлетворена ли эта потребность другим препаратом, который уже применяется в данном лечебном учреждении?
- ◆ Может ли эта потребность удовлетворяться более безопасным и эффективным путем с помощью других средств, уже включенных в формуляр?
- ◆ Проводились ли надлежащие клинические испытания данного препарата?
- ◆ Оправдывает ли применение данного лекарственного средства расходы на его закупку и затраты, связанные с его применением?
- ◆ Будет ли данный препарат назначаться также и амбулаторным пациентам, и если да, то есть ли гарантии возмещения затрат на его приобретение по условиям медицинского страхования или в рамках программы льготного обеспечения пациентов лекарственными средствами?

Для лекарственных препаратов, которые единичны в своем терапевтическом классе, необходимо четко определить, что именно эти препараты являются затратно-эффективным методом лечения и применяются по широкому спектру показаний.



### **11.1. Анализ терапевтических классов лекарственных средств**

Рекомендуется подготовить краткую оценочную справку по анализируемому лекарственному препарату. Обычно приводятся такие данные:

- ◆ К какому терапевтическому классу относится рассматриваемый лекарственный препарат в соответствии с принятой классификацией.

- ◆ Генерическое наименование - официально утвержденное наименование, а если препарат является комбинированным - отмечаются все активные составляющие под генерическими наименованиями.

- ◆ Торговое (ые) наименование (ия) - может быть значимо, когда препараты поступают от различных производителей. Происхождение препарата - название и адрес производителя (лей) данного препарата.

- ◆ Показание (ия) - убедитесь, что формулярные препараты охватывают все нозологии, встречающиеся в данном лечебном учреждении.

- ◆ Противопоказание (ия) - противопоказания могут служить основанием для исключения препарата из формуляра.

- ◆ Эффективность/Фармакология - при сравнении лекарственных средств с идентичными показаниями к применению, решение принимается на основе данных об эффективности препаратов.

- ◆ Побочные эффекты - желательно включать в формуляр лекарственные препараты с незначительными побочными эффектами и предоставить данные о частоте возникновения, тяжести побочных эффектов препарата.

- ◆ Схемы применения - существует тенденция к использованию лекарственных средств, которые можно применять один-два раза в день. Особенно это имеет значение для лекарственных средств, предназначенных для внутривенного введения.

- ◆ Продолжительность лечения - менее длительный курс терапии позволяет снизить расходы на закупку препарата и сократить срок пребывания пациента в больнице.

♦ Пути введения препарата - пероральные препараты зачастую дешевле инъекционных, как по закупочной стоимости, так и по затратам на их применение.

♦ Фармакокинетические характеристики. При рассмотрении препаратов важными факторами являются отличительные свойства метаболизма, возрастные особенности [52], пути элиминации.

♦ В разделе о путях элиминации необходимо включить данные о дозировке лекарства для пациентов с нарушенной функцией почек.

♦ Взаимодействие лекарств между собой и с пищевыми продуктами - предпочтение следует отдавать препаратам, в меньшей степени вступающим во взаимодействие.

♦ Идентичные препараты, которые могут заменить рассматриваемый лекарственный препарат.

♦ Составляется список всех терапевтически схожих лекарств, которые могут быть исключены из формуляра, если данный рассматриваемый препарат в него включен.

Далее приводятся возможные ограничения по использованию препарата:

*Ограничения по диагнозу* - определяют показания для правильного назначения лекарственных средств в конкретном лечебно-профилактическом учреждении. Может быть ситуация, когда использование по другим показаниям выходит за рамки утвержденных критериев по диагнозу.

*Ограничения по уровню квалификации* - определяют специалистов, имеющих право использовать данные формулярные лекарственные средства или терапевтические классы. Например, назначение некоторых инъекционных антибиотиков разрешается только инфекционисту, назначение тромболитических лекарственных средств допускается только кардиологами или реаниматологами.

*Фармакологические ограничения* - определяют утвержденные дозы, кратность приема, продолжительность курса лечения и другие аспекты, являющиеся специфичными для данного формулярного лекарственного препарата.

## **11.2. Экономические аспекты анализа терапевтических классов лекарственных средств**

Раньше экономические вопросы для принятия решений по формуляру в основном сводились к рассмотрению закупочной цены лекарственного препарата. Но с течением времени экономический анализ расширился и теперь включает в себя рассмотрение тех расходов и/или затрат, связанных с применением лекарственных препаратов, которые не столь очевидны. Подобные затраты являются частью общих расходов лечебного учреждения на лекарственную терапию, поэтому они должны обязательно учитываться. Это расширяет анализ затрат, связанных с проведением лекарственной терапии. Экономические аспекты лекарственной терапии включают в себя выявление, расчет и сравнение всех затрат и последствий при их применении (как позитивных, так и негативных). Решения по формулярным лекарствам должны учитывать различные виды расходов:

- стоимость лекарственного препарата на полный курс терапии;
- затраты на применение препарата, что включает в себя стоимость материалов, например, контейнеров или бутылок для внутривенного вливания, растворов, шприцев;
- стоимость лекарственных препаратов, которые необходимы для совместного применения с данным препаратом, например, лекарства для премедикации, а также и стоимость предметов медицинского назначения;
- затраты на закупку и введение лекарственных средств, используемых для лечения возможных побочных эффектов и осложнений;
- затраты, связанные с лабораторными исследованиями, включая оборудование для мониторинга и реактивы;
- затраты, связанные с хранением, если для препарата требуются особые условия хранения;
- возможное влияние на продолжительность пребывания в больнице.

Несмотря на то, что достаточно сложно просчитать финансовые затраты, связанные с возможной зависимостью воздействия препарата на продолжительность пребывания пациента в больнице, экономический анализ должен включать подобный расчет.

*Целью формуляра лекарственных средств не должно быть простое сокращение затрат на закупку препаратов, а сокращение затрат в целом, необходимых для лечения определенных заболеваний.*

Было доказано, что в некоторых случаях намного выгоднее применять более дорогостоящий лекарственный препарат. И подобное утверждение верно при условии, что прием данного препарата сократит использование параллельных лекарственных средств, снизит частоту возникновения побочных эффектов и/или затраты на оплату медицинского персонала, продолжительность пребывания пациента в больнице и/или частоту случаев повторной госпитализации.

Необходимо ознакомить медицинский персонал лечебного учреждения с возможностью регулирования и использования неформулярных лекарственных препаратов, включения и исключения лекарств из формулярного списка, осуществления генерических и терапевтических замен.

#### ***Использование неформулярных препаратов***

Как правило, для лечения пациентов в больнице разрешается использовать только вошедшие в формулярный список препараты. Тем не менее, терапевтические потребности отдельных больных не всегда могут быть удовлетворены лекарствами утвержденного формуляра. Для таких больных устанавливается регулирование использования неформулярных препаратов. Запрос на использование неформулярного препарата всегда соотносится с индивидуальным назначением. Если обоснование лечащего врача признается достаточно веским, больница закупает необходимое количество препарата для пациента. На регулярной основе проводится анализ всех поступивших запросов на неформулярные препараты. Если в результате такого анализа выявляются частые запросы определенного препарата и выясняется, что он превосходит существующий формулярный препарат, то может быть принято решение о его включении в формуляр и исключении старого лекарства.

Если, напротив, использование неформулярного препарата необоснованно, то это потребует разработки специальных учебных программ для врачей или направления отдельных врачей на курсы по клинической фармакологии.

### **Генерические и терапевтические замены**

*Генерической* называют замену биоэквивалентных лекарственных препаратов, содержащих активное одинаковое вещество химически идентичное по силе, концентрации, дозировке, лекарственной форме и путям введения. Например, замена одного наименования *верапамила* на другое является *генерической*, если количество активного вещества, концентрация и лекарственная форма идентичны. Замена свиного инсулина на человеческий не является генерической, как не является генерической и замена препарата быстрого действия на препарат пролонгированного действия.

*Терапевтической* называют замену препаратов одинакового терапевтического действия, не являющихся генерическими аналогами.

Примером *терапевтической замены* может быть замена *цефрадина* на *цефалексин*. В этом случае оба препарата обладают одинаковым спектром действия, частотой употребления, лекарственной формой и путями введения. Можно считать терапевтически эквивалентными и препараты с разной концентрацией активного вещества. Например, *циметидин* 400 мг два раза в день часто считается терапевтически эквивалентным *ранитидину* 150 мг два раза в день, или *фамотидину* 40 мг один раз в день. Следует отметить, что официально признанных терапевтических эквивалентов не существует. Решение должно приниматься самостоятельно каждым ФТК отдельной больницы. Другими примерами препаратов, часто рассматриваемых как терапевтически эквивалентные, являются: *фуросемид* 40 мг и *этакриновая кислота* 50 мг, *преднизолон* 5 мг и *триамцинолон* 4 мг, *пропранолол* 160 мг и *атенолол* 100 мг, *ампициллин* 250 мг и *амоксциллин* 250 мг.

**В дальнейшем следует сделать выбор между использованием в ЛПУ только формулярного списка или разработкой формулярного справочника.** Врачи не станут пользоваться формулярным справочником, если поиск необходимой информации затруднен.

Применение различных типов индексирования значительно упрощает пользование формулярным справочником:

***Перекрестное индексирование генерических наименований и синонимов.*** Данный индекс должен быть организован в алфавитном

порядке и содержать генерические и наиболее распространенные фирменные наименования. Он позволит врачу быстро найти сведения по генерическому или фирменному наименованию. Перечисляются все лекарственные препараты по терапевтическим классам. Это удобно при выборе терапевтических альтернатив при определенных заболеваниях, например, при аллергии на отдельное лекарственное средство.

**Индексирование по показаниям.** При таком индексировании в алфавитном порядке перечислены заболевания с перечнем лекарственных препаратов для их терапии. Такой индекс нужен врачу, если он хочет узнать, какие препараты внесены в формулярный список для лечения данного заболевания.

Формулярный процесс не заканчивается созданием и распространением в лечебном учреждении формулярного справочника. В дальнейшем необходимо провести следующие этапы:

- ♦ разработку и внедрение стандартов качества лечения;
- ♦ внедрение программы оптимального использования лекарственных средств (ОИЛС);
- ♦ мониторинг побочных реакций;
- ♦ обновление формулярного списка и формулярного справочника.

**Стандартами качества лечения** называют схемы лечения распространенных заболеваний и синдромов, разработанные и утвержденные ведущими специалистами для достижения максимального терапевтического эффекта наименее затратным способом. Стандарты качества лечения включают рекомендации по лекарственной терапии, включая лекарства выбора, указания по их применению и возможные альтернативы. Стандарты также должны включать указания по нелекарственной терапии, например, хирургическому вмешательству, указания по диетам, физиотерапии и лабораторному мониторингу. Стандарты качества лечения обеспечивают получение пациентами одинакового уровня лечения и соблюдение применения формулярных лекарственных препаратов. Очевидно, что не следует включать в стандарты качества лечения лекарственные средства, не внесенные в формулярный список. Стандарты качества лечения могут составить отдельный раздел форму-

лярного справочника. Для обеспечения соблюдения стандартов лечения могут потребоваться специальные образовательные программы по клинической фармакологии, поскольку сам по себе формулярный список не гарантирует правильного использования лекарственных средств. В дальнейшем формулярный список может составить основу ОИЛС.

Внедрение программы ОИЛС позволяет выявить нерациональную и неоправданно дорогостоящую лекарственную терапию.

Одной из задач внедрения программы является мониторинг нежелательных реакций организма на введение лекарственных средств. Могут быть следующие их разновидности:

1. *Определенная* - когда реакция возникает через четко выраженный промежуток времени после приема лекарства, и возобновляется при его повторном применении.

2. *Вероятная* - когда реакция развилась после приема препарата и не могла быть вызвана другим лекарственным средством или патологическим состоянием больного, а также ситуация, когда проявления реакции исчезают после отмены препарата.

3. *Возможная* - когда реакция проявляется без четкой связи с приемом препарата и может иметь другую этиологию, не связанную с действием лекарственного вещества.

4. *Сомнительная* - реакция, вызванная, скорее всего, другой причиной. Программа ОИЛС позволяет более профессионально и рационально относиться к включению и исключению лекарств.

При обновлении формулярного справочника желательно ориентироваться на систематические обзоры и различать такое понятие, как мета-анализ.

## 12. Систематические обзоры и мета-анализ

В настоящее время в мире ежегодно публикуется около 2 млн научных медицинских статей, не считая книг и материалов многочисленных конференций. Необходимость синтеза информации, представленной в виде обзора литературы по той или иной проблеме очевидна.

В этой ситуации чисто описательный подход к синтезу информации, имеющий место в несистематических, или качественных обзорах литературы не всегда оправдан. Такие обзоры литературы трудно воспроизводимы и часто отражают лишь субъективное мнение их авторов. Эта проблема нередко усугубляется различными или даже противоречивыми результатами некоторых клинических исследований, недостаточным числом больных, участвующих в рандомизированных контролируемых исследованиях. В таких исследованиях возникает большая вероятность  $\beta$ -ошибки (ложное утверждение об отсутствии значимого различия между эффективностью сравниваемых методов лечения, когда на самом деле они различаются). Классическим примером расхождения результатов описательного обзора литературы и ее систематического анализа (количественного анализа, или мета-анализа) является публикация вывода лауреата Нобелевской премии профессора Л. Полинга о снижении простудных заболеваний в результате применения аскорбиновой кислоты [53].

Впоследствии этот вывод не был подтвержден результатами соответствующего мета-анализа [54].

Мета-анализ, появившейся в конце 80-х годов, сегодня относится к одной из самых популярных и быстро развивающихся методик системной интеграции результатов отдельных научных исследований. В мета-анализе для объединения и обобщения результатов нескольких оригинальных исследований применяются статистические методы.

Мета-анализ можно характеризовать как новый научный метод обобщения количественных данных однородных исследований для получения суммарных статистических показателей. Его целесообразно применять при проведении систематического обзора литературы для интегральной оценки эффективности того или иного метода лечения, профилактики и диагностики, а также для получения суммарных показателей заболеваемости, смертности и факторов риска развития того или иного заболевания. *Составление на основе мета-анализа систематических обзоров литературы считают разновидностью научной деятельности* [55].



Мета-анализ определен как "количественный систематический обзор литературы ... или ... количественный синтез первичных данных с целью получения суммарных статистических показателей" [56]. Это методика "объединения результатов различных исследований, складывающаяся из качественного компонента (например, использование таких заранее определенных критериев включения в анализ, как полнота данных, отсутствие явных недостатков в организации исследования) и количественного компонента (статистическая обработка имеющихся данных)".

Кроме установления эффективности того или иного лечебного воздействия, мета-анализ используется для выяснения причинно-следственных связей между заболеванием и факторами риска, определения побочных эффектов изучаемого вмешательства и выявления прогностически значимых факторов развития того или иного исхода заболевания. Результаты постоянно обновляемых мета-анализов могут помочь организаторам здравоохранения и экспертам в выработке рекомендаций и подготовке законодательных актов, касающихся использования определенных диагностических и лечебных методов, а врачу получить максимально объективную информацию о результатах исследований в интересующих его областях. Однако необходимо помнить, что неправильно проведенный мета-анализ может привести и к ошибочным результатам.

Поскольку процесс подготовки обзора сам по себе (как и любое другое научное исследование) подвержен систематическим ошибкам, обзор ценен только тогда, когда в нем детально излагаются методы его подготовки. Общеизвестны следующие основные характеристики, используемые при отборе клинических исследований для включения в мета-анализ:

1. Четкие принципы критерий включения и исключения больных.
2. Место проведения исследования (университетская клиника, стационар общего типа, поликлиника, общая популяция).
3. Продолжительность.
4. Основные характеристики больных, включенных в исследование, дополнительное лечение и наличие сопутствующих заболеваний.

5. Диагностические критерии заболевания, являющегося предметом исследования в мета-анализе.

6. Тип, доза, частота применения, путь введения, время начала использования лекарственного препарата (или метода лечения), а также продолжительность лечения.

7. Отклонения от протокола исследования (если они имелись).

8. Клинические исходы и критерии их оценки.

9. Наличие в материалах исследования абсолютных величин, характеризующих число включенных в него больных и клинические исходы.

Исследования, включенные в мета-анализ, должны быть максимально однородными, т.е. вид вмешательства (лечение), популяция (включенные больные) и изучаемые исходы во всех этих исследованиях должны быть сходными. Нельзя объединять исследования, если в них применялось различное лечение, оценивались несопоставимые клинические исходы или существенно различалась структура их проведения (например, данные популяционных исследований нельзя объединять с данными, полученными в условиях стационара). Во всех оригинальных исследованиях должны быть приведены сведения о факторах, имеющих существенное значение для изучаемого исхода (так называемые прогностические факторы). Авторы мета-анализа должны иметь возможность учитывать их при статистической обработке данных. Такая информация также должна быть доступна и читателям мета-анализа, чтобы они могли судить о достоверности и применимости его результатов.

Прежде всего отметить, что в основе высококачественного систематического обзора лежит правильно сформулированный клинический вопрос, на который можно получить ответ. Формулировка вопроса определяет, какие исследования будут включены в обзор, каковы будут методы поиска оригинальных исследований и какие данные необходимо отобрать из этих исследований. При плохо сформулированном вопросе обзор вряд ли будет достаточно информативным. Кроме того, четко сформулированный вопрос позволяет читателю быстро понять, применимы ли результаты обзора в его клинической практике.

Оценка отобранных для обзора исследований преследует три цели:

- 1) сделать вывод о достоверности исследований;
- 2) выявить причины различий в результатах исследований;
- 3) обеспечить читателей достаточной информацией для принятия решения о том, применимы ли результаты систематического обзора в обычной клинической практике.

Систематические обзоры объективно обобщают большое количество информации, определяют пробелы в медицинских научных исследованиях и выявляют пользу или вред вмешательства. Руководители здравоохранения, распространяя знания среди врачей о существовании таких обзоров и обучая критически оценивать их результаты, могут рассчитывать на поддержку и концептуально принимать решения о видах медицинской помощи, которую следует предоставлять. Кроме того, участие руководителей здравоохранения в организации, проведении и публикации систематических обзоров поможет сделать эти обзоры более полезными и понятными для целевых аудиторий.

Вместе с этим, следует помнить, что все научные исследования подвержены влиянию предубеждений исследователей и специфических интересов финансирующих организаций [57].

Как и все прочие научные исследования, систематические обзоры варьируют по методологическому качеству. Обзоры, содержащие систематические ошибки, могут привести к потенциально опасным результатам [58, 59].

Полное раскрытие источников финансирования и учет возможности наличия конфликта интересов у авторов помогает потребителям медицинских услуг и руководителям здравоохранения отобрать наиболее взвешенные систематические обзоры [60].

Поэтому имеет смысл ограничить источники информации наиболее надежно проведенными систематическими обзорами. Самое простое — *научиться отличать систематические обзоры от обычных обзоров литературы* [61].

Основные отличия традиционные обзоры литературы от систематических обзоров приведены в табл. 5.

Таблица 5

**Различия между обзорами литературы и систематическими обзорами**

Характеристика	Обзор литературы	Систематический обзор
Освещаемые вопросы	Часто рассматривается широкий спектр вопросов	Часто посвящен определенному клиническому вопросу
Источники данных и стратегия поиска	Источники не всегда указаны, Стратегия может быть ошибочной	Источники обычно всеобъемлющи, а стратегия поиска точно изложена
Принцип отбора данных	Не всегда указан, может быть ошибочным	Отбор основан на определенных критериях, применяемых одинаковым образом
Методы оценки данных	Различные	Строгие, критические методы оценки
Обобщение данных	Часто качественное	Количественное (мета-анализ)
Выводы	Иногда научно-обоснованные	Как правило, научно обоснованные

Все обзоры относятся к наблюдательным ретроспективным исследованиям и поэтому подвержены систематическим и случайным ошибкам. Соответственно качество обзора и, следовательно, его ценность зависят от использования при его проведении научных методов, снижающих до минимума эти ошибки.

Систематические обзоры проводят, чтобы найти ответы на определенные, часто довольно узкие клинические вопросы, в точной формулировке которых должны быть отражены 4 параметра:

- определенная популяция и клиническая база (например, пожилые амбулаторные больные);
- изучаемое заболевание (например, артериальная гипертония);

- используемый метод исследования или лечения (например, определенная лекарственная терапия);
- один или более определенный клинический исход (например, сердечно-сосудистая и цереброваскулярная заболеваемость и смертность).

Условно различают следующие стадии использования систематических обзоров:

- осведомленность о существовании систематических обзоров; осознание преимуществ и сложностей их использования;
- выявление нужных обзоров;
- критическая их оценка;
- участие в создании систематических обзоров и включение их в процесс принятия решений.

Множество факторов свидетельствует о том, что принятие решений по проблемам здравоохранения - это сложный процесс, в котором информации отводится лишь весьма скромная и определенная роль [63].

К настоящему времени исследований, где оценивалось бы непосредственное влияние систематических обзоров на процесс принятия решений проведено недостаточно.

Можно полагать, что вовлечение заинтересованных людей в исследовательский процесс позволит подготавливать систематические обзоры, которые отвечают на более актуальные для потребителей медицинских услуг и руководителей здравоохранения вопросы. Научные исследования с участием потребителей медицинских услуг и руководителей здравоохранения более прагматичны и учитывают местные приоритеты, процессы и перспективы.

Таким образом, систематические обзоры в перспективе могут быть очень ценным инструментом принятия решений, поскольку они объективно обобщают большие объемы информации, выявляют пробелы в научных исследованиях, определяя пользу или вред медицинских вмешательств. Использование систематических обзоров как источника информации полезно для принятия клинических решений в повседневной практической работе, планирования будущих исследований и выработки политики здравоохранения.

## СЛОВАРЬ ТЕРМИНОВ [ 62 ]

*Альфа-ошибка* ( $\alpha$ -error) или *ошибка I рода* (Type I error) - вероятность ошибочного отклонения нулевой гипотезы. В клинических исследованиях это вероятность того, что одно лечение будет считаться эффективнее другого, когда в действительности это различие отсутствует.

*Анализ в зависимости от назначенного лечения* (intention to treat analysis) - способ анализа данных в контролируемом испытании. Анализ проводится в зависимости от того, к какой группе - экспериментальной или контрольной - был отнесен пациент при рандомизации независимо от того, получил он в действительности исследуемую терапию или нет.

*Анализ дожития* (survival analysis, time-to-event analysis) - статистические методы анализа времени, прошедшего от заданного начального момента до момента наступления определенного исхода (обычно - смерти). Метод позволяет описать любой исход, дихотомический по своему характеру и имеющий место лишь однажды в процессе наблюдения, например, время до развития инфаркта миокарда или до рецидива злокачественного новообразования. К способам анализа времени наступления события относятся анализ Каплана-Мейера (Kaplan-Meier analysis), регрессионная модель пропорционального риска Кокса (Cox proportional hazards regression model), лог-ранк тест (log-rank test) и другие.

*Анализ решений* (decision analysis) - статистический подход к принятию решений, осуществляемый путем сопоставления альтернатив с учетом экономических затрат и последствий для больного в заданных условиях.

*Анализ чувствительности* (sensitivity analysis) - метод, используемый в мета-анализе, фармакоэкономике и анализе решений для оценки влияния различных параметров на конечный результат. Анализ «*наилучший вариант - наихудший вариант*» (best case / worst case analysis) представляет собой специальный вид анализа чувствительности.

*Бета-ошибка* ( $\beta$ -error), или *ошибка II рода* (Type II error) - вероятность ошибочного принятия нулевой гипотезы. В клинических исследованиях это вероятность того, что оба лечения будут считаться равно эффективными, когда в действительности одно лечение лучше, чем другое.

*Величина  $p$*  ( $p$  value) - вероятность того, что полученный результат абсолютно случаен. Величина  $p$  может изменяться от 1 (результат точно случаен) до 0 (результат точно не случаен). Величина  $p$ , меньшая или равная заданному уровню альфа-ошибки (например, 0,05), свидетельствует о статистической значимости полученного различия. В то же время, она не характеризует клиническую значимость результата.

*Величина  $p_\alpha$*  - вероятность альфа-ошибки. Обычно, когда пишут  $p$ , имеют в виду именно  $p_\alpha$ .

*Величина  $p_\beta$*  - вероятность бета-ошибки.

*Вероятностная выборка* (probability sample) - это такая выборка, вероятность попадания в которую для каждого индивидуума известна заранее.

*Вмешивающийся фактор, или конфаундер, конфаундинг-фактор, неучтенный фактор* (confounding factor) - фактор, связанный с известным исследуемым фактором, и некоторым образом влияющий на результат (исход). Например, известно, что у мужчин чаще, чем у женщин возникает ИБС. Однако это может быть связано не с исследуемым фактором (пол), а, допустим, с тем, что мужчины чаще курят, в значительной степени подвергаются стрессам, потребляют больше пищи, богатой холестерином. Влияние вмешивающихся факторов приводит к возникновению систематической ошибки (confounding bias).

*Выборка* (sample) - это часть популяции, полученная путем отбора. Клинические исследования обычно выполняются на выборках. Оценку характеристик популяции на практике приходится осуществлять путем оценки этих характеристик по выборке.

*Выборка вероятностная* (probability sample) - выборка, вероятность попадания в которую для каждого индивидуума известна заранее.

*Выборка случайная* (random sample) - такая выборка, вероятность попадания в которую для каждого индивидуума в популяции одинакова.

*Выборка смещенная* (biased sample) - это такая выборка, которая вследствие систематической ошибки отличается от популяции, представляющей предмет исследования, или от популяции, по отношению к которой должны применяться результаты исследования. Например, лица, добровольно участвующие в программах профилактики заболеваний, отличаются от прочих людей в популяции и представляют собою смещенную выборку из популяции.

*Выборки, размер* (sample size) - число больных, которое необходимо включить в исследование для невелирования случайности в качестве объяснения полученного результата. Размер выборки зависит от четырех характеристик исследования: величины различия в частоте исходов между группами,  $p_{\alpha}$ ,  $p_{\beta}$  и природы данных. Размер выборки следует учитывать исследователю, планирующему эксперимент, и читателю, решающему, следует ли доверять опубликованным результатам.

*Группа контрольная* (control group), или *группа сравнения* - группа испытуемых, получающих обычное лечение, или не получающих лечения, или получающих плацебо. Результаты измерений в контрольной группе сравниваются с показателями в экспериментальной группе для оценки эффекта исследуемого метода лечения.

*Группа экспериментальная* (experimental group) - это группа, подвергающаяся вмешательству (лечению) в ходе исследования. Иначе она еще называется *группой лечения или группой вмешательства* (treatment group или intervention group).

*Данные качественные* (nominal data) - это такие признаки, которые нельзя выразить количественно. Например: диагноз, группа крови, страна проживания. Частным случаем качественных данных являются *дихотомические данные*.

*Данные количественные* (numerical, or interval data) - это такие величины, которым присущ естественный порядок расположения с равными интервалами между последовательными значениями, независимо от их места на шкале (масса тела, концентрация глюкозы в крови).



*Данные дискретные* (discrete data) - это количественные данные, которые выражаются целыми числами. Например: число беременностей; число рождений живых младенцев; число судорожных припадков у больного в течение месяца.

*Данные дихотомические* (dichotomous data) - признаки, которые могут иметь только два значения (присутствует, отсутствует, да-нет, жив-умер).

*Данные непрерывные* (continuous data) - это количественные данные, которые могут принимать любое значение на непрерывной шкале. Примеры: масса тела, артериальное давление, парциальное давление кислорода в артериальной крови.

*Данные порядковые* (ordinal data) - это величины, которые могут быть расположены в естественном порядке (ранжированы), например, от малого до большого или от хорошего до плохого, но размер интервала между такими категориями не может быть выражен количественно (например: стадии болезни; оценки «высокий, средний, низкий» или «отсутствует, слабый, умеренный, тяжелый»).

*Добавочная доля популяционного риска* (population attributable risk fraction) - доля заболеваемости (смертности, инвалидности) в популяции, связанная с данным фактором риска; рассчитывается путем деления добавочного популяционного риска на общую заболеваемость (смертность, инвалидность) в популяции.

*Добавочный, или атрибутивный риск* (attributable risk) - определяется как заболеваемость для лиц, подвергавшихся воздействию фактора риска, минус заболеваемость для лиц, не подвергавшихся этому воздействию. Добавочный риск - это дополнительные случаи развития заболевания, обусловленные воздействием фактора риска. Учитывая способ вычисления добавочного риска, его также называют *разницей рисков* (risk difference), или *непосредственным риском*.

*Доверительный интервал* (confidence interval) - статистический показатель, позволяющий оценить, в каких пределах может находиться истинное значение параметра в популяции; диапазон колебаний истинных значений. Величины, полученные в исследованиях на выборке больных, отличаются от истинных величин в популяции вслед-

ствие влияния случайности. Так, 95% доверительный интервал означает, что истинное значение величины с вероятностью в 95% лежит в его пределах. Доверительные интервалы помогают ориентироваться, соответствует ли данный диапазон значений представлениям читателя о клинической значимости эффекта и каких результатов можно ожидать, применив описанную методику на сходной группе больных. Величина доверительного интервала характеризует степень доказательности данных, в то время как величина  $p$  указывает на вероятность ошибочного отклонения нулевой гипотезы (см. Величина  $p$ ).

*Достоверность (validity)* - характеристика, показывающая, в какой мере результат измерения соответствует истинной величине. Достоверность исследования определяется тем, в какой мере полученные результаты справедливы в отношении данной выборки (internal validity). Это внутренняя характеристика, она касается именно данной группы больных и не обязательно распространяется на другие группы.

*Исследование «до-после» (before-after study)* - неконтролируемое клиническое исследование, описывает течение заболевания в одной группе пациентов, подвергающейся изучаемому вмешательству. Подход основан на предположении, что любое улучшение, наблюдаемое после лечения, обусловлено именно лечением. Это предположение может оказаться ложным, что делает данный метод весьма уязвимым.

*Исследование распространенности (prevalence study)* - вариант описательного исследования, проводимого в определенный момент времени для оценки распространенности заболевания или исхода, изучения течения заболевания и т.п.

*Исследование серий случаев (case series)* - описательное исследование, представляющее собой количественный анализ группы больных.

*Исследование случай - контроль (case control study)* - ретроспективное исследование, в котором по архивным данным или воспоминаниям, суждениям пациентов производится сравнение двух групп, в одну из которых отобраны пациенты с определенной пато-

логией, а в другую - лица без нее. При этом сначала осуществляется отбор группы пациентов с изучаемым заболеванием и сходной по остальным признакам группы лиц без этого заболевания. Затем ретроспективно оценивается частота воздействия возможного фактора риска в обеих группах. Полученные данные позволяют рассчитать относительный риск развития заболевания в связи с изучаемым фактором.

*Исследование совокупного риска* (aggregate risk study), или *экологическое исследование* (ecological study) - исследование, в котором фактор риска характеризуется средним воздействием на группу; его задача - оценить влияние на людей факторов окружающей среды в целом.

*Исторический, или непараллельный контроль* (historical or nonconcurrent control) - способ клинического исследования, при котором результаты современного лечения сравниваются с данными наблюдения аналогичных пациентов в прошлом. Противоположность *параллельному контролю* (concurrent control), при котором контрольная группа набирается одновременно и по тем же правилам, что и экспериментальная группа

*Историческое когортное исследование* (historical cohort study) - когорта выделена по архивным документам и прослежена до настоящего времени.

*Клиническая эпидемиология* (clinical epidemiology) - это наука, разрабатывающая методы клинических исследований, которые дают возможность делать справедливые заключения, контролируя влияние систематических и случайных ошибок.

*Клинические испытания* (clinical trials) - специальный вид когортных исследований для оценки результатов разных вмешательств, условия проведения которых (отбор групп вмешательства, характер вмешательства, организация наблюдения и оценка исходов) направлены на устранение влияния систематических ошибок на получаемые результаты.

*Когорта* (cohort) - группа лиц, изначально объединенных каким-либо общим признаком (например, здоровые лица, или больные на определенной стадии заболевания) и наблюдаемых в течение опре-

деленного периода времени, чтобы проследить, что с ними произойдет в дальнейшем.

*Когорта дожития* (survival cohort) - когорта, в которую включаются пациенты, имеющие заболевание и доступные для наблюдения - например, пациенты специализированной клиники. Другой термин для таких групп - *доступные когорты* (available patients cohort).

*Когортное исследование* (cohort study) - исследование, в котором определенная когорта пациентов прослеживается в течение некоторого периода времени. Когортные исследования называют также *продольными* или *лонгитудинальными* (longitudinal study). Термином *проспективные* (prospective study) подчеркивается, что пациенты прослеживаются во времени - имеется в виду, что группа сформирована в настоящее время и будет прослежена в будущем. Термином *исследованиями заболеваемости* (incidence study) обращается внимание на то, что основным способом оценки является регистрация новых случаев заболевания в течение определенного срока.

*Мета-анализ* (meta-analysis). Методология объединения разнородных и выполненных различными авторами исследований, относящихся к одной теме, для повышения достоверности оценок одноименных результатов. Количественный анализ объединенных результатов нескольких клинических испытаний одного и того же вмешательства. Такой подход обеспечивает большую статистическую мощность, чем в каждом отдельном испытании за счет увеличения размера выборки. Используется для обобщенного представления результатов многих испытаний и для увеличения доказательности результатов испытаний. *Кумулятивный мета-анализ* — метод расчета обобщенной точечной оценки изучаемого эффекта, при котором исследования добавляются по одному в определенном порядке. *Мета-регрессионный анализ* - регрессионный анализ, в котором в качестве единицы наблюдения рассматриваются отдельные исследования.

*Многофакторное моделирование* (multivariable modeling) - метод, обеспечивающий математическое выражение сочетанного влияния многих переменных. Метод называется *многофакторным*, так как

он оценивает эффект многих переменных. Термин «*моделирование*» означает, что это математическая конструкция, которая рассчитывается из полученных данных, но основана она на упрощающих допущениях о характеристиках этих данных (например, что все переменные имеют нормальное распределение и одинаковую дисперсию).

*Многофакторный анализ* (multivariable analysis) - это совокупность статистических методов, которые одновременно рассматривают влияние многих переменных на какой-либо один фактор. Если после устранения влияния этих переменных действие фактора сохраняется, его воздействие считается независимым. Кроме того, эти методы применяются для выделения из большого числа признаков малого подмножества, которое вносит независимый и существенный вклад в конечный результат (исход), что используется для ранжирования переменных по силе их воздействия на исход и для других целей.

*Модификация эффекта* (effect modification) - особый тип взаимодействия факторов, при котором сила корреляции между двумя переменными зависит от уровня некоторой третьей переменной, называемой *модификатором эффекта* (effect modifier).

*Мощность, статистическая* (statistical power) - вероятность того, что в исследовании будет найдено статистически достоверное различие, когда это различие действительно существует. Статистическая мощность равняется 1 минус  $p_{\beta}$  (см. *Величина  $p_{\beta}$* ). Мощность исследования аналогична чувствительности диагностического теста.

*Обобщаемость* (external validity or generalizability) - внешняя характеристика, она определяется тем, в какой мере результаты данного исследования применимы к другим группам больных.

*Описание случая* (case report, case study) - простое изложение материала о нескольких больных.

*Открытое исследование* (open, open-label, unmasked study) - исследование, в котором не применялся слепой метод.

*Относительный риск* (relative risk), или *отношение рисков* (risk ratio) отношение заболеваемости среди лиц, подвергавшихся и не подвергавшихся воздействию факторов риска. Относительный риск не несет информации о величине абсолютного риска (заболеваемости). Даже при высоких значениях относительного риска абсолютный риск может быть совсем небольшим, если заболевание редкое. Относительный риск показывает силу связи между воздействием и заболеванием.

*Отношение правдоподобия* (likelihood ratio) - это отношение вероятности получить положительный результат диагностического теста у больных к вероятности получить положительный результат теста у здоровых лиц. Отношение правдоподобия для положительного результата теста - это чувствительность, деленная на 1 минус специфичность. Таким образом, отношение правдоподобия отражает одновременно и чувствительность, и специфичность теста. Если отношение правдоподобия положительного результата теста равно 1, то это значит, что вероятность положительного результата теста у больного такая же, как вероятность положительного результата теста у здорового. Если отношение правдоподобия положительного результата теста равно 2,5, то это значит, что вероятность положительного теста у больного в 2,5 раза выше, чем вероятность положительного результата теста у здорового.

*Отношение шансов* (odds ratio) - определяется как отношение шансов события в одной группе к шансам события в другой, или как отношение шансов того, что событие произойдет, к шансам того, что событие не произойдет. В исследованиях случай контроль отношение шансов используется для оценки относительного риска.

*Ошибка, систематическая, смещение* (systematic error, bias) - это неслучайное, однонаправленное отклонение результатов от истинных значений. Систематическая ошибка может возникать вследствие отбора (при создании выборки) - sampling, или assembling bias, вследствие измерений - measurement bias, при воздействии неучтенных факторов - confounding bias и во многих других случаях. О систематической ошибке говорят также, имея в виду предвзятость при публикации положительных результатов исследований и отклонении отрицательных - publication bias. Для борьбы с систематическими

ошибками и получения достоверных данных используются организационные методы (например, рандомизация, «слепой метод» и т.п.), а также внесение поправок, учитывающих величину смещения.

*Ошибка случайная или вариабельность* (random variation) - отклонение результата отдельного наблюдения (измерения) от его истинного значения, обусловленное исключительно случайностью.

*Ошибка экологическая* (ecological fallacy) - систематическая ошибка, связанная с тем, что заболевшие лица в группе обследования на самом деле могли не испытывать влияния факторов риска. Возникает, когда группу выделяют по общим косвенным признакам (проживание на одной территории, принадлежность к одной профессии), а не по наличию фактора риска у отдельных субъектов (например, по потреблению алкоголя).

*Плацебо* (placebo) - лекарственная форма, неотличимая от исследуемого препарата по внешнему виду, цвету, вкусу и запаху, но не оказывающая специфического действия (например, таблетки глюкозы или инъекции изотонического раствора хлорида натрия), или иное безразличное вмешательство, используемое в медицинских исследованиях для имитации лечения с целью устранения систематической ошибки, связанной с *плацебо-эффектом*.

*Плацебо - эффект* (placebo effect) - изменение состояния пациента (отмечаемое самим пациентом или лечащим врачом), связанное с фактом лечения, а не с биологическим действием препарата.

*Популяция* (population) - это совокупность индивидуумов, из которой отбирается выборка и на которую могут быть распространены результаты, полученные для этой выборки. Популяция может представлять собой все население (обычно таковы популяции в эпидемиологических исследованиях причин заболеваний), или же состоять из пациентов, госпитализированных в определенную клинику или из пациентов с определенным заболеванием (что чаще имеет место в клинических исследованиях). Таким образом, можно говорить об общей популяции, госпитальной популяции или популяции пациентов с конкретным заболеванием. Эпидемиологическое определение популяции отличается от биологического (экологического).

*Претестовая вероятность* - вероятность наличия заболевания до применения диагностического теста.

*Проверка гипотез* (hypothesis testing) о равенстве средних методов статистического доказательства гипотезы о равенстве средних величин. Нулевая гипотеза ( $H_0$ ) заключается в том, что различий между двумя выборками, оцененным по их средним значениям, нет. Противоположная гипотеза ( $H_1$ ) утверждает обратное. Принятие нулевой гипотезы означает, что обнаруженная разница случайна, отклонение нулевой гипотезы что разница неслучайна. Вероятность ошибочного отклонения нулевой гипотезы называется *альфа-ошибкой*, уровень статистической значимости обнаруженных различий обозначается как *величина p*, вероятность ошибочного принятия нулевой гипотезы называется *бета-ошибкой*.

*Проверка гипотез* (hypothesis testing) о равенстве средних методов статистического доказательства гипотезы, о равенстве средних величин. Нулевая гипотеза ( $H_0$ ) заключается в том, что различий между двумя выборками, оцененным по их средним значениям, нет. Противоположная гипотеза ( $H_1$ ) утверждает обратное. Принятие нулевой гипотезы означает, что обнаруженная разница случайна, отклонение нулевой гипотезы что разница неслучайна. Вероятность ошибочного отклонения нулевой гипотезы называется *альфа-ошибкой*, уровень статистической значимости обнаруженных различий обозначается как *величина p*, вероятность ошибочного принятия нулевой гипотезы называется *бета-ошибкой*.

*Прогностическая ценность* (predictive value) - вероятность наличия заболевания при известном результате диагностического теста. Прогностическая ценность называется также *апостериорной* (или *посттестовой*) *вероятностью*, поскольку это вероятность наличия (или отсутствия) болезни после того, как стали известны результаты теста.

*Прогностическая ценность отрицательного результата* (negative predictive value) - вероятность отсутствия заболевания при отрицательном («нормальном») результате теста.

*Прогностическая ценность положительного результата* (positive predictive value) - это вероятность заболевания при положительном («ненормальном») результате теста.



*Профилактика, вторичная* (secondary prevention) - это вмешательства, направленные на то, чтобы замедлить или остановить развитие заболевания у больных.

*Профилактика, первичная* (primary prevention) - вмешательства, направленные на предупреждение возникновения болезни, обычно за счет устранения ее причин.

*Профилактика, третичная* (tertiary prevention) - мероприятия, направленные на предотвращение ухудшения течения или осложнений заболевания после того, как болезнь проявилась.

*Рандомизация* (randomization) - процедура, обеспечивающая случайное распределение больных в экспериментальную и контрольную группы. Случайным распределением достигается отсутствие различий между двумя группами и, таким образом, снижается вероятность систематической ошибки в клинических исследованиях вследствие различий групп по каким-либо признакам.

*Распространенность* (prevalence) - это частота некоторого состояния у пациентов в группе. Рассчитывается как отношение числа лиц, у которых на момент обследования наблюдается изучаемое состояние (болезнь или исход), к числу всех лиц в группе (популяции, обследованным, работникам предприятия).

*Риск* (risk) - вероятность нанесения ущерба здоровью или причинения дискомфорта.

*Риск, добавочный (атрибутивный)* - см. *Добавочный риск*.

*Риск, относительный* - см. *Относительный риск*.

*Риск, популяционный добавочный (атрибутивный)* (population attributable risk) - рассчитывается как произведение добавочного риска на распространенность фактора риска в популяции. Этот показатель отражает дополнительную заболеваемость в популяции, связанную с фактором риска.

*Скрининг* (screening) - массовое обследование лиц, не считающих себя больными, для выявления скрыто протекающих заболеваний или других состояний (факторов риска будущих заболеваний). Обычно проводится с использованием дешевых, простых, неинвазивных диагностических процедур, имеющих высокую чувствительность.

*Слепой, или маскированный метод* (blinding, or masking) - процедура, обеспечивающая отсутствие информации о том, к какой группе - экспериментальной или контрольной отнесен каждый испытуемый. При *простом слепом методе* информация отсутствует только у пациента, при *двойном слепом* - у пациента и исследователя, при *тройном слепом* - у пациента, исследователя и лиц, проводящих статистическую обработку результатов исследования. Применяется для устранения систематической ошибки в клинических исследованиях.

*Смещение к среднему* (regression to the mean) - закономерность в биомедицинских исследованиях, выражающаяся в том, что величины, отклоняющиеся от среднего значения, при последующих измерениях оказываются ближе к среднему.

*Специфичность диагностического теста* (specificity) - вероятность отрицательного результата диагностического теста при отсутствии болезни.

*Точка разделения* (cut - off point, cutpoint) - величина, используемая для разделения ряда величин на две части. В диагностическом процессе точка разделения отделяет «нормальные» показатели от «ненормальных». Результаты теста, лежащие в области «нормы», называются «отрицательным результатом», а лежащие в области «патологии» - «положительным результатом».

*Точность теста* (test accuracy) - доля правильных результатов теста (истинно положительных и истинно отрицательных) в общем количестве полученных результатов.

*Фактор риска* (risk factor) - особенность организма или внешнее воздействие, приводящие к увеличению риска возникновения заболевания или иному неблагоприятному исходу.

*Фармакоэпидемиология* (pharmacoepidemiology). Наука о применении эпидемиологических знаний, методов и доводов к изучению благоприятных и нежелательных эффектов для рационального использования лекарств в различных человеческих популяциях. Предметом исследования являются результаты наблюдений, собранных при повседневном клиническом обслуживании больных, в отличие от рафинированных данных рандомизированных, двойных слепых клинических испытаний. Цель исследований – определить реальную эффективность лекарств в клинической практике.

*Фармакоэкономика* (pharmacoeconomics). Анализ стоимости лекарственной терапии для системы здравоохранения и общества с целью выявить наиболее выгодные препараты (виды лечения) для финансирующих организаций, пациентов и производителей. На результатах специальных (см. ниже) типов фармакоэкономического анализа, основываются решения о планах финансирования различных программ здравоохранения, о приоритетах среди возможных альтернатив, о возможностях пациента использовать оптимальный для него вариант лечения.

*Характеристическая кривая* (receiver operating characteristic (ROC) curve) - график, характеризующий диагностическую точность теста. На оси ординат откладывается чувствительность (доля истинно положительных результатов), на оси абсцисс - 1 минус специфичность (доля ложноположительных результатов). Чем дальше кривая отстоит от диагонали, тем выше точность теста.

*Частота новых случаев* (incidence) - это отношение числа лиц, у которых в течение определенного времени развилось изучаемое состояние, ко всем обследованным в группе, где исходно этого состояния никто не имел. Частота новых случаев болезни в популяции, где это заболевание исходно отсутствовало, называется *заболеваемостью*. Частота новых случаев исходов при определенном заболевании указывает, соответственно, на инвалидность или летальность при этом заболевании. Для оценки частоты новых случаев отбирают группу лиц без заболевания (или без изучаемого исхода) и периодически обследуют ее, подсчитывая число новых случаев, появившихся за определенный период.

*Частотное распределение* (frequency distribution) - величины, расположенные в порядке возрастания и частота, с которой каждая величина встречается в популяции или выборке. Обычно представляется в виде графика, на котором по горизонтальной оси откладываются наблюдаемые величины, а по вертикальной - частоты с которой эти величины наблюдаются.

*Число больных, которых необходимо лечить* (number needed to treat, NNT) - способ оценки относительной эффективности двух методов лечения. Показывает, какое количество больных необходи-

мо подвергнуть лечению исследуемым методом для предотвращения одного случая изучаемого исхода. Например, если для предотвращения одного случая инфаркта миокарда надо лечить антигипертензивным средством 33 пациентов с артериальной гипертензией в течение 5 лет, то NNT равно 33 за 5 лет. Этот показатель является величиной, обратной снижению абсолютного риска.

*Чувствительность диагностического теста (sensitivity)* - вероятность положительного результата диагностического теста при наличии болезни.

*Шансы (odds)* - отношение вероятности того, что событие произойдет, к вероятности того, что событие не произойдет. Шансы и вероятности содержат одну и ту же информацию, но по-разному выражают ее. Если вероятность того, что событие произойдет обозначить  $P$ , то шансы этого события будут равны  $P/(1 - P)$ . Например, если вероятность выздоровления 0,3, то шансы выздороветь равны  $0,3/(1 - 0,3) = 0,43$ . Шансы удобнее использовать для некоторых расчетов, чем вероятности.

*Экспериментальное исследование (experimental study)* - сравнительное исследование, заранее спланированное и посвященное изучению влияния по крайней мере одного вмешательства. Сравнения могут производиться между двумя и более группами или внутри одной группы до и после назначения вмешательства.

*Эффект ярлыка (labeling)* - психологическое воздействие результатов теста или диагноза на пациента - участника скрининга.

## ЛИТЕРАТУРА

1. *Sackett D., Haynes B. et al.* Clinical Epidemiology: A Basic Science for Clinical Medicine. 2nd ed. Boston, Mass.: Little Brown, 1991.
2. *Бащинский С.Е.* Evidence-based medicine и международный журнал медицинской практики // Международный журнал медицинской практики. - 1996. - № 1. - С. 6-11.
3. *Lie K.I., Wellens H.J. et al.* Lidocain in the prevention of primary ventricular fibrillation: a double-blind, randomized study of 212 consecutive patients. N Engl J Med 1974;291:1324.
4. *Antman E.M., Lau J., Kupelnick B. et al.* A comparison of results of meta-analyses of randomized trials and recommendations of clinical experts: treatment of myocardial infarction. JAMA 1992;268:240-248.
5. *Боборыкин В.М., Нугманова Д.С.* Гипербарическая оксигенация: Материалы национальной конференции Казахстана по качеству медицинской помощи. 15-16 мая, 1996. - Алма-Ата, 1996. - С. 77-80.
6. *Лисицын Ю.В., Нугманова Д.С.* Лазерная терапия: Там же. - С. 61-76.
7. *Двойрин В.В.* Какие публикации заслуживают доверия практического врача? // Международный журнал медицинской практики. - 1997. - № 1. - С. 17-19.
8. *Schultz K.F., Chalmers I. et al.* The methodologic quality of randomisation as assessed from reports of trials in specialist and general medical journals. Online J Clin Trails 1995 (doc N 197).
9. *Temple R.J.* A regulatory authority's opinion about surrogate endpoints. In: Nimmo W.S., Tucker G.T., eds. Clinical Measurement in Drug Evaluation. New York: J Wiley; 1995.
10. *Fleming T.R., DeMets D.L.* Surrogate end points in clinical trials: are we being misled? Ann Intern Med 1996;125:605-13.
11. Preliminary Report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction. The Cardiac Arrhythmia Suppression Trial (CAST) Investigators. N Engl J Med 1989;321:406—12.

12. *Echt D.S., Liebson P.R., Mitchell L.B. et al.* Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. *N Engl J Med* 1991;324:781—8.

13. Effect of the antiarrhythmic agent moricizine on survival after myocardial infarction. The Cardiac Arrhythmia Suppression Trial II Investigators. *N Engl J Med* 1992;327:227—33.

14. *Coplen S.E., Antman E.M., Berlin J.A. et al.* Efficacy and safety of quinidine therapy for maintenance of sinus rhythm after cardioversion. A meta-analysis of randomized control trials. *Circulation* 1990;82:1106—16.

15. *Hine L.K., Laird N. et al.* Meta-analytic evidence against prophylactic use of lidocaine in acute myocardial infarction. *Arch Intern Med* 1989;149:2694—8.

16. *MacMahon S., Collins R., Peto R. et al.* Effects of prophylactic lidocaine in suspected acute myocardial infarction. An overview of results from the randomized, controlled trials. *JAMA* 1988;260:1910—6.

17. *Law M.R., Thompson S.G., Wald N.J.* Assessing possible hazards of reducing serum cholesterol. *BMJ* 1994;308:373—9.

18. *Law M.R., Wald N.J., Thompson S.G.* By how much and how quickly does reduction in serum cholesterol concentration lower risk of ischemic heart disease? *BMJ* 1994;308:367—72.

19. *Rossouw J.E., Lewis B., Rifkind B.M.* The value of lowering cholesterol after myocardial infarction. *N Engl J Med* 1990;323:1112—9.

20. *Der Simonian R., Charette L.J. et al.* Reporting on methods of clinical trials. *N Engl J Med* 1982;306:1332—7.

21. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet* 1994;344:1383—9.

22. Five-year findings of the hypertension detection and follow-up program. I. Reduction in mortality of persons with high blood pressure, including mild hypertension. Hypertension Detection and Follow-up Program Cooperative Group. *JAMA* 1979;242:2562—71.

23. Prevention of stroke by antihypertensive drug treatment in older persons with isolated systolic hypertension: final results of the Systolic Hypertension in the Elderly Program (SHEP). SHEP Cooperative Research Group. *JAMA* 1991;265:3255—64.

24. *Furberg C.D., Berglund G. et al.* Overtreatment and undertreatment of hypertension. *J Intern Med* 1994;235:387—97.

25. *Psaty B.M., Heckbert S.R., Koepsell T.D. et al.* The risk of incident myocardial infarction associated with antihypertensive drug therapies [Abstract]. *Circulation* 1996;91:925.

26. *Held P.H., Yusuf S., Furberg C.D.* Calcium channel blockers in acute myocardial infarction and unstable angina: an overview. *BMJ* 1989; 299: 1187—92.

27. *Riggs B.L., Hodgson S.F., O'Fallon W.M. et al.* Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis. *N Engl J Med* 1990;322:802—9.

28. *Riggs B.L., Seeman E., Hodgson S.F. et al.* Effect of the fluoride/calcium regimen on vertebral fracture occurrence in postmenopausal osteoporosis. Comparison with conventional therapy. *N Engl J Med* 1982;306:446—50.

29. Coronary artery surgery study (CASS): a randomized trial of coronary artery bypass surgery. Survival data. *Circulation*.1983;68:939—50.

30. *Freedman L.S., Graubard B.L., Schatzkin A.* Statistical validation of intermediate endpoints for chronic diseases. *Stat Med* 1992;11:167—78.

31. Единые требования к рукописям, представляемым в биомедицинские журналы // Международный журнал медицинской практики. - 1997. - № 5. - С. 53-64.

32. *Шпигель А.С., Вакурова Н.В.* Лечение и профилактика силикоза гомеопатическими лекарственными средствами // Самарскому государственному университету - 80 лет: Сборник тезисов докладов юбилейной научно-практической конференции / Под редакцией Г.П.Котельникова. - Самара; СамГМУ, 1999. - С.297 -298.

33. *Власов В.В.* Как читать медицинские статьи. Часть 1. Общий алгоритм оценки статьи // Международный журнал медицинской практики. - 1996.

34. *Власов В.В.* Как читать медицинские статьи. Часть 2. Исследования, посвященные методам диагностики // Там же. - 1997. - № 1. - С. 11-16.

35. *Власов В.В.* Как читать медицинские статьи. Часть 4. Исследования этиологии и патогенеза заболеваний // Там же. - № 3. - С. 7-10.

36. Власов В.В. Как читать медицинские статьи. Часть 5. Испытания методов лечения и профилактики заболеваний // Там же. - № 6. - С. 9-13.

37. Statistical Methods in Cancer Research: Part 1. The analysis of case-control studies. IARC Sci Publ. N.32. Lyon: WHO, IARC, 1980: 1-338

38. Власов В.В. Факторы риска и стадии развития заболевания // Клиническая медицина. - 1991. - № 10. - С. 98-99.

39. Власов В.В. Эффективность диагностических исследований. - М., 1988. - С. 245.

40. Bailar J.C., Louis T.A. et al. A classification for biomedical research reports. N Engl J Med 1984;311:23: 1482-1487.

41. Brown G.W., Baca G.M. A classification of original articles. Am J Dis Child 1986;140:641-645.

42. CONSORT Group. Improving the quality of reporting of randomised controlled trials: the CONSORT statement. JAMA 1996; 276:637—9.

43. Guyatt G.H., Sackett D.L., Cook D.J. User's guides to the medical literature: II. How to use an article about therapy or prevention. A. Are the results of the study valid? JAMA 1993;270:2598—601.

44. Guyatt G.H., Sackett D.L., Cook D.J. User's guides to the medical literature: II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients. JAMA 1994;271:59—63.

45. Применение селегилина при болезни Паркинсона (Письмо в редакцию с комментарием) // Международный журнал медицинской практики. - 1997. - № 3. - С. 3-4.

46. Detsky A.S., Sackett D.L. When was a “negative” clinical trial big enough? How many patients you needed depends on what you found. Arch Int Med 1985;145:709—12.

47. Spodick D.H. Randomized controlled clinical trials: the behavioral case. JAMA, 1982;247:2258—60

48. Rosenbaum P.R. Discussing hidden bias in observational studies. Ann Int Med. - 1991;115:901—5.

49. Feinstein A.R. Meta-analysis: Statistical alchemy for the 21st century. J Clin Epidemiol. - 1995; 48:71—9.

50. Remein Q.R., Wilkerson H.L. The efficiency of screening tests for diabetes. J Chron Dis 1961;13:6—21.



51. *Шнигель А.С.* Нейрогормональная дисрегуляция при воздействии профессиональных вредностей физической и химической природы: Автореф. дис. ... докт-ра мед. наук. - М., 1990. - 45 с.

52. *Котельников Г.П., Яковлев О.Г., Захарова Н.О.* Геронтология и гериатрия: Учебник для студентов медицинских вузов, слушателей учреждений дополнительного профессионального образования и повышения квалификации специалистов. - М., 1997. - 800 с.

53. *Paling L.* How to live longer and feel better. New York: Freeman; 1986.

54. *Kleijnen J. and Knipschild P.* The comprehensiveness of Medline and Embase computer searches. Searches for controlled trials of homoeopathy, ascorbic acid for common cold and ginkgo biloba for cerebral insufficiency and intermittent claudication. *Pharm Wekbl (Sci)* 1992;14:316—20.

55. *Chalmers and Altman D.G.* Systematic reviews. London: BMJ Publishing Group; 1995:1.

56. *Last J.M.* A dictionary of epidemiology. New York: Oxford University Press; 1988:81

57. *Greer A.L.* The state of the art versus the state of the science. The diffusion of new medical technologies into practice. *Int J Technol Assess Health Care* 1988,4:5—26.

58. *Sacks H.S., Berrier J., Reitman D., Pagano D., Chalmers T.C.* Meta-analysis of randomized controlled trials: an update of the quality and methodology. In: *Baird J.C., Mosteller F., eds. Medical Uses of Statistics*, 2d ed. Boston: NEJM Books; 1992:427—42.

59. *Jadad A.R., McQuay H.J.* Meta-analyses to evaluate analgesic interventions: a systematic qualitative review of their methodology. *J Clin Epidemiol* 1996;49:235—43.

60. *Phillips K., Bero L.A.* Improving the use of information in medical effectiveness research. *Int J Qual Health Care* 1996;8:21—30.

61. The Cochrane Collaboration. The Cochrane Library [database on disk and CD ROM], v. 2 Oxford: Update Software: 1996.

62. *Флетчер Р., Флетчер С., Вагнер Э.* Клиническая эпидемиология: Монография / Пер. с англ. *С.Е. Бащинского, С.Ю. Варшавского.* - М.: Медиа Сфера, 1998. - 352 с.

63. *Королюк И.П.* Введение в медицинскую информатику: Лекция для студентов медицинского института. - Куйбышев, 1989. - 41 с.

## СОДЕРЖАНИЕ

1. Клиническая эпидемиология - новая область знаний .....	3
2. Оценка качества клинической информации и ее интерпретация .....	6
3. Понятия: истинный клинический исход и косвенные критерии оценки .....	12
4. Современные методические требования для авторов публикаций, посвященных оценке эффективности вмешательства .....	18
5. Методы и критерии отбора информации .....	25
6. Взаимосвязь дизайна и структуры клинических испытаний ....	28
7. Принципы планирования и составления программ клинических испытаний .....	37
7.1. Фазы клинических испытаний .....	41
7.2. Использование плацебо и рандомизации в слепых исследованиях - возможность выявить снижение преднамеренных и непреднамеренные искажений .....	44
8. Методологические стандарты испытаний методов лечения (критерии достоверности) .....	48
9. Оценка публикаций, посвященных методам диагностики .....	54
10. Трудности внедрения научно-обоснованной медицинской практики и пути их преодоления .....	66
11. Практическая реализация доказательной медицины - разработка формулярной системы (совместно с С.И. Кузнецовым) .....	69
11.1. Анализ терапевтических классов лекарственных средств .....	81
11.2. Экономические аспекты анализа терапевтических классов лекарственных средств .....	83
12. Систематические обзоры и мета-анализ .....	87
СЛОВАРЬ ТЕРМИНОВ .....	94
ЛИТЕРАТУРА .....	109