

ISBN 5-89084-015-0

Путеводитель читателя медицинской литературы

М., Медиа Сфера, 2003

Основой настоящей книги послужила популярная серия статей, опубликованных в журнале JAMA под общим названием *User's Guides to the Medical Literature* (*Путеводитель читателя медицинской литературы*). В создании книги участвовали более 50 самых известных теоретиков и практиков доказательной медицины из разных стран. "Путеводитель читателя медицинской литературы" — результат почти двадцатилетнего опыта учебной и научной работы. Эта книга содержит подробные и в то же время доступные для практикующего врача сведения, необходимые для осмысленного применения данных медицинской литературы в конкретных клинических ситуациях. "Путеводитель читателя медицинской литературы" снабжен словарем терминов, предметным указателем, приложением с формулами расчета статистических показателей и представляет собой идеальное учебное пособие для студентов, врачей и преподавателей медицинских вузов.

Перевод с английского: П.В. Безменов, И.В. Журавлев, Е.А. Султанова, Е.А. Фролова

Редакторы: О.Н. Анкудинова, О.Л. Лобачев, М.Н. Соловова

© American Medical Association, 2003

Все права защищены. Репродуцирование любым способом, распространение или передача данного издания запрещена без предварительного согласия издательства Медиа Сфера.

© Медиа Сфера, Москва, 2003 — пер. с англ.; оформление
Отпечатано в типографии «Информполиграф», Москва

СОДЕРЖАНИЕ

Введение.....	V
Предисловие.....	IX
Авторы.....	XIII
Как пользоваться книгой.....	XVI

ЧАСТЬ 1

ОСНОВНЫЕ СВЕДЕНИЯ: ИСПОЛЬЗОВАНИЕ МЕДИЦИНСКОЙ ЛИТЕРАТУРЫ

1 * Введение: философская концепция доказательной медицины.....	3
2 • Поиск доказательной информации.....	17
3 * Лечение и риск: введение.....	63
4 * Исследования, посвященные эффективности лечения.....	71
5 • Оценкариска.....	105
6 * Диагностика.....	133
7 » Дифференциальная диагностика.....	145
8 • Диагностические методы.....	161
9 • Прогноз.....	189
10* Обобщение доказательств.....	207
11 * От доказательств к практике.....	233

ЧАСТЬ II

БОЛЕЕ ДЕТАЛЬНЫЕ СВЕДЕНИЯ: ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ ПРИНЦИПОВ
ДОКАЗАТЕЛЬНОЙ МЕДИЦИНЫ И ОБУЧЕНИЕ ЕЕ ОСНОВАМ

1 • Лечение и риск: причины недостоверности результатов — систематические и случайные ошибки.....	273
2 • Лечение и понимание результатов исследования: проверка гипотез.....	285
3 * Лечение и понимание результатов исследования: доверительные интервалы.....	297
4 * Лечение и понимание результатов исследования: измерение корреляции.....	309
Приложение.....	329
Словарь терминов.....	337
Предметный указатель.....	376

Введение

Школьные годы я провел в Британии в годы второй мировой войны, где на фоне скудного питания (главным образом вареной капустой) основными занятиями моих сверстников были не только обучение математике, холодные обливания, питание вареной капустой, долгие путешествия по стране, но также изучение латинского и французского языков. Несомненно, занятия латынью представляли собой чисто теоретический интерес, ведь Римской империи давно не существовало. Изучение французского языка не многим отличалось от штудирования латыни, поскольку Франция, хотя и находилась совсем рядом, по ту сторону пролива, в течение долгих лет была по той или иной причине недоступна. Ни я, ни мои учителя не полагали всерьез, что мне придется воспользоваться полученными знаниями на практике и когда-либо говорить по-французски.

Аналогичным образом многие практикующие врачи относятся к регулярному знакомству с новинками медицинской литературы — они всегда под рукой, но фактически недоступны. Все понимают, что работа врача должна основываться на данных, которые публикуются в медицинских журналах. С другой стороны, хорошо известно, что каждые несколько лет объем литературы удваивается, а у нас остается все меньше времени на знакомство с ней [1]; поэтому внимательное чтение многих журналов постепенно становится почти нереальным. Внедрение в повседневную практику данных, излагаемых в сотнях тысяч статей, мы оставляем другим. По мере того как возможность использования литературы становится все более призрачной, сама мысль о том, что результаты конкретного исследования можно применить при обследовании или лечении конкретного больного, начинает казаться фантастикой.

Данная книга призвана помочь вам поверить в обратное. Изучив ее, вы научитесь понимать язык любых медицинских статей, сможете избавиться от работы по шаблону, по наитию и сообразно личному опыту, который далеко не всегда бывает достаточным. Вы освободитесь от диктата фармацевтических компаний, навязывающих свои препараты, и от давления больных, вдохновленных новыми методами лечения, оценить эффективность которых у вас нет возможности. Вы перестанете целиком и полностью зависеть от мнений малограмотных "экспертов". У вас появится возможность использовать данные, опубликованные в медицинской литературе, при решении частных клинических проблем, оценивать актуальность и достоверность этих данных. Другими словами,

перед вами раскроется возможность использования самого мощного информационного ресурса из всех, существующих в медицине.

СЕРИЯ СТАТЕЙ, ПОСВЯЩЕННЫХ ПРИНЦИПАМ РАБОТЫ С МЕДИЦИНСКОЙ ЛИТЕРАТУРОЙ И ОПУБЛИКОВАННЫХ В ЖУРНАЛЕ JAMA

История создания серии статей под общим названием "Путеводитель читателя медицинской литературы", которые публиковались в журнале *JAMA*, и самой этой книги изложена в предисловии, написанном Гордоном Гайяттом — инициатором, главным редактором и соавтором большинства материалов. Но как получилось, что частью этой истории стал журнал *JAMA*?

В конце 80-х годов XX века я воспользовался приглашением моего друга Дэвида Сакетта и посетил возглавляемое им подразделение в университете МакМастер, чтобы обсудить другую его идею — о публикации в журнале *JAMA* серии статей, в которых анализировались бы все имеющиеся данные об информативности сбора анамнеза и обследования. Результатом этой работы стала серия статей и систематических обзоров под общим названием "Рациональное клиническое обследование" (*Rational Clinical Examination*), которую начали публиковать в *JAMA* в 1992 г. при поддержке главного редактора Джорджа Ландберга [1]. К этому времени у меня установились очень тесные контакты с блестящими специалистами из университета МакМастер. Все они, как и их руководитель Д. Сакетт, стремились преодолеть существующие стереотипы, были готовы к совместной работе, вовлекали в свою группу новых талантливых сотрудников и доводили начатое до конца.

Поэтому, когда стало известно о намерении этой группы обновить прекрасную серию статей о принципах изучения медицинской литературы, которые публиковались в журнале *Canadian Medical Association Journal* в 1981 г., я предложил подготовить и выпустить новую серию в *JAMA*. Так, в тесном сотрудничестве с Д. Сакеттом и Энди Оксманом, которого впоследствии сменил Г. Гайяйт (поскольку Э. Оксман переехал в Осло), была создана серия статей под общим названием "Рекомендуемый подход к изучению медицинской литературы". Ее публикация в *JAMA* началась в 1993 г. [2,3].

Вначале мы предполагали подготовить 8 или 10 статей по данной тематике, однако читатели восприняли выход серии с большим энтузиазмом, а методологические особенности исследований, которые следовало рассмотреть, оказались настолько разнообразными, что спустя 7 лет я продолжал получать, направлять на рецензирование и редактировать

все новые статьи этой серии. В конце концов Г. Гайятт и я завершили серию 25-й статьей; вся серия была опубликована в 33 номерах журнала.

Тот факт, что от начала до завершения этого проекта прошло долгих восемь лет, сыграл положительную роль. Многие проблемы, которые в 1992—1993 г. едва затрагивались в основных медицинских журналах, но за последнее время привлекли внимание медицинской общественности, смогли получить достойное освещение в статьях серии. Так, в 2000 г. были опубликованы две статьи о том, как читатель должен относиться к результатам медицинских исследований, в которых использовались качественные методы оценки [4, 5]. Другим примером может служить возросший интерес к систематическим обзорам и мета-анализам, чему в значительной мере способствовала деятельность Кокрановского сотрудничества. Одна из статей серии была посвящена подходу к оценке результатов таких исследований [6]. В состав серии вошла также статья об электронных источниках медицинской информации [7].

КНИГА

Уже после появления первых статей читатели предлагали издать их в виде книги. Мы и сами собирались сделать это, однако данное намерение откладывалось по мере подготовки каждой новой статьи. Такое промедление пошло будущей книге только на пользу. Когда в 1981 г. в *Canadian Medical Association Journal* появилась первая серия статей, еще отсутствовала концепция доказательной медицины, предложенная позднее Г. Гайяттом; лишь незначительная часть медицинских работников имели доступ к компьютерам. Интернета не существовало, о самой идее электронных публикаций можно было только мечтать. В 1992 г. Интернет только начинали использовать для получения информации; электронные адреса не были на слуху, как сейчас; компьютерная грамотность пускала в медицине первые ростки. Но уже в конце 90-х годов Г. Гайятт и я, предложив издательству АМА Press выпустить книгу не только в печатном варианте, но также на сайте в Интернете и на компакт-диске, немедленно получили согласие. Публикация этого издания в электронном виде стала возможной благодаря усилиям доктора Роба Хэйуорда из университета Альберты.

Кроме того, за последние 8 лет доказательная медицина также бурно развивалась, что нашло свое отражение на страницах книги.

К серии статей и главам книги предъявляются разные требования. Чтобы при работе с отдельной статьей избавить читателя от необходимости обращаться к другим источникам информации, она снабжается справочным материалом; одной из задач Г. Гайятта как редактора книги

было устранение выявляемых повторов. Наряду с этим члены Рабочей группы по разработке и внедрению принципов доказательной медицины обновили опубликованные ранее материалы, дополнив их современными сведениями.

Прежде всего мне хочется выразить признательность Г. Гайяту, который за многие годы стал моим хорошим другом; благодаря ему моя работа в качестве редактора серии статей в *JAMA* и члена редакционной коллегии книги была исключительно легкой и приятной. За время совместной деятельности мне удалось многому у него научиться. За 9 лет у меня сложились очень теплые личные отношения со многими из его коллег; к сожалению, объем этого раздела не позволяет назвать их всех поименно. Не могу не упомянуть Аннетт Флэнаджин — исключительно деятельную, творческую и дипломатичную сотрудницу *JAMA*. Создание книги было бы невозможным без помощи Тони Франкоса, а также без энтузиазма и усилий, приложенных Барри Боулюсом; оба они являются сотрудниками издательства АМА Press. Я благодарен образцовому редактору, аккуратному, грамотному специалисту Сюзан Мосс, которая прошла с нами весь путь создания книги. Они с Г. Гайяттом взвалили на себя всю ношу, оставив меня практически без работы. Хочу также поблагодарить Джин Роберте, Катарин Дворак и Эми Роберте из АМА Press. Наконец, я признателен моему очаровательному начальнику Кэти ДеАнджелис, новому главному редактору *JAMA*, за поддержку, так необходимую мне, моим коллегам и всему проекту в целом, за ее терпимость и шутки, которые так поднимали наш дух.

Драммонд Ренни

*Университет Калифорнии, Сан-Франциско
Заместитель главного редактора журнала JAMA*

Литература

1. *Durack D.T.* The weight of medical knowledge. *N Engl J Med* 1978;298:773–775
2. *Sackett D.L., Rennie D.* The science of the art of the clinical examination. *JAMA* 1992;267:2650–2652.
3. *Guyatt G.H., Rennie D.* Users' guides to the medical literature. *JAMA* 1993;270:2096–2097.
4. *Giacomini M.K., Cook D.J.* Users' guides to the medical literature, XXIII: qualitative research in health care. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 2000;284:357–362.
5. *Giacomini M.K., Cook D.J.* Users' guides to the medical literature. XXIII qualitative research in health care. B What are the results and how do they help me care for my patients? Evidence-Based Medicine Working Group. *JAMA* 2000;284:478–482.
6. *Oxman A.D., Cook D.J., Guyatt G.H.* Users' guides to the medical literature, VI: how to use an overview Evidence-Based Medicine Working Group. *JAMA* 1994;272:1367–1371.
7. *Hunt D.L., Jaeschke R., McKibbin K.A.* Users' guides to the medical literature, XXI using electronic health information resources in evidence-based practice. Evidence-Based Medicine Working Group *JAMA* 2000;283:1875–1879.

Предисловие

Основой данной книги послужила серия из 25 статей, опубликованных в журнале *JAMA* в период с 1993 по 2000 г. под общим названием "Путеводитель читателя медицинской литературы". Эта серия, над которой трудились более 50 экспертов в области доказательной медицины (ДМ), получила самую высокую оценку как практикующих врачей, так и преподавателей.

Предлагаемая читателю книга отличается от исходной серии статей. Во-первых, при подготовке книги опубликованные ранее материалы были переработаны и дополнены появившимися за последние 7 лет новыми данными о методологии проведения клинических исследований, в том числе о причинах возникновения систематических ошибок, о количественной оценке преимуществ и недостатков лечебных вмешательств, об учете предпочтений больных. Во-вторых, книга содержит разделы, посвященные вопросам обучения теоретическим и практическим основам ДМ. В-третьих, стоит отметить структуру изложения материала. Первая часть книги включает в себя сведения, знание которых, на наш взгляд, необходимо любому врачу. Во второй части книги (печатный вариант) приводится более детальное обсуждение, позволяющее глубже понять принципы ДМ. Специалисты и преподаватели могут найти всю необходимую для них информацию в расширенном электронном варианте второй части книги, изданной на компакт-диске. В-четвертых, все сведения представлены сжато, без лишних повторов; книга выдержана в едином стиле. В-пятых, основное внимание уделяется процессу принятия клинического решения; этому способствует включение большого числа клинических примеров и подчеркивание той роли, которую должен играть учет предпочтений больного при выборе оптимального вмешательства.

Оригинальное англоязычное издание книги доступно в печатном и электронном вариантах. Ответственным за создание электронного варианта, выпускаемого на компакт-дисках и публикуемого в Интернете (последняя версия включает в себя значительный объем дополнительной информации), был Роберт Хэйуорд. Он проделал колоссальную работу и в качестве одного из редакторов; вклад этого человека, затраченные им время и силы нельзя переоценить.

КАК НАЧИНАЛАСЬ РАБОТА НАД КНИГОЙ

В конце 70-х годов XX века группа эпидемиологов из университета МакМастер в составе Брайана Хэйнса, Питера Тагвелла и Виктора Ной-

фельда под руководством Дэвида Сакетта решила написать серию статей, призванных помочь врачам в изучении отчетов об исследованиях, которые публикуются в медицинских журналах. Первые статьи появились в *Canadian Medical Association Journals* 1981 г. При описании основных принципов разработанного ими подхода к практическому применению полученных в ходе исследований данных авторы использовали термин *"критическая оценка"*. Спустя несколько лет члены группы убедились в необходимости крупномасштабной программы, реализация которой позволила бы научить врачей не просто "листать" журналы, но использовать их для решения проблем, возникающих при лечении конкретных больных. Процесс практического применения опубликованных в литературе данных Д. Сакетт назвал *"переносом критической оценки к постели больного"*.

В 1990 г. меня назначили руководителем ординатуры по внутренним болезням в университете МакМастер. Благодаря усилиям Д. Сакетта критическая оценка информации у постели больного переросла в концепцию, подразумевающую принятие любого клинического решения только на основе тщательного поиска и изучения опубликованных данных. Эта концепция представляла собой принципиально новый подход к медицинской практике и поэтому, на наш взгляд, нуждалась в названии, которое бы подчеркнуло ее новизну.

Весной 1990 г. я представил членам руководства медицинского факультета, многие из которых не разделяли наши взгляды, предложения по изменению программы обучения врачей новым принципам. Для обозначения нового подхода я предложил термин "научная медицина", покоробивший наших оппонентов неприкрытым намеком на то, что вся прочая медицина была "ненаучной". Я изменил название на "доказательную медицину", что не встретило возражений.

В печати термин "доказательная медицина" впервые появился осенью 1990 г., когда вышел информационный бюллетень для поступающих в ординатуру по внутренним болезням в университете МакМастер. В этом документе говорилось:

"При повседневном использовании методов диагностики, лечения и прогнозирования ординаторы должны придерживаться просвещенного скептицизма. Такой подход, получивший название доказательной медицины, основан на концепции, которая изложена в руководстве "Клиническая эпидемиология". Он заключается в тщательном изучении актуальных научно обоснованных данных, оценке их достоверности и практической значимости. Врач должен уметь четко формулировать клинический вопрос, осуществлять поиск ответа на него в медицинской

литературе, проводить критическую оценку найденных фактов, определять возможность их использования при лечении конкретного больного и непосредственно применять найденные данные на практике."

Позднее термин "доказательная медицина" появился на страницах журнала *ACP Journal Club* в 1991 г. [1]. Тем временем группа энтузиастов из университета МакМастер, среди которых следует также упомянуть Дебору Кук, Романа Йешке, Джима Нишикаву, Пэта Брилл-Эдвардса и Акбара Панджу, продолжала разрабатывать вопросы практического внедрения принципов ДМ и обучения ее основам. Работа оказалась весьма продуктивной; поэтому мы решили, что эта концепция может быть интересной большему числу преподавателей.

Нам удалось установить контакты со многими университетскими врачами, главным образом из США, и впервые сформировать международную Рабочую группу по разработке и внедрению принципов ДМ. Затем была опубликована статья, в которой эти принципы получили дальнейшее развитие, в ней фактически было провозглашено изменение подхода к медицинской науке и практике [2].

Рабочая группа поставила своей целью подготовить серию статей, которые послужили бы продолжением материалов, опубликованных в *Canadian Medical Association Journal*, и позволили бы ознакомить врачей с принципами практического применения данных медицинской литературы. В написании статей принимали участие многие специалисты, работавшие вне стен университета МакМастер, однако особенно хочется отметить Скотта Ричардсона, Марка Уилсона, Роберта Хэйурда и Вирджинию Мойер. Итогом усилий Рабочей группы при активном содействии главного редактора журнала *JAMA* Драммонда Ренни, помощь которого поистине невозможно переоценить, стало появление на свет серии из 25 статей под общим названием "Путеводитель читателя медицинской литературы"; они публиковались в *JAMA* с 1993 по 2000 г. [3]. Книга, которую вы держите в руках, служит прямым продолжением этой серии.

Прошло не так много времени, и стало ясно, что принципы ДМ с одинаковым успехом могут применять в своей повседневной деятельности практически все медицинские работники (в том числе медицинские сестры, стоматологи, физиотерапевты, специалисты в области профессиональных заболеваний, мануальные терапевты, ортопеды и многие другие). Таким образом, наиболее правильно было бы говорить о "медицинской помощи, основанной на принципах доказательной медицины", поскольку этот термин включает в себя все аспекты ДМ, касающиеся практического использования научно обоснованных данных

при лечении больных. Однако книга адресована в первую очередь врачам; поэтому мы остановились на термине "доказательная медицина".

ДЛЯ КОГО ПРЕДНАЗНАЧЕНА ЭТА КНИГА?

Данная книга будет полезна любому врачу, который хотел бы научиться грамотно оценивать публикуемые в медицинской литературе данные и с наибольшей эффективностью использовать их для решения частных клинических проблем. Печатный вариант книги, включающий в себя основы ДМ, рассчитан на специалистов, нуждающихся в базовых знаниях. Этого уровня вполне достаточно для большинства студентов, ординаторов и врачей общей практики. В первой части они найдут множество клинических примеров, облегчающих понимание изложенного теоретического материала, а также описание методов поиска наиболее достоверных данных и их практического применения.

ОБ АВТОРСТВЕ

К новшествам настоящей книги следует отнести понятие авторства, которое в равной степени разделяют все 50 членов Рабочей группы. Большинство глав предваряется фамилиями авторов соответствующих статей в *JAMA*, поскольку именно эти авторы обычно готовили новый, пересмотренный текст статей для настоящего издания. После написания каждую главу тщательно обсуждали остальные члены Рабочей группы, не принимавшие непосредственного участия в ее создании. Нередко вклад последних был настолько велик, что их с полным правом можно считать соавторами. В других случаях имена членов Рабочей группы указаны сразу после авторов оригинальных статей. На стр. XI выражается благодарность всем членам Рабочей группы, участвовавшим в подготовке книги.

Данная книга — плод совместных усилий множества специалистов. Мы надеемся, что она в полной мере продемонстрирует преимущества использования на практике принципов ДМ.

Гордон Гайятт
Университет Мак Мастер

Литература

1. *Guyatt G.H.* Evidence-based medicine. ACP Journal Club 1991,114:A—16
2. *Evidence-based medicine working group.* Evidence-based medicine a new approach to the teaching of medicine. *JAMA* 1992;268:2420—2425.
3. *Guyatt G.H., Rennie D.* User's Guides to the Medical Literature (editorial) *JAMA* 1993,270:2096-2097.

Авторы

Барратт Александра (Alexandra Barratt, MBBS, MPH, PhD - Public Health and Community Medicine, University of Sydney, Sydney, NSW, Australia)

Басе Эрик (Eric Bass, MD, MPH — Division of Internal Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA)

Бачер Хейнер (Heiner Bucher, MD — Department of Internal Medicine, University Hospital Basel, Basel, Switzerland)

Боссют Патрик (Patrick Bossuyt, PhD — Clinical Epidemiology and Biostatistics, University of Amsterdam, Amsterdam, the Netherlands)

Гайятт Гордон (Gordon Guyatt, MD, MSc — Chair, Evidence-Based Medicine Working Group, Clinical Epidemiology and Biostatistics and Medicine, McMaster University, Hamilton, Ontario, Canada)

Глажиу Пол (Paul Glasziou, MBBS, PhD — Department of Social and Preventative Medicine, University of Queensland Medical School, Herston, QLD, Australia)

Грин Ли (Lee Green, MD, MPH — Department of Family Medicine, University of Michigan Medical Centre, Ann Arbor, Michigan, USA)

Гринхол Триша (Trisha Greenhalgh, MA, MD, FRCP, FRCGP - Unit for Evidence-Based Practice and Policy, Royal Free and University College Medical School, London, England, UK)

Дане Антонио (Antonio Dans, MD — Clinical Epidemiological Unit, University of the Philippines, Manila, Philippines)

Дане Леонилла (Leonila Dans, MD — Clinical Epidemiological Unit, University of the Philippines, Manila, Philippines)

Деворо П.Дж. (P.J. Devereaux, BSc, MD — Department of Medicine, McMaster University, Hamilton, Ontario, Canada)

Детски Алан (Alan Detsky, MD, PhD — Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada)

Джакомини Мита (Mita Giacomini, PhD — Centre for Health Economics and Policy Analysis, McMaster University, Hamilton, Ontario, Canada)

Джунипер Элизабет (Elizabeth Juniper, MCSP, MSc — Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada)

Драммонд Майкл (Michael Drummond, BSc, MCom, DPhil - Centre for Health Economics, University of York, York, England, UK)

ИрвигЛес (Les Irwig, MBBS, PhD, FFRPHM — Department of Public Health and Community Medicine, University of Sydney, Sydney, NSW, Australia)

Йешке Роман (Roman Jaeschke, MD, MSc — Department of Medicine, McMaster University, Hamilton, Ontario, Canada)

Камминг Роберт (Robert Cumming, MB BS, MPH, PhD, FAFPHM - Department of Public Health and Community Medicine, University of Sydney, Sydney, NSW, Australia)

Крэйг Джонатан (Jonathan Craig, FRACP, PhD — Department of Public Health & Community Medicine, Faculty of Medicine, University of Sydney, Sydney, NSW, Australia)

Кук Дебора (Deborah Cook, MD, MSc - Department of Medicine, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada)

Куш Регина (Regina Kunz, MD, MSc — Department of Nephrology, Chante Humboldt-University, Berlin, Germany)

Лачетти Кристина (Christina Lacchetti, MHS — Department of Clinical Epidemiology and Biostatistics and Medicine, McMaster University, Hamilton, Ontario, Canada)

Левайн Митчелл (Mitchell Levine, MD, MSc — Centre for Evaluation of Medicines, St Joseph's Hospital, Hamilton, Ontario, Canada)

Летельер Луз (Luz Letelier, MD — General Internal Medicine, Servicio de Medicina, Hospital Dr Sotero del Rio, U D A Universidad Catohca de Chile, Santiago, Chile)

Леунг Рэймонд (Raymond Leung, MDCM — Centres for Health Evidence, University of Alberta, Edmonton, Alberta, Canada)

Ли Хьюи (Нш Lee, MD, MSc, FRCPC - Department of Medicine, McMaster University, Hamilton, Ontario, Canada)

Лиймер Джероен (Jeroen Lijmer, MD, Clinical Epidemiology and Biostatistics, University of Amsterdam, Amsterdam, The Netherlands)

Лопасис Андреас (Andreas Laupacis, MD, MSc, Departments of Medicine and Epidemiology and Community Medicine, University of Ottawa, Ottawa, Ontario, Canada)

МакАлистер Финлей (Finlay McAlister, MD, MSc, FRCPC - Department of Medicine, University of Alberta, Edmonton, Alberta, Canada)

Макджин Томас (Thomas McGinn, MD — Primary Care Medicine, Mount Sinai Medical Center, New York, New York, USA)

МакКиббон Энн (Ann McKibbon, MLS — Department of Clinical Epidemiology and Biostatistics, HealthInformation Research Unit, McMaster University, Hamilton, Ontario, Canada)

Мид Морин (Maureen Meade, MD, FRCPC, MSc — Department of Medicine, McMaster University, Hamilton, Ontario, Canada)

Мойер Вирджиния (Virginia Moyer, MD, MPH — Department of Pediatrics, University of Texas, Houston, Texas, USA)

Монтори Виктор (Victor Monton, MD — Department of Medicine, Mayo Clinic and Foundation, Rochester, Minnesota, USA)

Нишикава Джим (Jim Nishikawa, MD, FRCPC — Departments of Clinical Epidemiology and Biostatistics and Medicine, McMaster University, Hamilton, Ontario, Canada)

Ньюмен Томас (Thomas Newman, MD, MPH — Departments of Epidemiology and Biostatistics, Pediatrics and Laboratory Medicine, University of California San Francisco, San Francisco, California, USA)

Нэйлор Дэвид (David Naylor, MD DPhil — Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada)

О'Брайен Берни (Bernie O'Brien, PhD — Centre for Evaluation of Medicines, St Joseph's Hospital, Hamilton, Ontario, Canada)

Оксман Эндриу (Andrew Oxman, MD, MSc — Health Services Research Unit, National Institute of Public Health, Oslo, Norway)

Проновост Питер (Peter Pronovost, MD, PhD — Anesthesiology and Critical Care Medicine, Johns Hopkins University, Baltimore, Maryland, USA)

Ренни Драммонд (Drummond Rennie, MD — Institute for Health Policy Studies, University of California San Francisco, San Francisco California, USA)

Ричардсон Скотт (Scott Richardson, MD — Audie L Murphy Memorial Veterans Hospital, San Antonio, TX, USA)

Рэндольф Адриенн (Adnenne Randolph, MD, MSc — The Children's Hospital, Boston, Massachusetts, USA)

Синклер Джек (Jack Sinclair, MD — Departments of Clinical Epidemiology and Biostatistics and Pediatrics, McMaster University, Hamilton, Ontario, Canada)

Стоклер Мартин (Martin Stockier MBBS, MSc, FRACP — Department of Medicine, University of Sydney, Sydney, NSW, Australia)

Строе Шэрон (Sharon Straus, MD — Department of Medicine University of Toronto and Mount Sinai Hospital, Toronto, Ontario, Canada)

Тагвелл Питер (Peter Tugwell, MD, MSc — Clinical Epidemiology Unit and Departments of Medicine and Epidemiology, University of Ottawa, Ottawa, Ontario, Canada)

Уайер Питер (Peter Wyer, MD — Department of Medicine, Columbia University College of Physicians and Surgeons, Pelham, New York, USA)

Уайятт Джереми (Jeremy Wyatt, MD — Knowledge Management Centre, University College London, London, England, UK)

Уилсон Марк (Mark Wilson, MD, MPH — Department of Medicine, Wake Forest University School of Medicine, Winston-Salem North Carolina, USA)

Уолтер Стивен (Stephen Walter, PhD — Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada)

Уэллс Джордж (George Wells, MSc, PhD — Clinical Epidemiology Unit and Departments of Medicine and Epidemiology, University of Ottawa, Ottawa, Ontario, Canada)

Хант Дерек (Dereck Hunt, MD — Health Information Research Unit, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada)

Хатала Роуз (Rose Hatala, MD, MSc, FRCPC — Department of Medicine, McMaster University Hamilton, Ontario, Canada)

Хейланд Дарен (Daren Heyland, MD, FRCPC, MSc — Medicine and Community Health and Epidemiology, Queen's University, Kingston, Ontario, Canada)

Холбрук Энн (Anne Holbrook, MD, PharmD, MSc — Centre for Evaluation of Medicines, Faculty of Sciences, McMaster University, Hamilton, Ontario, Canada)

Хэйнс Брайан (Brian Haynes, MD, MSc, PhD — Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada)

Хэйнс Тед (Ted Haines, MD, MSc - Departments of Clinical Epidemiology and Biostatistics, and Occupational Health Program, McMaster University, Hamilton, Ontario, Canada)

Хэйуорд Роберт (Robert Hayward, MD, MPH, FRCPC - Centre for Health Evidence, University of Alberta Edmonton, Alberta, Canada)

Хэслам Дэвид (David Haslam, MD, FRCPC — North Bay Psychiatric Hospital, North Bay, Ontario, Canada)

Шунманн Холгер (Holger Schbnemann, MD, PhD, MSc — Departments of Medicine, and Social and Preventive Medicine, University at Buffalo, Buffalo, New York, USA)

Эттиа Джон (John Attia, MD, MSc, PhD — Centre for Clinical Epidemiology and Biostatistics, Royal Newcastle Hospital, Newcastle, NSW, Australia)

Как пользоваться книгой

Г. Гайятт

В подготовке данной главы принимали также участие члены Рабочей группы по разработке и внедрению принципов доказательной медицины Д. Эттиа, Р Йешке, М Мид, Э. Оксман, Т Гринхол, Дж Синклер, Э. МакКиббон, Д. Кук и Э Басе

Данная книга, как и сама доказательная медицина, призвана облегчить принятие клинических решений. Авторы попытались показать, каким образом можно с наибольшей эффективностью использовать данные, опубликованные в различных источниках, для выбора оптимальных подходов к оказанию медицинской помощи. Таких источников великое множество: это оригинальные журнальные статьи, обзоры или рефераты материалов оригинальных исследований, практические рекомендации, а также хорошо известные или новые учебные руководства. Уже сейчас практически любую научную информацию можно найти в Интернете; в будущем он может остаться единственным доступом к некоторым источникам данных.

ЧАСТЬ 1. ОСНОВНЫЕ СВЕДЕНИЯ: ИСПОЛЬЗОВАНИЕ МЕДИЦИНСКОЙ ЛИТЕРАТУРЫ

Первая часть адресована всем, кому приходится иметь дело с изучением медицинской литературы (студентам, интернам, ординаторам или практикующим врачам), и содержит начальные сведения об эффективной работе с разнообразными источниками данных. Мы попытались сделать эту часть книги максимально простой и понятной. На наш взгляд, ее можно рассматривать в качестве краткого учебного пособия для студентов; лиц, обеспечивающих уход за больными; врачей, которые готовятся к получению "узкой" медицинской специальности или повышают свою квалификацию в рамках программы последипломного образования.

В этой части книги изложен систематический подход к изучению медицинской литературы, включающий в себя три основных этапа. Сначала следует определить степень достоверности любой новой информации, затем сделать выводы о предлагаемом подходе к ведению больных, после чего решить вопрос о том, каким образом можно использовать эту информацию на практике. На первом этапе врач оценивает достоверность представленных данных и вероятность наличия систематической ошибки, на втором этапе приходит к пониманию клинической сущности результатов исследования, а на третьем оценивает возможность их применения при лечении конкретного больного. Эти принципы символизируют три колонны, изображенные на обложке книги. Чтобы облегчить понимание материала, каждая глава начинается с описания возможной клинической ситуации, после чего рассматривается процесс поиска соответствующей информации. Кроме того, каждая глава содержит таблицу, в которой изложены критерии, которые необходимо применять на любом из трех вышеуказанных этапов.

Несмотря на краткость первой части, после знакомства с ней вы сможете с успехом использовать принципы доказательной медицины в своей работе. Вы научитесь:

- отличать более достоверные данные от менее достоверных;
- проводить в полном объеме критическую оценку и обобщение данных, а также соотносить преимущества и недостатки предполагаемого вмешательства перед принятием решения о тактике лечения;
- осуществлять поиск данных, предварительная оценка которых была проведена авторами обзоров и клинических рекомендаций;
- применять опубликованные данные в своей повседневной практике, в частности при лечении конкретного больного.

В настоящее время существует множество источников доказательной информации, уже прошедшей критическую оценку; большинство из них доступны в электронном виде. В ближайшие годы объем и качество этих источников существенно возрастут. В главе 2 *Поиск доказательной информации* излагаются критерии выбора оптимальных баз данных (т.е. содержащих информацию не только достоверную, но и применимую в определенных ситуациях) и методы работы с ними. После знакомства с главами 3—11 читатель сможет понять, как с наибольшей пользой для больного использовать найденные данные.

В сущности совсем не обязательно читать всю первую часть. Книга построена таким образом, что каждая глава более или менее самостоятельна и закончена. Если вас интересуют только принципы формулирования клинического вопроса и поиска соответствующих данных, вы

можете ограничиться изучением главы 2. Если вы ищете только материалы оригинальных исследований, посвященных методам лечения и уже включенных в систематические обзоры, необходимую информацию можно получить в разделах об оценке эффективности лечения и риске в главах 3—5, а также в разделе о систематических обзорах в главе 10. Чтобы избежать ненужных повторов, приводятся ссылки именно на ту главу, в которой уже приводились соответствующие пояснения.

ЧАСТЬ 2. БОЛЕЕ ДЕТАЛЬНЫЕ СВЕДЕНИЯ: ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ ПРИНЦИПОВ ДОКАЗАТЕЛЬНОЙ МЕДИЦИНЫ И ОБУЧЕНИЕ ЕЕ ОСНОВАМ

Вторая часть книги адресована врачам, которые хотят более профессионально использовать в своей работе принципы доказательной медицины. При этом печатный вариант второй части включает в себя сведения, необходимые для углубления уже полученных знаний, а вариант на компакт-диске дополнен данными, которые могут оказаться полезными преподавателям (компакт-диск прилагается к оригинальному англоязычному изданию книги).

Изучение материалов, включенных во вторую часть, позволит лучше понимать методологические аспекты проведения исследований, осваивать статистические методы, а также более эффективно помогать больным выбрать предпочтительное для них лечение, опираясь на количественные данные о преимуществах и недостатках медицинских вмешательств. При написании второй части мы ориентировались также на читателей, которые преподают основы доказательной медицины. Многие материалы представлены в виде интерактивного обсуждения с небольшой группой курсантов на семинаре или у постели больного и действительно были созданы на основе таких занятий

Как пользоваться второй частью книги? Прежде всего следует учитывать, что обе части построены по аналогичному принципу. Каждый раздел второй части включает в себя более детальное обсуждение тех вопросов, которые уже рассматривались в соответствующем разделе первой части. Так, раздел *Лечение и риск* содержит данные об исследованиях, посвященных экономической оценке эффективности лечения, программам скрининга и компьютерным системам, облегчающим принятие клинических решений. Материалы, изложенные во второй части

книги, выходят за рамки основ доказательной медицины и дополняют сведения, содержащиеся в первой части.

Каждый из вопросов, рассматривающихся во второй части, имеет ссылку на соответствующий раздел первой части. Например, в главе 2А на компакт-диске вы найдете более подробное обсуждение принципов, положенных в основу философской концепции доказательной медицины. В большинстве случаев дополнительные данные, включенные во вторую часть, напрямую связаны с положениями, представленными в первой. Так, в первой части книги подчеркивается, что при оценке результатов исследований, посвященных лечению, необходимо ответить на вопрос, насколько полно авторы оценили все значимые клинические исходы (как благоприятные, так и неблагоприятные). Одним из важных критериев оценки служит качество жизни, связанное с состоянием здоровья, в первой части приведены краткие сведения о его измерении. Во второй части можно найти подробное описание показателей качества жизни и детальное обсуждение способов их определения.

Многим покажется более удобным использование компакт-диска, когда один щелчок мыши позволяет перейти от любой главы первой части к соответствующей главе второй. Независимо от того, какой вариант книги вы предпочтете, вашу работу значительно упростит словарь терминов, в котором приведены определения многих рассматриваемых понятий. Мы надеемся, что выбранный нами способ изложения материала удовлетворит запросы любого читателя, желающего освоить принципы доказательной медицины.

Более полную информацию можно найти на сайте www.userguides.org

ЧАСТЬ I

Основные сведения: использование медицинской литературы

- 1 • ВВЕДЕНИЕ; ФИЛОСОФСКАЯ КОНЦЕПЦИЯ
ДОКАЗАТЕЛЬНОЙ МЕДИЦИНЫ
- 2 • ПОИСК ДОКАЗАТЕЛЬНОЙ ИНФОРМАЦИИ
- 3 • ЛЕЧЕНИЕ И РИСК; ВВЕДЕНИЕ
- 4 • ИССЛЕДОВАНИЯ, ПОСВЯЩЕННЫЕ
ЭФФЕКТИВНОСТИ ЛЕЧЕНИЯ
- 5 • ОЦЕНКА РИСКА
- 8 • ДИАГНОСТИКА
- 7 • ДИФФЕРЕНЦИАЛЬНАЯ ДИАГНОСТИКА
- 8 • ДИАГНОСТИЧЕСКИЕ МЕТОДЫ
- 9 • ПРОГНОЗ
- "10 • ОБОБЩЕНИЕ ДОКАЗАТЕЛЬСТВ
- 11 • ОТ ДОКАЗАТЕЛЬСТВ К ПРАКТИКЕ

Врачи прежде всего заинтересованы в постановке правильного диагноза и выборе оптимальных методов лечения. Они также должны избегать неблагоприятных воздействий и предоставлять больным сведения о возможном прогнозе. Первая часть данной книги позволяет читателю овладеть навыками работы с медицинской литературой, необходимыми для успешного решения этих задач.

Прежде чем воспользоваться имеющейся доказательной информацией, ее следует найти. В начале первой части книги излагается рекомендуемый подход к правильному формулированию клинических вопросов и поиску соответствующих исследований. За последние годы существенно возросли число и объем источников информации; при написании главы 2 *Поиск доказательной информации* учитывались самые современные на тот момент возможности поиска данных.

ОРИГИНАЛЬНЫЕ ИССЛЕДОВАНИЯ

Следует различать материалы оригинальных (первичных) исследований и обзоры, в которых объединены результаты нескольких таких исследований. Ознакомившись с методами поиска самой достоверной информации, читатель изучит подходы к критической оценке и практическому применению результатов оригинальных исследований. При оценке исследований, посвященных эффективности или безопасности лечения, используются похожие критерии; то же можно сказать об исследованиях, посвященных диагностическим методам или прогнозу. Эти взаимосвязи постоянно подчеркиваются в ходе обсуждения.

СИСТЕМАТИЧЕСКИЕ ОБЗОРЫ

Если по интересующей теме опубликованы систематические обзоры оригинальных исследований, им следует отдавать безусловное предпочтение. Врач, опирающийся в своей работе на принципы доказательной медицины, должен в первую очередь рассматривать именно такие обзоры. Для этого нужно владеть методами поиска систематических обзоров, критической оценки их методологического качества и применения полученных результатов в клинической практике. В первой части книги после обсуждения оригинальных исследований приводятся сведения о поиске и использовании систематических обзоров.

КЛИНИЧЕСКИЕ РЕКОМЕНДАЦИИ, ПОСВЯЩЕННЫЕ ТАКТИКЕ ЛЕЧЕНИЯ

Еще более эффективным подходом следует считать первоначальное рассмотрение не систематических обзоров, а клинических рекомендаций, посвященных тактике лечения. Лучшие из них, разрабатываемые на основе практических рекомендаций или анализа принятия решений, включают в себя самые достоверные данные; в таких рекомендациях обязательно учитываются предпочтения и жизненные ценности больных. Однако следует помнить, что методологическое качество клинических рекомендаций может быть низким (такие рекомендации следует игнорировать) или высоким (таким рекомендациям нужно следовать). В главе 11 *От доказательств к практике* показано, каким образом можно оценить методологическое качество практических рекомендаций или анализа принятия решений, понять их недостатки и с наибольшей эффективностью использовать эти материалы в повседневной практике

1

ВВЕДЕНИЕ:
ФИЛОСОФСКАЯ КОНЦЕПЦИЯ
ДОКАЗАТЕЛЬНОЙ МЕДИЦИНЫ

Г. Гайятт, Б. Хэйнс, Р. Йешке, Д. Кук,
Т. Гринхол, М. Мид, Л. Грин, К.Д. Нэйлор,
М. Уилсон, Ф. Макалистер
и У.С. Ричардсон

В подготовке данной главы принимали также участие члены рабочей группы по разработке и внедрению принципов доказательной медицины В. Монтори и Х. Бачер

В ЭТОМ РАЗДЕЛЕ

Два **фундаментальных принципа доказательной медицины**

Для принятия клинического решения недостаточно только доказательной информации

Достоверность данных, полученных в ходе разных исследований, различна

Клинические навыки, **гуманизм**, социальная ответственность и доказательная **медицина**

Другие проблемы **доказательной медицины**

ВОЗМОЖНАЯ СИТУАЦИЯ

Кто прав?

В больничном кафе старший ординатор, профессор-консультант, младший и старший врачи беседовали о доказательной медицине.

— Доказательная медицина — это поистине революционный шаг вперед, — экзальтированно воскликнула ординатор и начала перечислять фундаментальные изменения, которые та привнесла в медицинскую практику.

— Звучит убедительно, — отозвался профессор.

— Постойте, — столь же горячо возразил младший врач и изложил иное мнение, заключающееся в том, что доказательная медицина всего лишь пополнила некоторыми методами традиционный арсенал врача.

— Вы привели достаточно веские доводы, — заметил профессор.

— Здесь что-то не так, — обратился старший врач к своему уважаемому коллеге, — высказанные мнения диаметрально противоположны. Одно из них должно быть ошибочным.

Профессор внимательно посмотрел на озадаченного врача и с легкой улыбкой произнес:

— Если задуматься, то и вы правы.

Доказательная медицина (ДМ) призвана облегчить принятие клинических решений [1]. В 1992 г. авторы данной главы охарактеризовали ДМ как изменение подхода к медицинской науке и практике [2], использование в процессе принятия врачебных решений не только интуиции, несистематизированного клинического опыта и знания основ патофизиологии. Концепция ДМ подразумевает также тщательный поиск и критическую оценку данных, полученных в ходе клинических исследований; для этого недостаточно обладать профессиональными навыками и здравым смыслом, необходимо использовать также целый ряд специально разработанных правил. Кроме того, в отличие от традиционного подхода ДМ придает гораздо меньшее значение "авторитетному мнению".

В целом подобная концепция ДМ справедлива и сегодня. Но, как видно из приведенной в самом начале беседы, мир в наши дни зачастую слишком сложен, чтобы согласиться с единственным суждением о том или ином явлении. В данной главе изложен современный взгляд на ДМ. Рассматриваются два основополагающих принципа ДМ: о важности принятия клинических решений с учетом многих факторов и о различной достоверности имеющихся данных. Затем обсуждаются условия, порой выходящие за рамки ДМ, но не менее важные для оказания высококачественной медицинской помощи, и проблемы, с которыми ДМ будет сталкиваться в новом тысячелетии.

ДВА ФУНДАМЕНТАЛЬНЫХ ПРИНЦИПА ДОКАЗАТЕЛЬНОЙ МЕДИЦИНЫ

ДМ представляет собой новый подход к оказанию медицинской помощи с использованием двух фундаментальных принципов. Во-первых, для принятия клинического решения недостаточно только доказательной информации. Делая выбор между существующими вмешательствами, врач всегда должен учитывать целый ряд факторов: соотношение пользы и риска, удобство того или иного метода для больного, затраты на обследование и лечение, а также предпочтения и жизненные ценности больного [1]. Во-вторых, при принятии клинического решения следует помнить о том, что достоверность данных, полученных в исследованиях с разной структурой, может существенно различаться.

1. ДЛЯ ПРИНЯТИЯ КЛИНИЧЕСКОГО РЕШЕНИЯ НЕДОСТАТОЧНО ТОЛЬКО ДОКАЗАТЕЛЬНОЙ ИНФОРМАЦИИ

Представим себе больную с хроническим болевым синдромом, который обусловлен наличием рака в терминальной стадии. Больная уже

смирилась со своим состоянием, сделала все необходимые распоряжения, попрощалась с близкими и теперь хотела бы получать только паллиативную терапию. В этот момент у нее возникает пневмококковая пневмония. Имеется множество весьма достоверных данных о том, что применение антибиотиков снижает частоту развития осложнений и смертность от пневмококковой пневмонии. Однако большинство врачей согласятся с тем, что в описываемой ситуации самые веские доводы не сделают обязательным назначение антибиотиков. Действительно, такая терапия могла бы уменьшить выраженность симптомов пневмонии и продлить жизнь, но сама больная предпочитает быструю и естественную смерть.

Рассмотрим еще один пример. Пневмококковая пневмония возникает у 85-летнего больного с тяжелой деменцией, недержанием мочи и мышечными контрактурами, который лишен дара речи, не имеет семьи и друзей, испытывает постоянные страдания. В такой ситуации многие врачи сочтут применение антибиотиков нецелесообразным, однако другие могут с этим не согласиться. Как и в предыдущем случае, наличие данных об эффективности лечения вовсе не означает его обязательного назначения. При принятии решения необходимо оценить соотношение риска и пользы; представления отдельных врачей и больных о клинической значимости преимуществ лечения в каждом конкретном случае могут существенно различаться, поэтому и решения будут разными.

Наконец, представим себе, что пневмококковая пневмония развивается у 30-летней женщины без сопутствующих заболеваний, имеющей двух детей. Никто из врачей не усомнится в необходимости назначения ей антибиотиков. Однако это не означает, что в процессе принятия решения можно не учитывать другие факторы, влияющие на выбор оптимальной тактики. В данном случае их анализ достаточно прост, поскольку мнение большинства врачей во многом совпадает с пожеланиями больной, а польза от применения антибиотиков существенно перевешивает риск, связанный с таким лечением.

В современной медицинской практике принятие решения о выборе того или иного вмешательства нередко отражает представления врачей или общества о целесообразности затрат на лечение [2]. Так, при остром инфаркте миокарда можно применять тканевый активатор плазминогена или стрептокиназу, при преходящем нарушении мозгового кровообращения — клопидогрел или аспирин. Результаты крупных клинических Испытаний свидетельствуют о том, что при обоих заболеваниях более Дорогостоящие вмешательства характеризуются большей эффективностью.

стью. Однако многие специалисты рекомендуют в качестве препаратов первого ряда более дешевые и менее действенные лекарственные средства, поскольку полагают, что высвободившиеся ресурсы здравоохранения можно с большей пользой потратить на другие нужды. Таким образом, в процессе принятия подобного решения на одной чаше весов оказываются уровень смертности и число предотвращенных инсультов, а на другой — стоимость лечения.

Под важными факторами, учитываемыми в процессе принятия клинического решения, в данной главе подразумеваются возможные благоприятные или неблагоприятные последствия такого решения для конкретного больного или общества в целом. Точная количественная оценка и сопоставление пользы и риска, связанных с медицинскими вмешательствами, лежат в основе ДМ и существенно облегчают выбор оптимальной врачебной тактики.

Попытки использовать этот подход при принятии любого решения выявили наше неумение определять и соотносить факторы, важные как для общества, так и конкретного больного. В разработке научных методов оценки предпочтений больных большую роль сыграли специалисты в области экономики здравоохранения [3, 4]. В некоторых пособиях, призванных облегчить принятие клинических решений, имеются косвенные упоминания предпочтений и жизненных ценностей больного; предполагается, что больной, правильно представляющий возможные благоприятные и неблагоприятные последствия применения разных вмешательств, примет решение, которое с большой долей вероятности будет отражать его предпочтения [5]. Такие разработки можно считать неплохим началом. Однако остается еще много неясных вопросов, касающихся выявления и учета предпочтений больного в процессе принятия решений, на который врачу зачастую отведено совсем немного времени. В ходе дальнейшего развития ДМ этой проблеме следует уделить самое пристальное внимание.

2, ДОСТОВЕРНОСТЬ ДАННЫХ, ПОЛУЧЕННЫХ В ХОДЕ РАЗНЫХ ИССЛЕДОВАНИЙ, РАЗЛИЧНА

Какими данными оперирует ДМ⁹ Можно предложить довольно широкое определение: любое эмпирическое наблюдение о наличии связи между какими-либо явлениями следует рассматривать в качестве потенциального доказательства. Таким образом, источником данных могут служить и несистематические наблюдения отдельных врачей, и результаты исследований, в которых регистрировались только физиологиче-

ские параметры. Некоторые клинические наблюдения заставляют по-новому, более пристально взглянуть на проблему; опытные врачи с должным вниманием относятся к мнению специалистов относительно разных аспектов диагностики и течения заболеваний, взаимоотношений с больными и коллегами. Некоторые клинические наблюдения могут быть весьма поучительны, хотя в медицинской литературе публикуются довольно редко.

При этом ценность несистематических наблюдений отдельных врачей ограничена небольшим числом больных и, что еще более важно, индивидуальными особенностями, присущими мыслительной деятельности каждого человека [6].

Предположения о влиянии определенного вмешательства на клинически значимый исход, которые были сделаны в ходе исследований с регистрацией только физиологических параметров, обычно верны, однако в отдельных случаях они совершенно не соответствуют действительности.

Учитывая ограничения, присущие несистематическим клиническим наблюдениям и результатам вышеуказанных исследований, эксперты в области ДМ предлагают подразделять все имеющиеся данные в зависимости от уровня их доказательности. В табл. 1 разные исследования, посвященные оценке эффективности лечебных вмешательств, расположены в порядке убывания уровня доказательности получаемых данных; исследования, посвященные оценке информативности диагностических методов или клиническому прогнозированию, пришлось бы расположить в совершенно ином порядке. В этом списке результаты клинических исследований по уровню своей доказательности стоят выше несистематических наблюдений, поскольку в ходе клинических исследований применяются методы, позволяющие избежать появления ошибочных данных или свести такую вероятность к минимуму. Поскольку очень немногие лечебные вмешательства (если таковые вообще имеются) эффективны у всех без исключения больных, в испытаниях методов лечения следовало бы включать участников, по своим характеристикам идеально соответствующих больным, которым планируется их назначать. Кроме того, врачи могут неправильно оценивать результаты открытых клинических испытаний из-за особенностей естественного течения заболеваний, эффекта плацебо, наличия систематических ошибок, связанных с ожиданиями больных и медицинских работников или со стремлением больного не разочаровать лечащего врача

Таблица 1. Исследования с различной структурой, посвященные оценке эффективности лечебных вмешательств и расположенные в порядке убывания достоверности их результатов

РКИ с участием отдельно взятого больного (РКИ "N=1")

Систематические обзоры РКИ

Отдельные РКИ

Систематические обзоры обсервационных исследований, в которых оценивались клинические исходы, значимые для больных

Отдельные обсервационные исследования, в которых оценивались клинические исходы, значимые для больных

Исследования, в которых оценивались только физиологические параметры (артериальное давление, величина сердечного выброса, переносимость физических нагрузок, минеральная плотность костной ткани и т.д.)

Несистематические клинические наблюдения

В исследованиях, включающих отдельных больных, для снижения вероятности возникновения систематической ошибки используется тот же подход, что и в обычных клинических испытаниях с большими выборками [7]. В РКИ с участием отдельно взятых больных (РКИ "N=1") каждый больной в течение определенного времени получает изучаемое лечебное вмешательство, а затем на протяжении такого же срока — плацебо или какое-либо другое вмешательство. При этом ни врачи, ни больные не знают о характере вмешательства, последовательность чередования таких периодов определяется рандомизированно; в течение каждого из них больные количественно оценивают выраженность симптомов. Исследование продолжается до тех пор, пока больной и врач не сделают вывод о наличии (или отсутствии) преимуществ изучаемого вмешательства. Подобные РКИ достаточно легко осуществимы [8, 9] и позволяют получить точные данные об эффективности вмешательства у отдельных больных; на основе результатов этих РКИ в ряде случаев может быть пересмотрена стандартная схема лечения [10].

Используя любые другие источники данных, врач должен помнить о возможных различиях в характеристиках участников исследований и его больных. Это неизбежно снижает достоверность данных об эффективности вмешательства; при этом возникает вопрос о применимости результатов испытания в конкретной клинической ситуации. Вместе с тем уровень достоверности данных может быть достаточно высоким, если они получены в ходе систематического обзора однородных РКИ надлежащего методологического качества. Результаты единственного РКИ, пусть даже крупного и включающего больных с разными характеристиками, менее достоверны (см. табл. 1). В обсервационных исследованиях

возможна недооценка эффекта вмешательства, которую трудно предусмотреть [11, 12], поэтому их результаты менее надежны, чем результаты РКИ. Наименее достоверны данные, полученные в ходе исследований с регистрацией только физиологических параметров, и несистематические клинические наблюдения.

Такая иерархия клинических исследований не претендует на абсолютную точность. Если эффект вмешательства достаточно выраженный и постоянный, результаты обсервационных исследований могут оказаться более убедительными, нежели данные, полученные в большинстве РКИ. Например, данные об эффективности применения инсулина при диабетическом кетоацидозе или об эффективности эндопротезирования тазобедренного сустава при тяжелом остеоартрите были получены в ходе обсервационных исследований и характеризуются чрезвычайно высокой достоверностью. Однако при стойкой противоречивости результатов РКИ и обсервационных исследований следует задуматься о достоверности последних. Вопрос о том, насколько врач может полагаться на результаты обсервационных исследований в отсутствие данных, полученных из других источников, остается открытым. Интерес к этой проблеме отнюдь не праздный, поскольку большая часть данных о побочных эффектах вмешательств получена именно в обсервационных исследованиях.

Предлагаемая классификация упрощает действия врача при решении той или иной проблемы: он всегда должен искать и применять на практике самые достоверные из имеющихся данных. Сама концепция иерархии клинических исследований заставляет усомниться в любом заявлении о том, что данные об эффективности или неэффективности вмешательства отсутствуют. Такие данные могут быть малоубедительными (несистематические наблюдения отдельных врачей или результаты исследований с регистрацией только физиологических параметров, имеющие косвенное отношение к проблеме), но не должны игнорироваться.

Ниже кратко охарактеризованы дополнительные навыки и черты, которыми врач должен обладать для оказания наиболее эффективной медицинской помощи, а также связь между этими характеристиками врача и принципами ДМ.

КЛИНИЧЕСКИЕ НАВЫКИ, ГУМАНИЗМ, СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ
И ДОКАЗАТЕЛЬНАЯ МЕДИЦИНА

Выбор оптимального решения той или иной клинической проблемы, проводимый на основе принципов ДМ, может быть успешным только при правильной формулировке вопроса. Один из авторов этой книги, врач стационара, незадолго перед важным докладом обнаружил у себя

на губе ссадину. Чтобы уточнить показания к применению ацикловира, следующие два часа он посвятил поиску самых достоверных данных, изучая результаты доступных РКИ. Так и не придя к окончательному заключению, он начал обсуждать свои сомнения с коллегой, опытным стоматологом, но та быстро положила конец дискуссии, воскликнув: "Дорогой мой, но это же не герпес!".

Этот пример свидетельствует о необходимости постановки правильного диагноза до начала поиска данных и их практического использования, о важности клинического опыта и возможности ошибочных суждений. Умение собрать анамнез, обследовать больного и четко сформулировать клинический вопрос достигается только при наличии соответствующей теоретической подготовки и практических навыков. Врачи применяют методы ДМ (например, рассчитывают отношение правдоподобия положительного или отрицательного результата диагностического теста) при оценке данных анамнеза и общего осмотра. Клинический опыт необходим и для того, чтобы определить все имеющиеся способы лечения конкретного заболевания, прежде чем искать данные об эффективности и безопасности каждого из них.

И наконец, клинический опыт помогает выявить факторы, которые могут повлиять на обобщаемость найденных данных и их применимость у конкретного больного. Выше уже отмечалось, что почти всегда (за исключением случаев, когда проводится РКИ "N=1") врач вынужден применять результаты, полученные у других больных, в конкретной клинической ситуации. При этом он должен оценить, каким образом различия в самом процессе лечения (обусловленные, например, более низкой квалификацией местных хирургов или высокой вероятностью того, что больной не будет соблюдать предписанную схему лечения), в оснащенности специальным оборудованием, позволяющим вести постоянное тщательное наблюдение за состоянием больного, или в характеристиках больного (таких, как возраст, наличие сопутствующих заболеваний или одновременное применение других препаратов) могут повлиять на соотношение пользы и риска, о котором сообщалось в материалах исследования.

Использование на практике методов ДМ необходимо, но недостаточно для оказания высококачественной медицинской помощи. Наряду с клиническим опытом врач должен обладать способностью к состраданию, умением внимательно слушать, обширными познаниями в гуманитарных и социальных науках. Все это позволяет не только поставить диагноз, но и понять больного, его переживания, личностные и культурные особенности. Подобный подход может по-разному проявляться при использовании принципов ДМ в повседневной практике. В некото-

рых случаях учет жизненных ценностей больного в процессе принятия важного клинического решения потребует тщательного взаимного обсуждения всех благоприятных и неблагоприятных последствий (польза, риск для здоровья или просто неудобство) применения разных вмешательств. Порой к такой беседе следует привлекать и членов семьи больного. Однако в ряде ситуаций (например, при обсуждении с пожилым мужчиной вопроса о необходимости скринингового обследования с определением концентрации простатспецифического антигена в крови) привлечение других членов семьи может идти вразрез с принятыми социальными и культурными нормами.

Нередко подробное обсуждение преимуществ и недостатков методов лечения вызывает негативную реакцию больных, которые считают такой подход попыткой переложить ответственность за решение на их плечи [13]. В тех случаях, когда больной желает, чтобы выбор того или иного вмешательства остался за врачом, последний должен принять ответственное решение, которое наилучшим образом соответствует жизненным ценностям и предпочтениям больного. Чтобы понять, какой вариант принятия решения и какой способ представления информации в большей степени соответствуют характеру больного, врач должен уметь слушать его и видеть в нем личность [14, 15]. Для ДМ, как для медицины в целом, всегда будет актуальным сочетание самых современных достижений науки и проверенного временем искусством врачевания.

В идеале врачи, владеющие методами ДМ и исповедующие принципы гуманизма, должны эффективно защищать интересы больных как при оказании непосредственной медицинской помощи, так и при решении общих вопросов здравоохранения. Многие врачи считают своей задачей лишь применение того или иного вмешательства. Даже решая вопрос о выборе методов профилактики, они ориентируются на коррекцию отдельных особенностей стиля жизни и поведения конкретного больного. Подобный подход представляется слишком узким.

Результаты обсервационных исследований свидетельствуют о наличии сильной постоянной связи между социально-экономическим положением больного и состоянием его здоровья. Здоровье населения в большей степени зависит от разницы в уровне дохода отдельных групп, чем от благосостояния общества в целом. Другими словами, общее состояние здоровья населения в более бедных странах с относительно равномерным распределением благ будет выше, чем в более богатых странах, в которых существует четкое разделение общества на богатых и бедных. Поэтому врач, заботящийся о здоровье своих больных и общества в целом, должен задумываться о способах уменьшения числа нуждающихся.

В ходе обсервационных исследований выявлена сильная постоянная связь между уровнем загрязнения окружающей среды и риском развития заболеваний органов дыхания и сердечно-сосудистой системы. При лечении больных с хроническими обструктивными заболеваниями легких врач убеждает их в необходимости отказа от курения. Но должен ли врач помнить также и об опасности, связанной с вдыханием загрязненного воздуха? Скорее всего, должен.

ДРУГИЕ ПРОБЛЕМЫ, С КОТОРЫМИ СТАЛКИВАЕТСЯ ДОКАЗАТЕЛЬНАЯ МЕДИЦИНА

В 1992 г. были определены основные навыки, которыми должен обладать каждый врач, стремящийся использовать в своей работе принципы ДМ. К таким навыкам отнесены умение четко определить проблему больного и характер информации, необходимой для ее разрешения; провести эффективный поиск данных, опубликованных в медицинской литературе; выбрать самые высококачественные исследования по интересующей проблеме; оценить уровень достоверности полученных в них результатов, руководствуясь специальными критериями ДМ; выбрать необходимые клинические данные и использовать их при лечении больного [1]. Однако следует признать, что сегодня этих навыков уже недостаточно. Для эффективной работы требуется умение вовлечь больного в процесс принятия решения и понимание того, как его жизненные ценности и предпочтения могут повлиять на соотношение пользы и риска от планируемого вмешательства.

За последние 10 лет так и не решена одна из основных проблем, препятствующих широкому внедрению принципов ДМ в клиническую практику. Речь идет о постоянной нехватке времени. К счастью, развитие информационных технологий позволило во много раз ускорить процесс поиска необходимых данных. При этом можно пользоваться следующей классификацией источников информации, распределенных в порядке возрастания их достоверности (англоязычному читателю предлагается пользоваться мнемоническим правилом 4S: Studies, Systematic reviews, Synopses, Systems): материалы одного или нескольких отдельных исследований; систематический обзор всех проведенных исследований по соответствующей теме; краткие обзоры, включающие результаты отдельных исследований и/или систематических обзоров; системные источники информации. Под последними подразумеваются статьи, подготовленные на основании нескольких кратких обзоров и посвященные опреде-

ленной медицинской проблеме (например, обследование и лечение больных с острым кровотечением из верхних отделов желудочно-кишечного тракта или амбулаторное ведение больных сахарным диабетом; табл. 2).

Принципы ДМ все чаще используются при отборе и обобщении данных, получаемых из любого вышеперечисленного источника (см. главу 2 *Поиск доказательной информации*)

Данная книга посвящена в первую очередь принятию решений относительно ведения конкретного больного. Однако принципы ДМ можно использовать также при формировании политики здравоохранения на государственном [16] и региональном уровне или управлении отдельными лечебно-профилактическими учреждениями. Независимо от уровня, на котором принимается решение, подобный подход облегчает достижение поставленной цели с минимальными затратами ресурсов. Однако следует отметить, что под знаменем ДМ и при весьма вольной трактовке ее принципов проводились в жизнь многочисленные реформы — от необоснованного снижения стоимости медицинских услуг до внедрения чрезвычайно дорогостоящих, но малоприбыльных или даже убыточных технологий.

При формировании политики здравоохранения требуется решать множество проблем; при этом приходится учитывать значительно большее число факторов, чем при лечении отдельного больного. Следует ли про-

Таблица 2. Источники доказательное информации

Материалы отдельных исследований	При предварительном поиске отбираются материалы только тех исследований, проблематика которых наиболее близка к изучаемой, а структура позволяет свести к минимуму вероятность наличия систематических ошибок и обеспечивает наибольшую достоверность полученных результатов
Систематические обзоры	Представляют собой наиболее полный свод доказательной информации по определенному клиническому вопросу
Краткие обзоры	Содержат основные методологические характеристики и результаты отдельных исследований или систематических обзоров, что позволяет применить полученные данные при лечении конкретного больного
Системные источники информации	К ним относятся практические руководства, алгоритмы принятия клинических решений или справочники доказательной медицины, которые посвящены определенной медицинской проблеме и содержат максимальное количество информации, необходимой для лечения конкретного больного

сто перераспределять средства ограниченного бюджета здравоохранения в зависимости от изменения ситуации? Стоит ли урезать финансирование медицинских служб, вводя при этом более низкие налоги для физических лиц и льготное медицинское обслуживание корпораций? Как следует относиться к результатам многочисленных обсервационных исследований, указывающим на то, что социальные и экономические факторы влияют на здоровье населения в большей степени, чем качество медицинской помощи? Как быть с противоречиями между интересами отдельных лиц и общества, членами которого они являются? Обсуждение этих вопросов относится к сфере политики здравоохранения, однако неизбежно влияет и на принятие решения на уровне конкретного больного.

Литература

1. *Haynes R.B., Sackett R.B., Gray J.M.A., Cook D.C., Guyatt G.H.* Transferring evidence from research into practice. I: the role of clinical care research evidence in clinical decisions. ACP Journal Club. Nov-Dec 1996;125A-14-15
2. *Napodano R.J.* Values in Medical Practice. New York, NY: Human Sciences Press; 1986.
3. *Drummond M.F., Richardson W.S., O'Brien B., Levine M., Heyland D.K., for the Evidence-Based Medicine Working Group.* Users' Guides to the Medical Literature XIII. How to use an article on economic analysis of clinical practice. A. Are the results of the study valid? JAMA 1997;277:1552-1557
4. *Feeny D.H., Furlong W., Boyle M., Torrance G.W.* Multi-attribute health status classification systems: health utilities index. Pharmacoeconomics 1995;7490-502.
5. *O'Connor A.M., Rosioni A., Fiset V., et al.* Decision aids for patients facing health treatment or screening decisions: systematic review. BMJ 1999;319.731-734
6. *Nisbett R., Ross L.* Human Inference. Englewood Cliffs, NJ: Prentice-Hall; 1980.
7. *Guyatt G.H., Sackett D.L., Taylor D.W., et al.* Determining optimal therapy: randomized trials in individual patients. N Engl J Med 1986;314.889-892.
8. *Guyatt G.H., Keller J.L., Jaeschke R., Rosenbloom D., Adachi J.D., Newhouse M.T.* The n-of-1 randomized controlled trial: clinical usefulness. Our three-year experience. Ann Intern Med 1990;112.293-299. PART 1: THE BASICS 11 Copyright © 2002 by the American Medical Association.
9. *Larson E.B., Ellsworth A.J., Oas J.* Randomized clinical trials in single patients during a 2-year period. JAMA 1993;270:2708-2712.
10. *Mahon J., Laupacis A., Donner A., Wood T* Randomised study of n of 1 trials versus standard practice BMJ 1996;312-1069-1074
11. *Guyatt G.H., DiCenso A., Farewell V., Willan A., Griffith L.* Randomized trials versus observational studies in adolescent pregnancy prevention. J Clin Epidemiol 2000;53:167-174
12. *Kunz R., Oxman A.D.* The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. BMJ 1998;317 1185-1190.
13. *Sutherland H.J., Llewellyn-Thomai H.A., Lockwood G.A., Trichler D.L., Till J.E.* Cancer patients' their desire for information and participation in treatment decisions. J R Soc Med 1989;82:260-263.
14. *Greenhalgh T.* Narrative based medicine: narrative based medicine in an evidence based world. BMJ 1999;318:323-325.
15. *Greenhalgh T., Hurwitz B.* Narrative based medicine: why study narrative? BMJ 1999;318-48-50.
16. *Muir Gray F.A., Haynes R.B., Sackett D.L., Cook D.J., Guyatt G.H.* Transferring evidence from research into practice, III. developing evidence-based clinical policy. ACP Journal Club 1997;A14.

2

ПОИСК ДОКАЗАТЕЛЬНОЙ ИНФОРМАЦИИ

А. МакКиббон, Д. Хант, У.С. Ричардсон,
Р. Хэйуорд, М. Уилсон, Р. Йешке, Б. Хэйнс,
П. Уайер, Дж. Крэйг, Г. Гайатт

В подготовке данной главы принимали также участие
члены Рабочей группы по разработке и внедрению
принципов доказательной медицины П. Боссюг,
Т. Гринхол, Ш. Строе и Д. Кук

В ЭТОМ РАЗДЕЛЕ

Способы использования медицинской литературы

Общие и частные вопросы, интересующие читателей

Просмотр медицинской периодики и поиск данных для решения конкретной проблемы

Формулирование клинического вопроса

Пример 1. Оптимальный уровень артериального давления (АД), который следует поддерживать у больных с артериальной гипертонией на фоне сахарного диабета

Пример 2. Тактика ведения больных при предполагаемой нестабильной стенокардии

Пример 3. Тактика ведения больных при плоскоклеточном раке легких

Поиск ответа

Определение типа вопроса

Целесообразен ли поиск данных в медицинской литературе?

Источники доказательной информации

Выбор оптимального источника данных

Медицинские электронные базы данных, в которые включаются только материалы, отвечающие определенным критериям методологического качества

Медицинские электронные базы данных, в которых при включении материалов не используются критерии методологического качества

Практическое применение найденной доказательной информации

СПОСОБЫ ИСПОЛЬЗОВАНИЯ МЕДИЦИНСКОЙ ЛИТЕРАТУРЫ

В настоящей главе мы рассмотрим, как можно использовать огромное количество данных, опубликованных в медицинской литературе. Вы убедитесь, что эти способы в корне отличаются от тех, которые бытуют среди студентов-медиков.

ОБЩИЕ И ЧАСТНЫЕ ВОПРОСЫ, ИНТЕРЕСУЮЩИЕ ЧИТАТЕЛЕЙ

Студенты-медики, особенно на младших курсах, довольно редко обращаются к оригинальным медицинским статьям. Причин тому несколько, однако основная заключается в том, что им не требуется брать на себя ответственность за ведение больных и решать связанные с этим проблемы. Даже при использовании учебных программ, базирующихся на решении конкретных задач, студентов в первую очередь интересуют нормальная физиология человека и общие патофизиологические механизмы развития отдельных заболеваний или состояний. После знакомства с основными принципами наступает черед изучения прогноза, существующих методов диагностики и лечения. И наконец, при столкновении с проблемой, значимой для конкретного больного, студенты начинают интересоваться такими вопросами, как, например, особенности течения сахарного диабета, причины возникновения полиурии у данного больного и возможная тактика лечения.

Напротив, опытные врачи, которым ежедневно приходится брать на себя ответственность за принимаемые клинические решения, задаются вопросами совсем другого характера. Их скорее интересуют не принципы диагностики того или иного заболевания, а интерпретация результатов конкретного диагностического теста; не общий прогноз при каком-либо хроническом заболевании, а перспективы конкретного больного, не перечень всех существующих методов лечения, а сравнительные преимущества и недостатки двух альтернативных вариантов.

Вопросы, возникающие у студентов, можно отнести к общим вопросам, а у практикующих врачей — к частным. В подавляющем большинстве случаев грамотно сформулировать частные вопросы можно лишь при четком понимании общих вопросов.

В первый день клинической практики студент будет задаваться преимущественно вопросами общего характера. Однако в дальнейшем, определяя способ лечения своего первого больного, он неизбежно столкнется и с частными клиническими вопросами. Точно так же опытному "Рачу, который в достаточном объеме владеет теоретическими знаниями

ми, необходимыми для решения привычных задач, может понадобиться дополнительная информация общего характера. Чаще всего это происходит при появлении нового, неизвестного ранее заболевания или состояния (всего два десятилетия назад квалифицированные врачи спрашивали друг друга, что такое синдром приобретенного иммунодефицита), нового метода диагностики или лечения (врач, например, должен представлять себе принцип полимеразной цепной реакции или механизм действия ингибиторов циклооксигеназы-2). На любом из этапов обучения, при любом опыте работы умение очертить круг общих вопросов, имеющих отношение к тому или иному заболеванию, определяет способность четко формулировать частные вопросы ведения конкретного больного.

На рис 1 представлено соотношение числа частных и общих вопросов, изменяющееся по мере приобретения врачом профессионального опыта (от почти исключительно общих вопросов у начинающих до почти исключительно частных у квалифицированных специалистов). Данная книга призвана помочь врачу использовать медицинскую литературу для того, чтобы искать и находить ответы на частные клинические вопросы.



Рис. 1. Соотношение числа общих и частных вопросов в зависимости от профессионального опыта врача.

ПРОСМОТР МЕДИЦИНСКОЙ ПЕРИОДИКИ И ПОИСК ДАННЫХ ДЛЯ РЕШЕНИЯ КОНКРЕТНОЙ ПРОБЛЕМЫ

Обычно врачи выписывают сразу несколько (иногда довольно много) медицинских журналов, в которых могут быть опубликованы статьи, так или иначе относящиеся к сфере их профессиональной деятельности. Материалы для прочтения обычно выбирают после беглого знакомства с оглавлением.

Такой подход к поиску информации неэффективен; его результаты, как правило, неудовлетворительны. Представьте себе стопку журналов, которые пылятся на столе в ожидании момента, когда врач выкроит время для их просмотра. Скорее всего он прекрасно понимает, что высоким методологическим качеством и практической ценностью обладают менее 10% статей, публикуемых в основных медицинских журналах. Времени для просмотра литературы как не было, так и нет; гора журналов растет с пугающей быстротой. В один прекрасный момент скопившаяся периодика отправляется в шкаф или корзину для мусора, и все начинается сначала.

Подобная ситуация несколько утрирована, но хорошо знакома большинству практикующих врачей. Каждый из них, вне зависимости от накопленного опыта, нередко чувствует себя бессильным перед лавиной медицинской литературы. Доказательная медицина (ДМ) предлагает два способа решения этой проблемы.

Во-первых, при поиске данных рекомендуется использовать журналы, содержащие предварительно обработанную информацию. Возможно, наиболее действенным способом следует считать просмотр только этих журналов. Так, в журнале ACP Journal Club (www.acronline.org/acprjc/jcmenu.htm) публикуются структурированные рефераты статей по внутренним болезням и общемедицинским вопросам, отвечающих критериям высокого методологического качества и клинической значимости. Особенности разных источников информации, содержащих предварительно обработанную информацию, будут детально обсуждаться ниже.

В подобных источниках до сих пор не освещаются многие вопросы, относящиеся к узким областям медицины. На наш взгляд, это проблема будет со временем решена (по крайней мере относительно публикаций По основным медицинским специальностям). Пока же врач может самостоятельно оценивать методологическое качество и клиническую значимость статей в интересующих его журналах. Большинство публикаций служат авторам средством общения как с другими исследователями, так и с врачами [1]. Только во втором случае статья имеет непосред-

ственное отношение к медицинской практике. В первой части этой книги излагаются методы, облегчающие поиск наиболее достоверных данных по интересующей тематике. Овладев этими методами, читатель будет удивлен не только небольшим количеством исследований, с материалами которых ему действительно следует ознакомиться, но и простотой их поиска.

Другой способ, позволяющий справиться с избытком информации и сократить время, затрачиваемое на просмотр медицинской литературы, заключается в поиске данных для решения конкретной проблемы. В этом случае сначала должен быть сформулирован клинический вопрос, возникший при лечении конкретного больного, после чего осуществляется поиск тех данных, которые помогут решить этот вопрос. Вне зависимости от применяемого способа (просмотр медицинской периодики или поиск данных для решения конкретной проблемы) эта книга поможет читателю судить о достоверности публикуемых материалов, лучше понимать результаты исследований и успешно применять их в своей работе.

Далее рассматриваются навыки, которые необходимы, чтобы поиск данных для решения конкретной задачи был наиболее эффективным. I

ФОРМУЛИРОВАНИЕ КЛИНИЧЕСКОГО ВОПРОСА

Нередко врачу приходится решать сложные клинические вопросы, найти ответ на которые в медицинской литературе кажется невозможным. К основным навыкам, необходимым для работы в соответствии с принципами ДМ, относится умение подразделить любой сложный вопрос на основные компоненты, что существенно облегчает поиск данных [2, 3]. В большинстве клинических вопросов должны быть отражены три составляющие:

1. *Больные/Популяция.* Каковы основные характеристики интересующих вас больных?
2. *Вмешательство/Воздействие* (особенности питания, применяемые диагностические тесты, лекарственные препараты, хирургические вмешательства и т.д.). Какие медицинские вмешательства или потенциально вредные воздействия вас интересуют? Если речь идет об оценке методов лечения или факторов риска, в клиническом вопросе следует указать не только изучаемое лечение или воздействие, но и особенности применяемого контроля (вмешательство или воздействие отсутствует, имитируется или заменено другим).

3. *Клинический исход.* Каковы интересующие вас последствия (чаще всего неблагоприятные), имеющие значение для больного и связанные с изучаемым вмешательством или воздействием?

Рассмотрим несколько примеров, позволяющих проследить превращение первичного, неструктурированного клинического вопроса в четко сформулированный, структурированный, что облегчает дальнейшее использование медицинской литературы.

ПРИМЕР 1. ОПТИМАЛЬНЫЙ УРОВЕНЬ АД, КОТОРЫЙ СЛЕДУЕТ ПОДДЕРЖИВАТЬ У БОЛЬНЫХ С АРТЕРИАЛЬНОЙ ГИПЕРТОНИЕЙ НА ФОНЕ САХАРНОГО ДИАБЕТА

У 55-летней представительницы европеоидной расы диагностированы сахарный диабет II типа и артериальная гипертония. Прием метформина обеспечивает поддержание нормального уровня глюкозы в крови; осложнения сахарного диабета отсутствуют. Для снижения уровня АД больная ежедневно принимает тиазидный диуретик в низкой дозе. На протяжении 6 мес АД составляет в среднем 155/88 мм рт. ст.

Первичный, неструктурированный клинический вопрос: до какого уровня необходимо снижать АД при лечении артериальной гипертонии?

Более глубокий анализ проблемы. Один из недостатков такой формулировки вопроса — отсутствие достаточно полного перечня характеристик интересующей вас популяции, соответствующих характеристикам данной больной. Следует помнить, что эффективность строгого контроля за уровнем АД может в значительной степени зависеть от наличия сахарного диабета, его типа и развития осложнений. Можно также указать, интересует ли нас эффективность применения какого-то определенного или любого гипотензивного препарата. Кроме того, важно указать желаемый уровень АД, который будет поддерживаться в ходе проведения гипотензивной терапии. Например, врача могут интересовать возможные различия в эффективности терапии, направленной на поддержание диастолического АД <80 или <90 мм рт. ст. Однако основной недостаток первичной формулировки вопроса заключается в отсутствии критериев оценки эффективности гипотензивной терапии. В данной клинической ситуации врача должны интересовать такие клинические исходы, как развитие инсульта или инфаркта миокарда, смерть от сердечно-сосудистых заболеваний или любых причин.

Усовершенствованный, структурированный клинический вопрос, пригодный для проведения поиска данных. В окончательном варианте вопроса должны быть указаны особенности интересующей врача популяции больных, тактика лечения с использованием определенных вмешательств и важные для больных последствия этих вмешательств. Структурированный вариант клинического вопроса может выглядеть следующим образом:

Больные: больные с артериальной гипертензией и неосложненным сахарным диабетом II типа.

Вмешательство: применение любого гипотензивного препарата для поддержания диастолического АД на уровне 80 мм рт. ст. по сравнению с его поддержанием на уровне 90 мм рт. ст.

Клинические исходы: развитие инсульта, развитие инфаркта миокарда, смерть от сердечно-сосудистых заболеваний, смерть от любой причины.

ПРИМЕР 2. ТАКТИКА ВЕДЕНИЯ БОЛЬНЫХ ПРИ ПРЕДПОЛАГАЕМОЙ НЕСТАБИЛЬНОЙ СТЕНОКАРДИИ

Мужчина в возрасте 39 лет, никогда ранее не предъявлявший жалоб на боли в грудной клетке, обратился вечером в отделение неотложной помощи. Находясь на работе, он отметил ухудшение самочувствия и тошноту; появились неприятные ощущения в грудной клетке и обильное потоотделение. Симптомы сохранялись в течение 2 ч, после чего исчезли; осталось лишь чувство усталости. В конце рабочего дня мужчина, обеспокоенный произошедшим, решил обратиться за медицинской помощью. Указания на ишемическую болезнь сердца у ближайших родственников отсутствуют. На протяжении 5 лет больной принимает тиазидный диуретик по поводу артериальной гипертензии; эффект такой терапии вполне удовлетворительный. Стаж курения 15 пачко-лет*; уровень липидов в крови не повышен. Результаты осмотра, электрокардиографии, определения уровня креатинкиназы и тропонина I в крови указывают на отсутствие отклонений от нормы.

Первичный, неструктурированный клинический вопрос: можно ли отпустить больного домой или его следует госпитализировать для дальнейшего наблюдения в условиях стационара?

* Данный показатель рассчитывается путем умножения числа пачек сигарет, выкуренных за сутки, на продолжительность курения в годах (*Примеч. ред*)

Более глубокий анализ проблемы. По первоначальной формулировке вопроса трудно даже предположить, где искать на него ответ. Можно конкретизировать вопрос, указав на предполагаемую нестабильную стенокардию, хотя целый ряд характеристик отличает этого больного от других больных с подобным диагнозом. В данном случае мужчина сравнительно молод; и хотя у него имеются некоторые факторы риска развития ишемической болезни сердца, отмечавшиеся днем клинические проявления были нетипичными, а при обращении в отделение неотложной помощи боль в грудной клетке или признаки сердечной недостаточности отсутствуют, результаты электрокардиографии и определения уровня сердечных изоферментов также указывают на отсутствие отклонений.

Мужчину можно либо оставить до утра в отделении, оснащенном соответствующей аппаратурой для мониторинга, либо отпустить домой, порекомендовав ему в дальнейшем пройти более полное обследование, включая пробы с физической нагрузкой. Однако при выборе оптимального варианта необходимо учитывать последствия, к которым может привести выписка из стационара. Насколько она безопасна? Насколько низка вероятность развития осложнений? При ответе на этот вопрос в первую очередь следует учитывать фактор времени. Влияние данного фактора обычно оценивается в ходе исследований, посвященных изучению прогноза при том или ином заболевании.

Какую цель преследует госпитализация такого больного в кардиологическое отделение? Она не предотвратит развития отдаленных осложнений (например, инфаркта миокарда через месяц). Оставляя больного в стационаре, врач в первую очередь думает об осложнениях, которые могут развиваться в течение последующих 72 ч; в отсутствие таких осложнений дальнейший мониторинг обычно считают нецелесообразным. Если за время пребывания в стационаре у больного возникнет интенсивная боль в грудной клетке, сердечная недостаточность или инфаркт миокарда, ему будет оказана своевременная медицинская помощь, а при развитии фибрилляции желудочков или другой опасной для жизни аритмии персонал сможет немедленно произвести спасительную в таких ситуациях дефибрилляцию.

Усовершенствованный, структурированный клинический вопрос, пригодный для проведения поиска данных. В окончательном варианте вопроса Должны быть указаны особенности интересующей врача популяции больных, возможная тактика их ведения (применяемые вмешательства или потенциально вредные воздействия), а также важные для больных

последствия такого ведения. В данном случае структурированный вариант клинического вопроса может выглядеть следующим образом:

- *Больные:* молодые мужчины с предполагаемой нестабильной стенокардией, не имеющие типичной симптоматики и с нормальными результатами электрокардиографии и определения уровня сердечных изоферментов в крови.
- *Вмешательство:* госпитализация в отделение, оснащенное аппаратурой для мониторинга, по сравнению с выпиской из стационара, т.е. наблюдение различной продолжительности.
- *Клинические исходы:* развитие тяжелой стенокардии, инфаркта миокарда, сердечной недостаточности или аритмии в течение 72 ч после обращения к врачу.

ПРИМЕР 3. ТАКТИКА ВЕДЕНИЯ БОЛЬНЫХ ПРИ ПЛОСКОКЛЕТОЧНОМ РАКЕ ЛЕГКИХ

Мужчина 60 лет со стажем курения 40 пачко-лет поступил с жалобами на кровохарканье. Других жалоб не предъявляет; при общем осмотре отклонений от нормы не найдено. При рентгенографии органов грудной клетки выявлено затемнение в легких без изменений структур средостения, результаты пункционной биопсии указывают на наличие плоскоклеточного рака легких.

Первичный, неструктурированный клинический вопрос: какие дополнительные обследования следует провести для решения вопроса о необходимости хирургического лечения?

Более глубокий анализ проблемы. К основным характеристикам больного относятся наличие у него немелкоклеточного рака легких в отсутствие регионарных или отдаленных метастазов (по данным анамнеза, общего осмотра и рентгенографии органов грудной клетки). Выбор дополнительных методов диагностики определяется вероятностью наличия у больного недиагностированных метастазов в средостение или другие органы. Для выявления скрытого метастатического поражения структур средостения можно выполнить либо первичную медиастиноскопию, либо компьютерную томографию грудной клетки, на основании результатов которой будет определяться дальнейшая тактика ведения больного.

На какие клинические исходы можно повлиять, выбирая тот или иной диагностический подход? Основная цель состоит в продлении жизни

больного, однако в данном случае главным прогностическим фактором служит распространенность злокачественного процесса, и ни один из диагностических методов не способен на нее повлиять. Выявить скрытые метастазы в средостение, если таковые имеются, необходимо потому, что при их наличии хирургическое удаление опухоли, вероятнее всего, окажется неэффективным. В таком случае будет целесообразным рекомендовать больному паллиативное лечение и не подвергать его риску, связанному с бесполезной торакотомией. Таким образом, в этом примере основным клиническим исходом будет вероятность выполнения необоснованной операции.

Усовершенствованный, структурированный клинический вопрос, пригодный для проведения поиска данных. В окончательном варианте вопроса должны быть указаны особенности интересующей врача популяции больных, возможная тактика их ведения (применяемые диагностические вмешательства) и важные для больных последствия такого ведения. В данном случае структурированный вариант клинического вопроса может выглядеть следующим образом:

Больные: больные с впервые выявленным немелкоклеточным раком легкого без отдаленных метастазов.

Вмешательство: первичная медиастиноскопия или компьютерная томография, по результатам которой определяется дальнейшая тактика ведения больного.

Клинический исход: выполнение необоснованной торакотомии.

Данный клинический вопрос можно структурировать иным образом, позволяющим уточнить информативность компьютерной томографии. При этом представленные характеристики больных останутся прежними, изучаемым вмешательством будет компьютерная томография, а клиническим исходом — наличие или отсутствие метастазов в средостение. Далее (см. главу 8 *Диагностические методы*) будет обсуждаться вопрос о том, почему при подобной формулировке клинического вопроса вероятность выбора оптимальной тактики ведения больного может быть ниже.

Из вышеприведенных примеров видно, что правильно сформулировать клинический вопрос, позволяющий провести эффективный поиск соответствующих данных в медицинской литературе, не всегда просто. Это требует от врача углубленного понимания самых разных аспектов, так или иначе связанных с тактикой лечения. Работа с конкретным больным неизбежно влечет за собой целый ряд вопросов, и врач должен четко представлять себе, что он хочет узнать. Чтобы правильно поставить

вопрос, необходимо всегда помнить схему его структурирования ("Большой, вмешательство, клинический исход").

После того как клинический вопрос надлежащим образом сформулирован, следует определить оптимальную стратегию поиска данных. Зная все 3 составляющие вопроса, сделать это гораздо проще.

ПОИСК ОТВЕТА

В этом разделе речь пойдет об электронных базах данных, с помощью которых можно достаточно быстро найти ответы на интересующие вопросы. Будет наглядно продемонстрировано, насколько правильная постановка вопроса (т.е. четкое определение характеристик больных, вмешательств и клинических исходов) облегчает разработку оптимальной стратегии поиска. Однако следует учитывать и четвертый, очень важный компонент клинического вопроса, а именно тип исследований, материалы которых станут объектом поиска. Под типом здесь подразумевается структура исследования, т.е. особенности его планирования и проведения.

ОПРЕДЕЛЕНИЕ ТИПА ВОПРОСА

Чтобы получить полное представление о структуре исследования, необходимо целиком ознакомиться с первой частью этой книги. Ниже-изложенное следует рассматривать лишь в качестве краткого введения.

Все клинические вопросы можно подразделить на 4 основных типа, относящихся к лечению, диагностике, оценке риска и прогноза.

- **Лечение.** Клинический вопрос посвящен оценке эффективности различных лечебных вмешательств; при этом интересующими клиническими исходами будет улучшение состояния больных или снижение риска развития осложнений.
- **Оценка риска.** Клинический вопрос посвящен оценке возможного неблагоприятного влияния различных воздействий (или применяемых лечебных вмешательств) на состояние больных, заболеваемость и смертность.
- **Диагностика.** Клинический вопрос посвящен оценке информативности того или иного метода диагностики, т.е. точности его положительного или отрицательного результатов, позволяющей установить наличие или отсутствие искомого заболевания.

- **Прогноз.** Клинический вопрос посвящен оценке дальнейшего течения заболевания.

Чтобы найти ответ на клинический вопрос об эффективности определенного лечебного вмешательства, проводят поиск рандомизированных контролируемых испытаний (РКИ), в которых распределение участников на группы осуществляется случайным образом, по принципу, схожему с подбрасыванием монеты (см. главу 4 *Исследования, посвященные эффективности лечения*). В ходе таких РКИ после формирования основной и контрольной групп за больными ведется постоянное наблюдение, направленное на выявление заранее определенных клинических исходов, например развития инсульта или инфаркта миокарда (рис. 2).

Теоретически наиболее достоверные данные о факторах риска и безопасности медицинских вмешательств также следует искать в отчетах об РКИ. Однако при изучении многих вредных воздействий и потенциально опасных вмешательств рандомизация невозможна или недопустима по соображениям этики. Едва ли кто-нибудь согласится на то, чтобы посторонние люди подбрасыванием монеты решали, суждено ли ему или ей на протяжении последующих 20 лет курить либо контактировать с источниками ионизирующего излучения. В таких случаях целесообразно искать исследования, в которых участники подвергаются потенциально вредным воздействиям добровольно или по стечению не зависящих от них обстоятельств. Данные, полученные в ходе таких обсервационных исследований, менее достоверны, чем результаты РКИ.

На рис. 3 представлена общая схема обсервационного исследования, в ходе которого участники, подвергающиеся или не подвергающиеся влиянию определенного фактора риска, наблюдаются на протяжении определенного периода времени с целью выявления изучаемых клини-

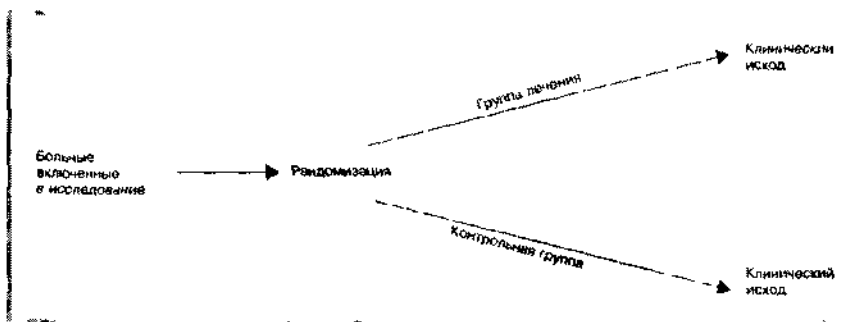


Рис. 2. Схема рандомизированного контролируемого испытания.

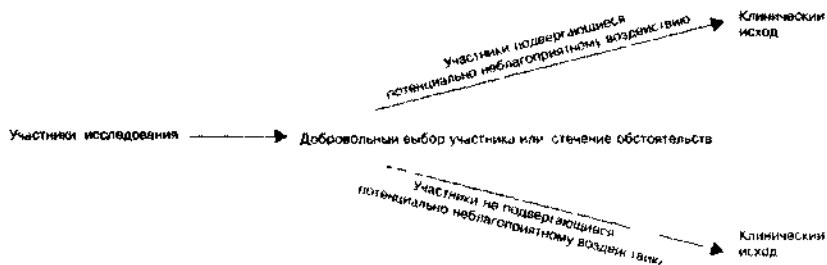


Рис. 3. Схема обсервационного исследования, посвященного оценке влияния потенциально неблагоприятного воздействия

ческих исходов. При оценке последствий курения и воздействия ионизирующего излучения таким клиническим исходом может быть развитие рака.

Для оценки информативности (свойств, рабочих характеристик) того или иного метода диагностики структура исследования должна быть иной. На первом этапе отбираются больные, у которых предполагается наличие искомого заболевания или состояния (например, туберкулеза, рака легкого или железодефицитной анемии). При обследовании больных сначала используется изучаемый диагностический метод, а затем — "золотой стандарт" (общепринятый метод диагностики определенного заболевания или состояния, называемый иначе диагностическим стандартом, или референсным методом). После этого результаты применения этих 2 методов сравниваются (рис. 4).

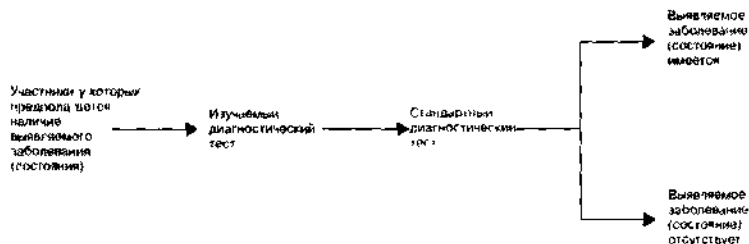


Рис. 4. Схема исследования, посвященного оценке информативности диагностического метода.

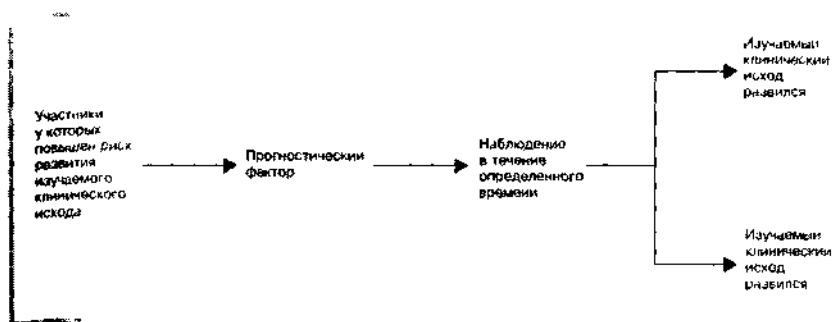


Рис. 5. Схема обсервационного исследования, посвященного изучению прогноза.

В ходе исследований последнего, четвертого типа изучается прогноз при тех или иных заболеваниях, а также выявляются факторы, которые могут повлиять на него. На первом этапе отбираются участники, относящиеся к определенной группе (например, беременные, больные, направляемые на операцию, или больные со злокачественными опухолями), а затем у них определяется наличие или отсутствие определенных факторов, которые могут повлиять на прогноз (например, пожилой возраст или наличие сопутствующих заболеваний). Однако основным определяющим фактором в таких исследованиях считается время, в течение которого ведется наблюдение за участниками с целью выявления изучаемого клинического исхода (например, нарушений родовой деятельности, развития инфаркта миокарда в послеоперационном периоде или смерти при злокачественных опухолях). Схема такого исследования представлена на рис. 5.

Одна из основных задач врача, приступающего к изучению медицинской литературы, заключается в правильном определении типа исследования, материалы которого могут содержать ответ на поставленный вопрос. Например, если врач, желая узнать о характеристиках того или иного диагностического метода (но не о влиянии его использования на эффективность лечения), станет искать материалы РКИ, его усилия едва ли увенчаются успехом.

Вернемся к вопросам, о которых шла речь в предыдущем разделе данной главы. Безусловно, при выработке оптимальной тактики ведения больных с артериальной гипертонией, следует искать исследования, посвященные оценке эффективности лечения. Если же врача интересует и риск развития редких или отдаленных побочных эффектов гипо-

тензивной терапии, необходимо также искать исследования, посвященные оценке безопасности лечения.

При рассмотрении второго примера клинический вопрос можно сформулировать по-разному. Если врач хочет узнать, насколько высок риск развития инфаркта миокарда или смерти у молодых мужчин с предполагаемой нестабильной стенокардией в отсутствие типичных симптомов, то необходимую информацию он может найти в материалах исследований, посвященных изучению прогноза. Если врача в первую очередь интересует сравнительная эффективность двух альтернативных подходов (отпустить больного домой или госпитализировать его для дальнейшего наблюдения), нужную информацию следует искать в отчетах об РКИ.

В третьем примере также возможны два варианта постановки вопроса. Если надо уточнить, какова вероятность того, что проведение компьютерной томографии позволит выявить у больных с немелкоклеточным раком легкого наличие метастазов в средостение, следует искать исследования, посвященные оценке информативности диагностических методов (см. рис. 4). Если же врача интересует, какова сравнительная частота проведения необоснованной торакотомии у больных с немелкоклеточным раком легких, которым выполнялась первичная медиастиноскопия по сравнению с компьютерной томографией, то предметом поиска будут РКИ (см. рис. 2).

ЦЕЛЕСООБРАЗЕН ЛИ ПОИСК ДАННЫХ В МЕДИЦИНСКОЙ ЛИТЕРАТУРЕ?

Время, которым мы располагаем для поиска, ограничено, поэтому желательно быть уверенным в том, что оно не будет потрачено напрасно. С этой точки зрения стоит рассмотреть следующие клинические вопросы.

Пример 1. Насколько наличие инфаркта легкого ухудшает прогноз при тромбоэмболии легочной артерии?

Прежде чем определять стратегию поиска данных и приступать к изучению медицинской литературы, следует задуматься, каким образом исследователи могут выявить инфаркт легкого. Поскольку в настоящее время не существует надежных способов его диагностики (за исключением аутопсии), нецелесообразность поиска данных становится очевидной.

Пример 2. Больная 50 лет, 4 дня назад перенесшая неосложненный инфаркт миокарда, перед выпиской из стационара интересуется, когда она сможет возобновить половую жизнь.

В данном случае клинический вопрос можно сформулировать примерно так:

- *Больные*: женщины, перенесшие неосложненный инфаркт миокарда.
- *Вмешательство*: разрешение возобновить половую жизнь в любое время по сравнению с советом воздержаться от этого, например, в течение 8 нед.
- *Клинические исходы*: развитие повторного инфаркта миокарда, развитие нестабильной стенокардии, смерть от сердечно-сосудистых заболеваний, смерть от любой причины, изменение связанного со здоровьем качества жизни.
- *Тип вопроса*: оценка эффективности вмешательства; следовательно, необходимо искать материалы РКИ.

Однако вероятность проведения РКИ по данному вопросу чрезвычайно низка. Чуть выше может быть вероятность проведения наблюдательных исследований, в ходе которых женщины, перенесшие инфаркт миокарда, сообщали бы врачам о времени возобновления половых отношений, а врачи сравнивали бы клинические исходы у тех, кто сделал это в ранние или более поздние сроки.

Из приведенных примеров видно, что попытки поиска данных в медицинской литературе не всегда оправданы. Такой поиск будет безрезультатным, если ни один из рассмотренных выше типов исследования не позволяет получить ответ на поставленный вопрос. Но даже при наличии теоретической возможности проведения соответствующих исследований значимость проблемы может быть настолько низкой, что ни один из исследователей не станет тратить время, пытаясь ее разрешить. Прежде чем начать поиск, нужно быть уверенным, что результат будет стоить потраченного на него времени.

ИСТОЧНИКИ ДОКАЗАТЕЛЬНОЙ ИНФОРМАЦИИ

Можно обратиться к врачам своей или иной специальности не только для того, чтобы узнать их мнение о каком-либо аспекте медицины, но и для того, чтобы получить информацию, необходимую для ответа на поставленный клинический вопрос (см. главу 1 *Введение: философская концепция доказательной медицины*). Это особенно важно в тех случаях, когда вероятность успешного поиска в медицинской литературе невелика. Специалисты, которые знакомы с последними исследованиями в своей области, могут достаточно быстро снабдить вас актуальными ссылками.

Конечно же, врачей не нужно убеждать в полезности советов более опытных коллег; этим способом получения информации пользуются достаточно активно. Проблемы, как правило, возникают при необходимости использовать электронные источники данных. В этой книге акцент сделан именно на них, а не на печатных изданиях, поскольку постоянно обновляемые компьютерные базы данных позволяют осуществить оперативный поиск необходимой доказательной информации (табл. 1) [4]. Многие из этих баз появились недавно, поэтому исследований, посвященных сравнительной оценке их достоинств и недостатков, пока мало. Принципы работы с базами данных, описываемые в этой книге, основаны на нашем личном опыте, а также на опыте наших коллег, в том числе преподавателей основ доказательной медицины.

Таблица 1 Постоянно обновляемые медицинские электронные базы данных

База данных	Адрес в Интернете	Стоимость годовой подписки, доллары США*
ACP Journal Club	www.acponline.org/journals/acpj/jcmenu.htm	65
Best Evidence	www.acponline.org/catalog/electronic/best_evidence.htm	85
Cochrane Library	www.update-software.com/cochrane/cochrane-frame.html	225
UpToDate	www.uptodate.com	495
MEDLINE		
PubMed	www.ncbi.nlm.nih.gov/PubMed	Бесплатная
Internet Grateful Med	lgn.nlm.nih.gov	»
Другие базы данных	www.medmatrx.org/info/medhnetable.asp	»
Scientific American Medicine	www.samed.com	245 (159 при работе только в режиме он-лайн)
Clinical Evidence	www.evidence.org	115
Harrison's Online	www.harnsonsonline.com	89
eMedicine	www.emedicine.com	Бесплатная
Medscape	www.medscape.com/Home/Topics/homepages.html	»
Medical Matrix	www.medmatrx.org/index.asp	»

Таблица 1 (продолжение)

База данных	Адрес в Интернете	Стоимость годовой подписки, доллары США*
ScHARR Netting the Evidence	www.shef.ac.uk/~scharr/ir/netting	Бесплатная
Medical World Search	www.mwsearch.com	»
Journal Listings	www.nthames-health.tpmde.ac.uk/connect/journals.htm www.pslgroup.com/dg/medjournals/htm	»
Clinical practice guidelines	www.guidelines.gov www.csta.ca/cpgs	»
MD Consult	www.mdconsult.com	200
Evidence-based Medicine Reviews (OVID)	www.ovid.com/products/chmcal/ebmr.cfm (доступна во многих медицинских библиотеках)	1995

Примечание * — по данным на 2000 г

ВЫБОР ОПТИМАЛЬНОГО ИСТОЧНИКА ДАННЫХ

Какую электронную базу данных можно считать наилучшей⁹ Это определяется главным образом типом интересующего вас вопроса и временем, которое вы можете затратить на поиск информации [5] Результаты наблюдательных исследований конца 80-х годов XX века указывают на то, что в амбулаторных условиях у врача возникает 1–2 вопроса по каждому больному [6], а в условиях стационара — от 1 до 5 [7] В ходе более поздних исследований, выполненных в Великобритании [8] и США [9], было показано, что в практике семейного врача этот показатель в среднем составляет 0,32

Следует убедиться, что в той базе данных, которую вы выбрали, можно найти самый полный ответ на интересующий вопрос Так, обращение к MEDLINE представляет собой не лучший способ поиска сведений о нормальной анатомии человека, а больничная информационная система служит источником исчерпывающей информации о лабораторных показателях у конкретного больного В табл. 2 приведены примеры клинических вопросов 4 основных типов, соответствующая им структура исследований, оптимальная база данных для поиска ответа и ключевые термины, принятые в базе данных MEDLINE для определения типа исследования.

Таблица 2. Организация поиска ответа на правильно сформулированный клинический вопрос

Тип вопроса	Больные/ популяция	Вмешательство/ воздействие	Клинический исход	Наиболее пригодная структура исследования	Оптималь- ные базы данных	Ключевой термин для определения типа исследования, принятый в базе данных MEDLINE
Диагностика	У больных раком легкого	каковы характеристики компьютерной томографии	при выявлении метастазов в средостение?	Перекрестное аналитическое исследование	Best Evidence, UpToDate, MEDLINE	Sensitivity.tw
Оценка риска	У мужчин	приводит ли вазэктомия	к развитию рака яичек?	Когортное исследование, популяционное исследование случай—контроль	Best Evidence, UpToDate, MEDLINE	Risk.tw
Прогноз	У молодых мужчин с нетипичной болью в грудной клетке	в течение 72 ч после выписки домой из отделения неотложной помощи	увеличивается ли риск развития нестабильной стенокардии, сердечной недостаточности, аритмии, инфаркта миокарда или внезапной смеси?	Когортное исследование	Best Evidence, UpToDate, MEDLINE	Explode cohort studies

Таблица 2 (продолжение)

Тип вопроса	Больные/ Популяция	Вмешательство/ Воздействие	Клинический исход	Наиболее пригодная структура исследования	Оптималь- ные базы данных	Ключевой термин для определения типа исследования, принятый в базе данных MEDLINE
Лечение	У больных с артериальной гипертонией в сочетании с сахарным диабетом II типа	гипотензивная терапия, направленная на поддержание диастолического АД на уровне 80 мм рт.ст. по сравнению с 90 мм рт.ст..	снижает ли риск развития инсульта, инфаркта миокарда, смерти от сердечно-сосудистых заболеваний и смерти от любой причины?	РКИ или систематические обзоры РКИ	Cochrane Library, Best Evidence, UpToDate, MEDLINE	Meta-analysis.pt (для систематических обзоров) или Clinical trial.pt (для РКИ)

При поиске ответа на частный, правильно сформулированный клинический вопрос лучше всего начинать с тех баз данных, в которые включаются только материалы, отвечающие определенным критериям методологического качества, например Best Evidence, Cochrane Library или Clinical Evidence (см. табл. 1). Использование специальных критериев ("фильтров") обеспечивает включение в базу данных исследований высочайшего методологического качества. Эти базы данных организованы таким образом, чтобы максимально облегчить поиск необходимых сведений. Информация регулярно обновляется (с интервалом от нескольких месяцев до 2 лет) за счет включения результатов высококачественных и клинически значимых исследований.

Учебные базы данных. При поиске ответов на вопросы общего характера вряд ли пригодны электронные источники информации, в которых используются методологические "фильтры". Целесообразнее обратиться к постоянно обновляемым учебным базам данных, включающим в себя большой объем информации. Примером таких информационных ресурсов служат базы данных UpToDate и *Scientific American Medicine*, пополняющиеся каждые несколько месяцев или лет по мере появления новых важных научных фактов. Статьи в них содержат подробные библиографические списки, что позволяет определить время написания того или иного раздела учебника; более того, читатель может ознакомиться с текстом оригинальных публикаций. Среди других учебных баз данных, доступных в электронном виде, стоит упомянуть *Harrison's Principles of Internal Medicine*; с ее помощью можно найти ответ на многие общие вопросы. Появились также медицинские учебники, существующие только в Интернете (например, *emedicine*). По мере увеличения удельного объема доказательной информации, обновляющейся почти сразу после появления новых данных, электронные информационные ресурсы становятся все более важным источником для поиска ответов как на частные, так и на общие вопросы. Наш собственный опыт позволяет утверждать, что UpToDate и Clinical Evidence уже сегодня служат достаточно надежным способом решения частных клинических проблем.

MEDLINE. Эта библиографическая электронная база данных организована и поддерживается Национальной медицинской библиотекой США. Обращение к ней позволяет наиболее эффективно решать четко сформулированные частные клинические вопросы. Однако огромный объем содержащейся в ней информации и довольно сложная структура затрудняют поиск нужных данных, требуя существенных затрат време-

ни. Поэтому мы рекомендуем обращаться к **MEDLINE** только в случае безуспешного поиска в базах данных, где используются методологические "фильтры" (или когда об отсутствии в них необходимых сведений можно предположить заранее).

Рассмотрим алгоритм поиска информации в электронных базах данных на конкретном клиническом примере о выборе оптимального уровня ДД, который следует поддерживать у больных с артериальной гипертензией на фоне сахарного диабета.

МЕДИЦИНСКИЕ ЭЛЕКТРОННЫЕ БАЗЫ ДАННЫХ, В КОТОРЫЕ ВКЛЮЧАЮТСЯ ТОЛЬКО МАТЕРИАЛЫ, ОТВЕЧАЮЩИЕ ОПРЕДЕЛЕННЫМ КРИТЕРИЯМ МЕТОДОЛОГИЧЕСКОГО КАЧЕСТВА

Сначала следует попытаться найти систематические обзоры по данной теме. Систематические обзоры содержат информацию, необходимую для ответа на частный клинический вопрос, и проводятся таким образом, чтобы уменьшить вероятность возникновения систематической ошибки. Авторы высококачественных систематических обзоров отбирают и обобщают только самые достоверные данные (как появившиеся на страницах печатных или электронных изданий, так и неопубликованные), полученные в исследованиях высокого методологического качества. Обращение к Best Evidence и Cochrane Library позволяет осуществить быстрый и эффективный поиск таких систематических обзоров. Кроме того, обе базы данных дают возможность ознакомиться с текстом оригинальных статей.

Best Evidence

В этой базе данных можно быстро найти систематические обзоры и оригинальные статьи, необходимые для решения частного клинического вопроса. Она доступна как на компакт-дисках, так в Интернете, и представляет собой объединенную электронную версию двух печатных журналов *ACP Journal Club* и *Evidence-Based Medicine*, содержащих предварительно обработанную информацию. В январе 2000 г. в Северной, Южной и Центральной Америке эти журналы были объединены в одно издание *ACP Journal Club*; стехпор *Evidence-Based Medicine* выходит только за пределами США. Редакционная коллегия этих изданий осуществляет регулярный систематический поиск исследований, характеризующихся высоким методологическим качеством и клинической значимостью (особенно тех из них, которые посвящены диагностике и лечению самых распространенных заболеваний и состояний), в 170 меди-

цинских журналах. При оценке методологического качества исследования используются специальные критерии (см. главы 4 *Исследования, посвященные эффективности лечения*, 5 *Оценкариска*, 6 *Диагностика*, 8 *Диагностические методы* и 9 *Прогноз*). Так, в раздел, посвященный лечению, включаются только те РКИ, в которых до конца исследования наблюдались не менее 80% участников, а в раздел, посвященный диагностике, — только те исследования, в которых проводилось сравнение информативности изучаемого диагностического теста и "золотого стандарта" с обязательным привлечением независимых экспертов и с использованием слепого метода.

ACP Journal Club и *Evidence-Based Medicine* содержат структурированные рефераты высококачественных исследований и комментарии специалистов с обсуждением перспектив практического использования полученных результатов. В разделе базы данных Best Evidence, озаглавленном Other articles noted ("Другие статьи"), врач может найти данные, полученные в исследованиях достаточно высокого методологического качества, но признанные менее достоверными. Best Evidence обновляется ежегодно и в настоящее время включает в себя более 2000 рефератов статей по терапии, опубликованных с 1991 г., а также материалы, публикуемые с 1995 г., по таким дисциплинам, как акушерство и гинекология, семейная медицина, педиатрия, психиатрия и хирургия. Каждые 5 лет все материалы оцениваются заново с точки зрения их соответствия современному уровню знаний.

Поскольку в базу данных Best Evidence включаются только исследования высокого методологического качества, ее объем значительно меньше, чем у многих аналогичных источников информации; это существенно облегчает поиск. С другой стороны, в связи с небольшим количеством включенных материалов эта база не охватывает весь диапазон имеющихся данных; если врач ограничится только Best Evidence, поиск не будет полным, что повышает вероятность возникновения систематической ошибки, связанной с отбором исследований. Однако, на наш взгляд, существенные достоинства Best Evidence (высокое методологическое качество работ и быстрота поиска) компенсируют этот недостаток.

Пример использования базы данных Best Evidence. Открываем титульную страницу Best Evidence 4 (рис. 6). Для поиска данных о гипотензивной терапии при сахарном диабете II типа выбираем функцию "Search" ("Поиск") и вводим термины, отражающие основные компоненты клинического вопроса:

"hypertension AND diabetes AND mortality"
("артериальная гипертензия И сахарный диабет И смертность").

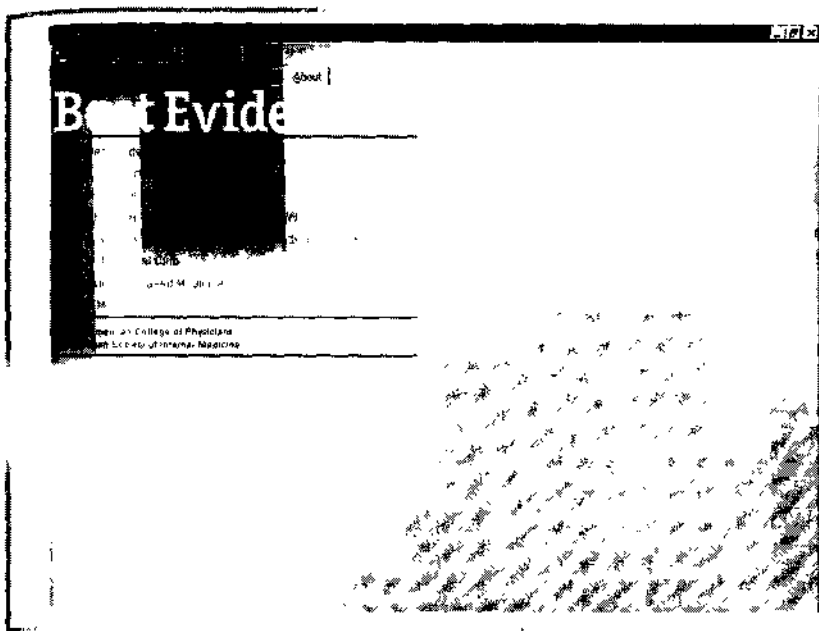


Рис. 6. Титульная страница базы данных Best Evidence, распространяемой на компакт-дисках.

Воспроизводится с разрешения American College of Physicians — Society of Internal Medicine.

На экране появляется список из 109 публикаций. Однако многие из них посвящены прогнозу при сахарном диабете и не имеют прямого отношения к интересующему нас вопросу. Поэтому мы возвращаемся к предыдущему окну и меняем условия поиска: вместо требования "All topics" ("Все статьи") выбираем "Selected topics" ("Отдельные статьи"), а затем щелкаем по кнопкам "Therapeutics" ("Лечение") и "Search" ("Поиск"). В результате повторного поиска получаем список из 27 публикаций; все они посвящены оценке эффективности лечебных вмешательств (рис. 7). Данный список включает в себя 5 систематических обзоров, но ни в одном из них не содержится интересующая нас информация. Первая же из 22 оригинальных статей озаглавлена "Tight blood pressure control reduced diabetes mellitus-related death and complications and was cost-effective in diabetes" ("Тщательная коррекция артериального давления снижает риск развития осложнений, связанных с сахарным диабетом, и характеризуется высокой эффективностью затрат"; рис. 8). Двойной щел-

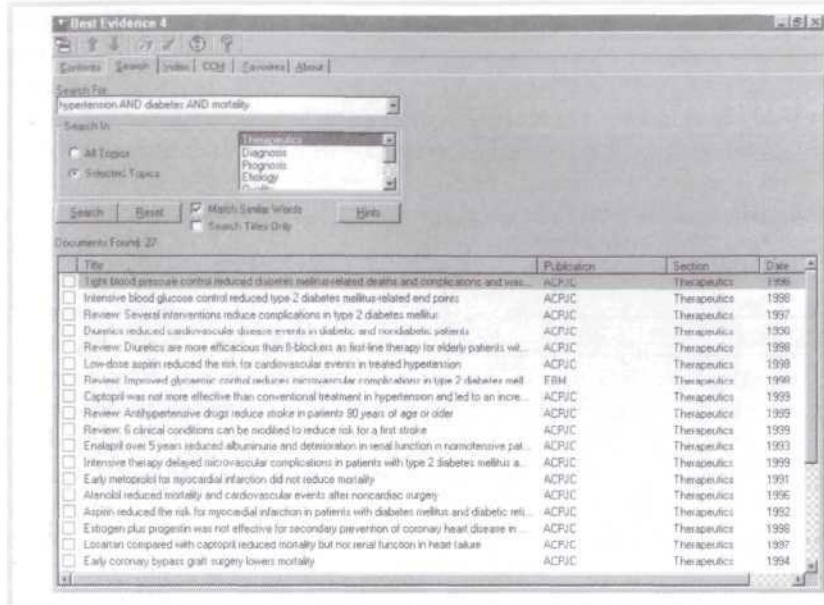


Рис. 7. Список публикаций, найденных в базе данных Best Evidence при повторном поиске.

Воспроизводится с разрешения American College of Physicians — Society of Internal Medicine.

чок по названию статьи открывает текст структурированного реферата с описанием РКИ, в ходе которого больным с сахарным диабетом II типа и артериальной гипертензией назначалась гипотензивная терапия, направленная на поддержание АД либо на уровне $< 150/85$ мм рт. ст., либо на уровне $< 180/105$ мм рт. ст. (рис. 9). Продолжительность наблюдения составила в среднем 9 лет. В ходе РКИ было показано (рис. 10), что более тщательная коррекция АД, т.е. проведение гипотензивной терапии, направленной на поддержание более низкого уровня АД, снижает риск смерти от осложнений, связанных с сахарным диабетом, на 32% (при 95% ДИ от 8 до 50%; $p=0,019$).

Не всегда поиск в базе данных Best Evidence оказывается успешным. Иногда это связано с отсутствием исследований высокого методологического качества по интересующему вопросу, в некоторых случаях — с тем, что их результаты не были опубликованы ни в одном из 170 журналов, материалы которых используются при составлении этой базы данных. Кроме того, нужные вам материалы могли появиться уже после выхода в свет последней обновленной версии Best Evidence или до 1991 г.

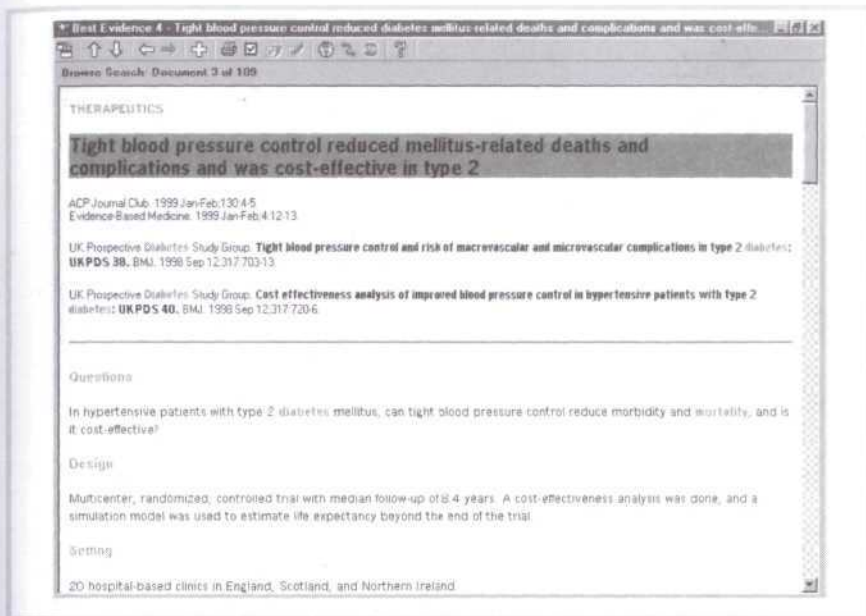


Рис. 8. Развернутый текст реферата из базы данных Best Evidence (начало).
Воспроизводится с разрешения American College of Physicians — Society of Internal Medicine.

Даже если исследование обладает достаточно высоким методологическим качеством, а его результаты были опубликованы после 1991 г., отчет о нем может отсутствовать в Best Evidence, если редакторы сочли, что его тематика не соответствует основной направленности издания (внутренние болезни). Несмотря на перечисленные ограничения, обычно удается найти ответ на возникший клинический вопрос, особенно когда он имеет отношение к какому-либо распространенному заболеванию или состоянию. Но даже если поиск безрезультатен, он отнимает так мало времени, что всегда остается возможность найти ответ в другом источнике информации.

Cochrane Library

Кокрановское сотрудничество (Cochrane Collaboration) представляет собой международную организацию, в задачи которой входят подготовка, проведение и распространение систематических обзоров, посвященных оценке эффективности и безопасности различных медицинских вмешательств, а также создание электронной базы данных, позволяю-

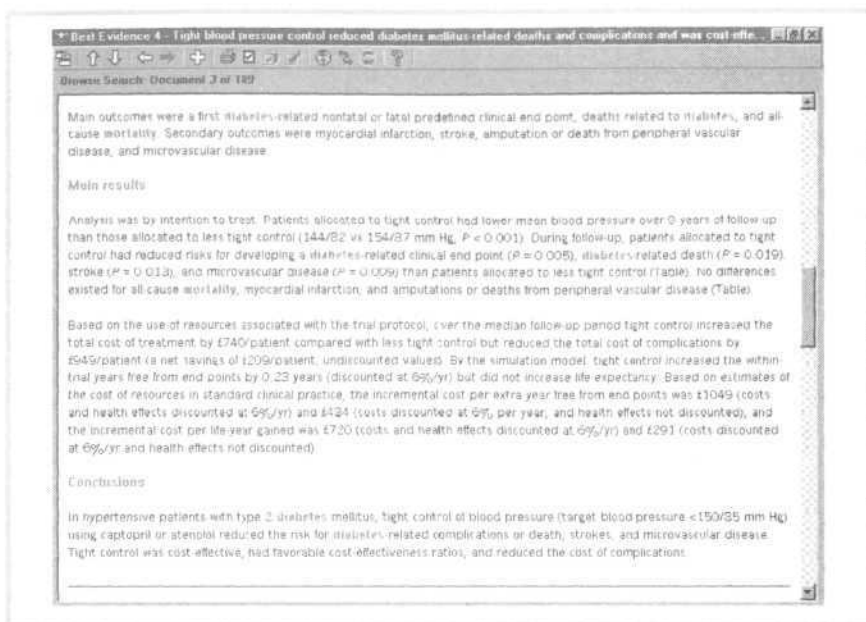


Рис. 9. Развернутый текст реферата из базы данных Best Evidence (продолжение).
Воспроизводится с разрешения American College of Physicians — Society of Internal Medicine.

щей осуществлять быстрый поиск необходимой информации. Эта база (Cochrane Library) включает в себя главным образом систематические обзоры РКИ, посвященные оценке различных аспектов лечения. Результаты других исследований (например, посвященных изучению новых методов диагностики или прогноза) в ней практически отсутствуют.

Cochrane Library публикуется на компакт-дисках и на сайте в Интернете, обновляется ежеквартально и состоит из 3 основных разделов. Первый из них, база данных Кокрановских систематических обзоров (Cochrane Database of Systematic Reviews — CDSR), содержит исчерпывающую информацию обо всех систематических обзорах, подготовленных членами Кокрановского сотрудничества (в первом выпуске за 2000 г. таких обзоров насчитывалось 716), а также о протоколах незавершенных обзоров. Второй раздел, база данных обзоров, посвященных оценке эффективности вмешательств (Database of Abstracts of Reviews of Effectiveness — DARE), включает в себя систематические обзоры, которые были выполнены не входящими в Кокрановское сотрудничество авторами (в первом выпуске за 2000 г. таких обзоров насчитывалось 2565).

Best Evidence 4 - Tight blood pressure control reduced diabetes mellitus-related deaths and complications and was cost-effective

Browser: Google, Document 2 of 100

Tight vs less tight blood pressure control in hypertensive patients with type 2 diabetes mellitus*

Outcomes over 8.4 y of follow-up	Tight control	Less tight control	RRR (95% CI)	NNT (CI)
Any diabetes-related clinical end point	34%	44%	22% (9 to 32)	11 (6 to 29)
Diabetes-related death	11%	16%	32% (8 to 50)	20 (10 to 100)
All-cause mortality	18%	21%	17% (-6 to 35)	Not significant
Myocardial infarction	14%	18%	20% (-5 to 39)	Not significant
Stroke	5%	9%	42% (10 to 63)	27 (14 to 141)
Peripheral vascular stroke	1%	2%	49% (-32 to 80)	Not significant
Microvascular disease	8%	14%	35% (9 to 54)	21 (11 to 34)

*Abbreviations defined in Glossary; RRR, NNT, and CI calculated from data in article.

Commentary

The UKPDS shows that meticulous blood pressure reduction is important in patients with type 2 diabetes. This finding has major implications for health care because of the projected dramatic increase in diabetes in the future and the clear relation between diabetes and hypertension. The Systolic Hypertension in the Elderly Program (SHEP) has shown that antihypertensive treatment reduces cardiovascular events in elderly patients with and without diabetes (1). The present trial consolidates this information and extends the findings to younger patients and to patients with newly detected diabetes. The results are consistent with the recently published Hypertension Optimal Treatment (HOT) study, in which intensive blood

Рис. 10. Развернутый текст реферата из базы данных Best Evidence с итоговой таблицей и комментарием (окончание).

Воспроизводится с разрешения American College of Physicians — Society of Internal Medicine.

Этот раздел представляет собой самостоятельный информационный ресурс и доступен не подписавшимся на Cochrane Library пользователям по адресу <http://nhscrd.york.ac.uk>; кроме того, он позволяет осуществлять поиск исследований, посвященных оценке экономических и технологических аспектов оказания медицинской помощи.

Третий раздел, реестр контролируемых испытаний (Cochrane Controlled Trials Registry — CCTR), содержит постоянно пополняющийся перечень клинических испытаний (в настоящее время более 268 000), выявленных членами Кокрановского сотрудничества в различных источниках, в том числе в библиографических базах данных MEDLINE и EMBASE/Excerpta Medica, посредством ручного поиска и изучения библиографических списков в оригинальных статьях и обзорах. Наряду с отчетами об РКИ, составляющими большую часть материалов данного раздела, в него также входят сообщения о некоторых обсервационных исследованиях. Возможно, скоро в данный раздел будут включаться и исследования, посвященные оценке информативности методов диагностики. Кроме того, в базе данных Cochrane Library содержится инфор-

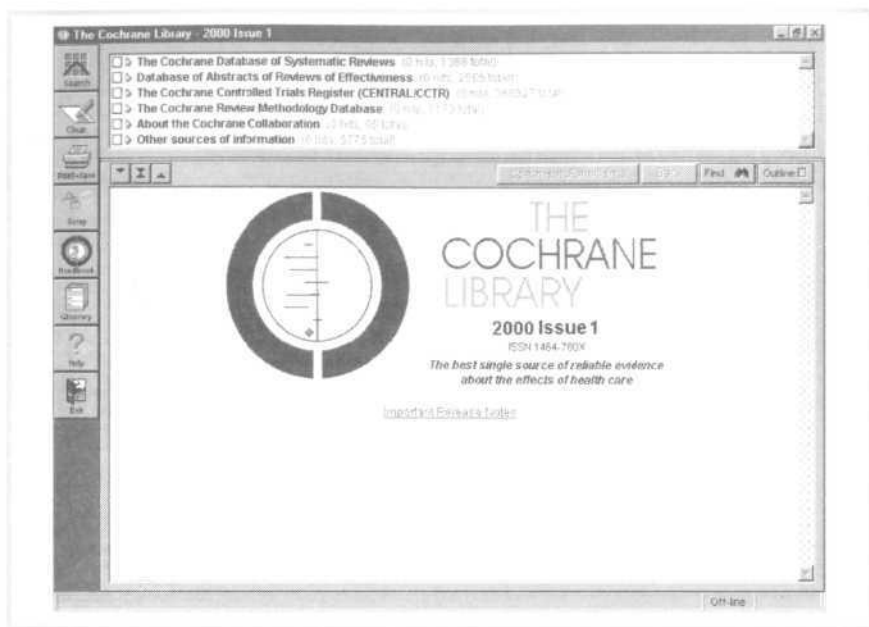


Рис. 11. Титульная страница базы данных Cochrane Library.

Воспроизводится с разрешения Update Software.

мация о Кокрановском сотрудничестве, о принципах и методологии проведения систематических обзоров.

Для проведения поиска данных в Cochrane Library (рис. 11) следует выбрать функцию "Search" ("Поиск"). При пользовании версией этой базы данных, издаваемой на компакт-диске, можно дважды щелкнуть левой кнопкой "мыши" по кнопке "Advanced search" ("Расширенный поиск"), что позволяет провести поиск по медицинским предметным рубрикам (Medical Subject Headings — MeSH) и использовать логические операторы запроса (подробнее эти возможности описаны ниже в разделе, посвященном работе с базой данных MEDLINE).

Пример использования базы данных Cochrane Library. Открываем титульную страницу Cochrane Library (1-й выпуск 2000 г.; рис. 12). Для поиска данных об эффективности гипотензивной терапии при сахарном диабете вводим термины, отражающие основные компоненты клинического вопроса:

"diabetes AND hypertension AND mortality"
 ("артериальная гипертензия И сахарный диабет И смертность").

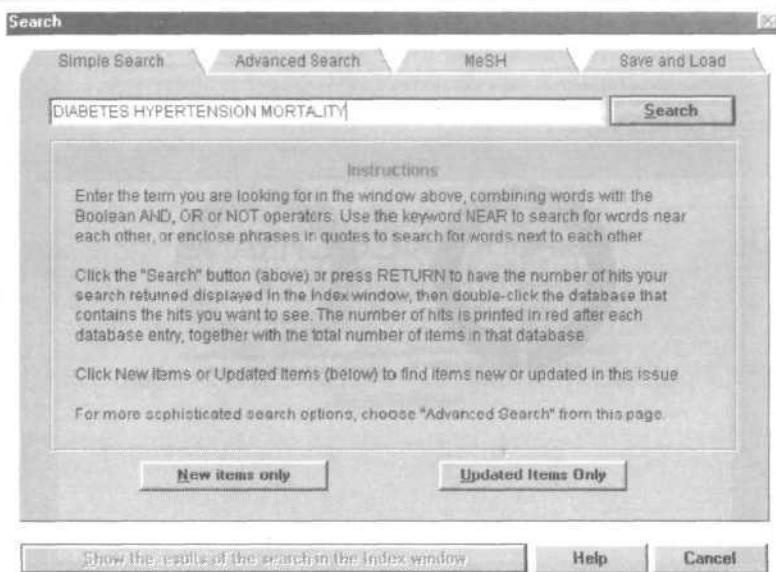


Рис. 12. База данных Cochrane Library (раздел Cochrane Database of Systematic Reviews): алгоритм поиска данных.

Воспроизводится с разрешения Update Software.

В результате такого поиска появится сообщение о том, что в разделе CDSR содержится 36 систематических обзоров, в разделе DARE — 6 обзоров, а в разделе CCTR — 130 публикаций (рис. 13). В разделе CDSR (рис. 14) выберем Кокрановский обзор "Antihypertensive therapy in diabetes mellitus" ("Гипотензивная терапия при сахарном диабете") [11]. Двойной щелчок на названии открывает полный текст систематического обзора, включая методы его проведения, критерии включения и исключения, результаты и обсуждение (рис. 15). Результаты представлены как в виде текста, так и в графической форме. Однако, как и в случае со статьей, найденной в Best Evidence, этот обзор не позволяет ответить на вопрос об оптимальном уровне АД, который следует поддерживать при артериальной гипертонии на фоне сахарного диабета.

Обратимся теперь к разделу CCTR. Двойной щелчок на его названии откроет список исследований. В числе первых 26 работ упоминаются исследования United Kingdom Prospective Diabetes Study Group (UKPDS)



Рис. 13. База данных Cochrane Library (раздел Cochrane Database of Systematic Reviews): результаты поиска.

Воспроизводится с разрешения Update Software.

10] и Hypertension Optimal Treatment (HOT) [12]. В ходе клинического опытыания HOT сравнивалась эффективность применения трех различных схем гипотензивной терапии. При анализе данных в подгруппах >блю установлено, что на фоне сахарного диабета поддержание диасто- [ического АД на уровне 81,1 мм рт. ст. по сравнению с его поддержани- м на уровне 85,2 мм рт. ст. снижает частоту развития сердечно-сосуди- тых заболеваний и смертности от них.

Приведем еще один пример использования раздела CCTR. В упоми- [авшейся выше клинической ситуации (см. *Пример 3. Тактика ведения ольных при плоскоклеточном раке легких*) возник вопрос о целесообразности применения двух дополнительных методов диагностики: первич- ой медиастиноскопии или компьютерной томографии грудной клет- и, на основании результатов которой будет определяться дальнейшая актика ведения больного. После введения ключевого слова "mediasti- oscopy" открывается список из 20 статей, 4-я и 5-я из которых пред- авляют собой индексированную в MEDLINE и EMBASE ссылку на



Рис. 14. База данных Cochrane Library (раздел Cochrane Database of Systematic Reviews): выявление интересующего систематического обзора.

Воспроизводится с разрешения Update Software.

одно и то же РКИ. В это РКИ были включены 685 больных, у которых немелкоклеточный рак легкого исходно расценивался как операбельный. В первой группе всем больным проводилась первичная медиастиноскопия, а во второй — компьютерная томография, по результатам которой решался вопрос о дальнейшем ведении больного (при выявлении незначительно увеличенных лимфатических узлов сразу выполнялась торакотомия, при выявлении крупных лимфатических узлов направлялся на медиастиноскопию). Показано, что во второй группе относительный риск выполнения необоснованной торакотомии составил 0,88 (при 95% ДИ от 0,71 до 1,10). При сравнении затрат на обследование и лечение разница между первой и второй группой составила 708 Долларов США (при 95% ДИ от -723 до +2140 долларов США). Исследователи сделали вывод о том, что "применение компьютерной томографии по сравнению с первичной медиастиноскопией характеризуется такой же или более низкой частотой выполнения необоснованной торакотомии и такими же или более низкими затратами" [13].



Рис. 15. База данных Cochrane Library (раздел Cochrane Database of Systematic Reviews): полный текст систематического обзора.

Воспроизводится с разрешения Update Software.

UpToDate

UpToDate представляет собой обширную учебную базу данных, обновляемую каждые 4 мес. Она существует только в электронном формате, поскольку в печатном виде была бы слишком объемной. В отличие от Best Evidence и Cochrane Library при подготовке этой базы данных не применяются специальные строгие критерии методологического качества; однако UpToDate содержит много высококачественных исследований, отобранных составителями разделов.

Пример использования базы данных UpToDate. Открываем титульную страницу UpToDate (версия 8.3). Для поиска данных об эффективности гипотензивной терапии на фоне сахарного диабета II типа вводим в окне поиска ключевое слово "diabetes". В открывшемся диалоговом окне (рис. 16) увидим 21 расширенный вариант этого ключевого слова. Выберем изданного перечня "diabetes mellitus, type II" ("сахарный диабет II типа"). В результате такого поиска появится список, включающий 64 статьи, одна из которых озаглавлена "Treatment of hypertension in diabetes" ("Ле-

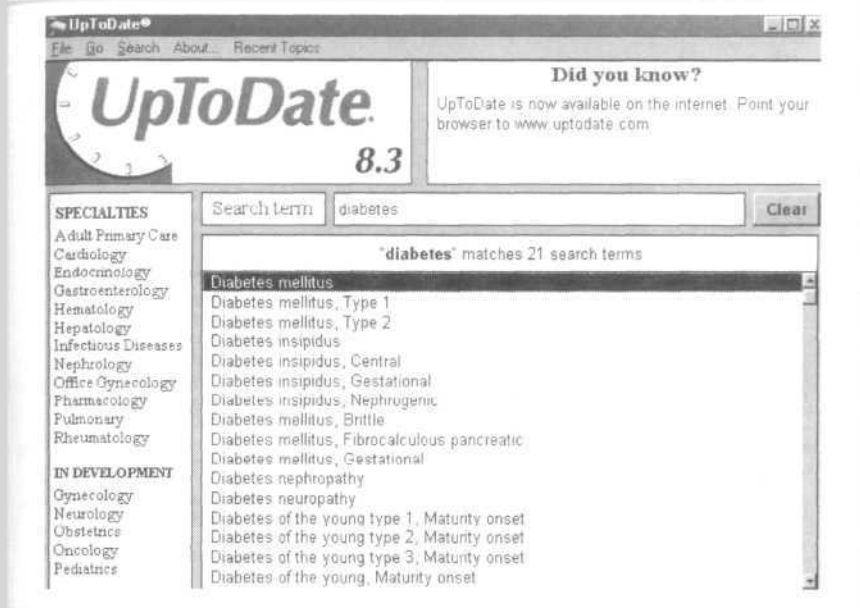


Рис. 16. База данных UpToDate: результаты первичного поиска по ключевому слову.

Воспроизводится с разрешения UpToDate.

ление артериальной гипертонии при сахарном диабете"). Двойной щелчок на названии этой статьи открывает ее полный текст. В статье рассматриваются патогенетические механизмы развития артериальной гипертонии на фоне сахарного диабета и имеется раздел, посвященный оптимальному уровню АД при проведении гипотензивной терапии (рис. 17 и 18). В этом разделе содержится подробное описание двух крупных РКИ (HOT [12] и UKPDS [10]), в которых сравнивались клинические исходы при более и менее выраженном снижении АД. В тексте приводятся данные о структуре испытаний и полученных результатах; щелчок на соответствующих ссылках позволяет открыть рефераты отчетов об этих РКИ.

Clinical Evidence

База данных Clinical Evidence выпускается издательской группой *BMJ* и одним из американских медицинских обществ American College of Physicians — American Society of Internal Medicine; эта база построена по тому

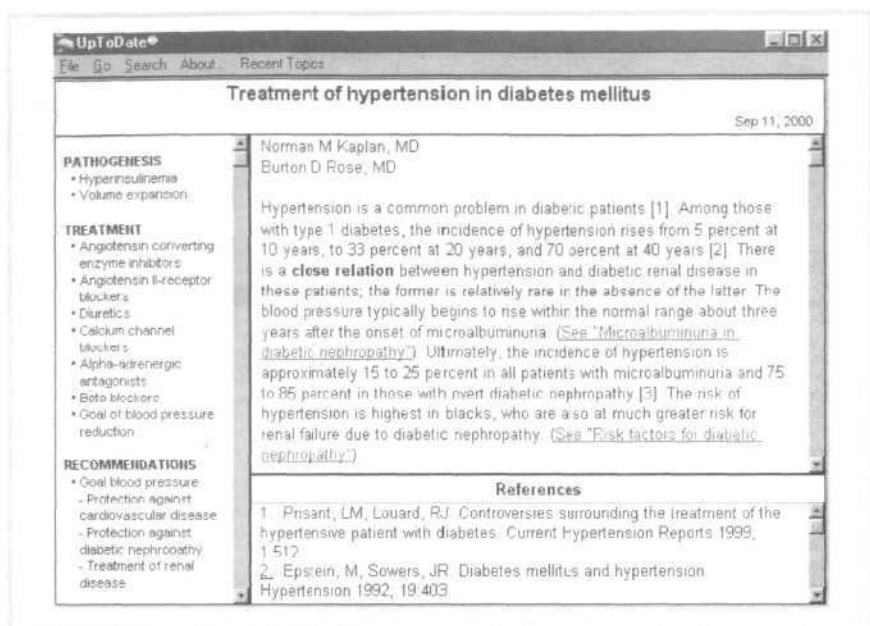


Рис. 17. База данных UpToDate: полный текст обзорной статьи (начало).

Воспроизводится с разрешения UpToDate.

же принципу, что и UpToDate, однако при ее составлении в меньшей степени учитывается мнение экспертов. Clinical Evidence доступна в печатном и электронном виде. Эта база данных представляет собой не учебное руководство, а краткий анализ самых современных знаний о профилактике и лечении наиболее распространенных и клинически значимых заболеваний или состояний; при этом особо оговариваются случаи, когда информация отсутствует или неубедительна. Clinical Evidence выпускается два раза в год, в том числе на сайте в Интернете (www.evidence.org).

Пример использования базы данных Clinical Evidence. Для поиска данных об эффективности гипотензивной терапии при сахарном диабете введем в окне поиска следующие термины:

"target blood pressure AND diabetes"

("поддержание АД на определенном уровне И сахарный диабет").

В открывшемся окне (рис. 19) увидим название основной темы "Cardiovascular disease in diabetes" ("Сердечно-сосудистые заболевания при сахарном диабете"), раздела "Which interventions improve cardiovascular

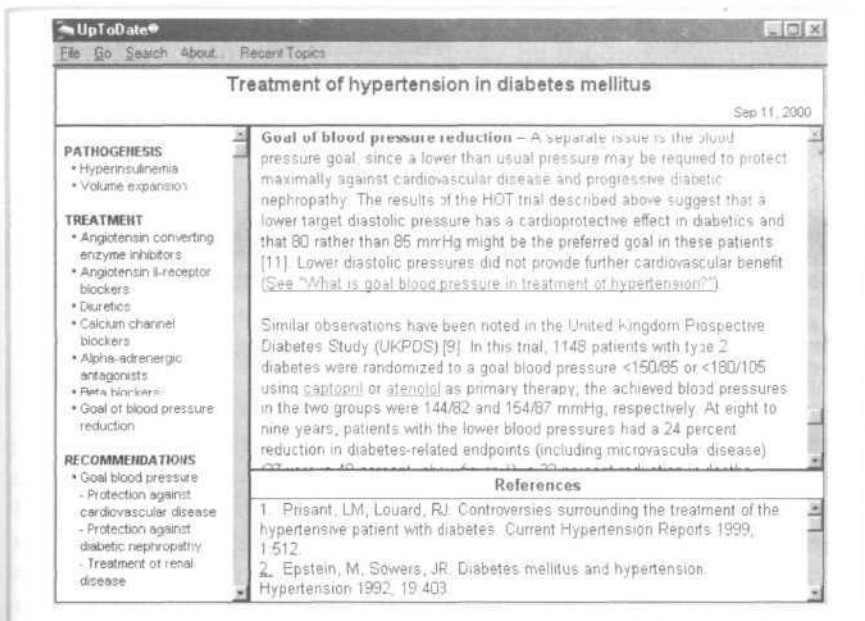


Рис. 18. База данных UpToDate: раздел обзора, посвященный оптимальному уровню АД при проведении гипотензивной терапии.

Воспроизводится с разрешения UpToDate.

outcomes in patients with diabetes?" ("Какие вмешательства улучшают клинические исходы, связанные с сердечно-сосудистыми заболеваниями, при сахарном диабете?") и подраздела "Antihypertensive treatment" ("Гипотензивная терапия"). В этом подразделе обсуждается оптимальный уровень АД, который необходимо поддерживать при проведении гипотензивной терапии на фоне сахарного диабета, и приводятся результаты уже упоминавшихся **РКИ** (HDT и UKPDS).

МЕДИЦИНСКИЕ ЭЛЕКТРОННЫЕ БАЗЫ ДАННЫХ, В КОТОРЫХ НЕ ИСПОЛЗУЮТСЯ КРИТЕРИИ МЕТОДОЛОГИЧЕСКОГО КАЧЕСТВА

MEDLINE

Если поиск ответа на четко сформулированный клинический вопрос в базах данных Best Evidence, Cochrane Library, UpToDate и Clinical Evidence оказался безрезультатным, пришло время обратиться к MEDLINE. Эта впечатляющая по объему библиографическая база данных, вклю-

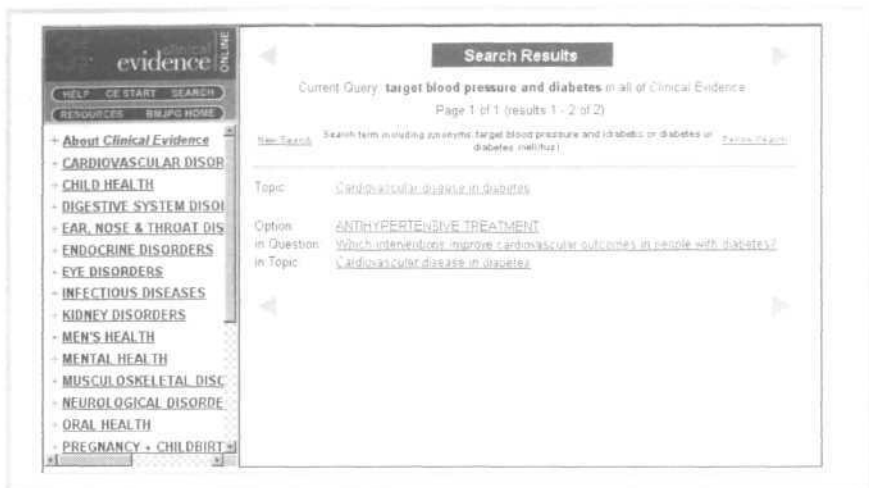


Рис. 19. База данных Clinical Evidence: результаты поиска.

Воспроизводится с разрешения BMJ Publishing Group.

чающая более 11 000 000 ссылок на публикации о клинических и доклинических исследованиях, создана и поддерживается Национальной медицинской библиотекой США. Ее дополнительный раздел PreMEDLINE содержит информацию о недавно опубликованных отчетах, которые еще не были индексированы. К несомненным достоинствам MEDLINE следует отнести почти полный охват всех медицинских журналов, а также ее доступность. Подключившись к Интернету, можно бесплатно осуществлять поиск информации в этой базе данных с помощью поисковых систем PubMed или Internet Grateful Med. Кроме того, большинство медицинских научных и больничных библиотек предоставляют доступ к MEDLINE, осуществляемый с помощью платных приложений, таких как OVID, Knowledge Finder или Silver Platter.

К недостаткам этой базы данных относятся прежде всего ее колоссальные размеры и неоднородность размещенных в ней материалов. Чтобы поиск в MEDLINE был эффективным, следует тщательно продумать его стратегию, хорошо разбираться в структуре этой базы данных и процедуре индексации статей. Необходимо уметь пользоваться медицинским предметным рубрикатом (MeSH) с разворачиванием списка, выбирать термины для поиска в тексте рефератов и логические операторы "AND" ("и") и "OR" ("или") для объединения результатов поиска по нескольким запросам.

Для знакомства с методами поиска информации в MEDLINE стоит прочесть статью Т. Greenhalgh [14]. В отсутствие необходимых навыков лучше потратить некоторое время на работу под руководством опытного библиотечного работника или пройти обучение на специальных курсах. Кроме того, можно посетить учебный сайт в Интернете (www.docnet.org.uk/dfelix/medtut.html), помогающий пользоваться различными разделами MEDLINE при разных уровнях компьютерной грамотности. Более подробная информация о поиске данных в MEDLINE и целом ряде других объемных баз данных, включая EMBASE/Excerpta Medica, можно найти в книге А. McKibbin [15]. В этом разделе представлены лишь основные советы по работе с MEDLINE.

При индексации каждой статьи составители MEDLINE используют рубрику MeSH, с помощью которой данную статью можно найти в процессе поиска. Однако при этом выбираются ключевые слова, наиболее точно отражающие содержание статьи (например, "нарушение функции левого желудочка", а не "нарушение функции желудочков сердца"). Если при формулировании параметров поиска будет указан термин, имеющий более широкое значение ("нарушение функции желудочков сердца"), многие интересующие статьи могут оказаться невыявленными. Для решения этой проблемы предназначена команда "explode" ("расширить поиск"), которая позволяет найти все рефераты, проиндексированные с использованием введенного термина и родственных ему понятий. Например, в PubMed MEDLINE насчитывается 10 806 ссылок, вносящихся к предметной рубрике "sports" ("спорт") и опубликованных с 1966 по 2000 г. Если же при определении параметров поиска к ключевому слову "sports" добавить команду "explode", список пополнится рефератами по отдельным видам спорта (от бейсбола и баскетбола до тяжелой атлетики и армрестлинга) и будет включать уже 37 043 ссылки.

Поиск можно проводить и с использованием функции "text word". В этом случае пользователь вводит определенное слово, которое затем будет найдено в заголовке и тексте всех имеющихся рефератов. По мере накопления опыта работы с MEDLINE каждый врач выбирает наиболее удобную для себя стратегию поиска. Для более полного выявления ссылок по интересующей теме целесообразно комбинировать оба основных способа поиска данных в MEDLINE.

Пример использования базы данных MEDLINE. Для получения данных об эффективности гипотензивной терапии при сахарном диабете II типа из базы данных MEDLINE воспользуемся поисковой системой PubMed. Введем ключевой термин "diabetes mellitus" ("сахарный диабет") и шелк-

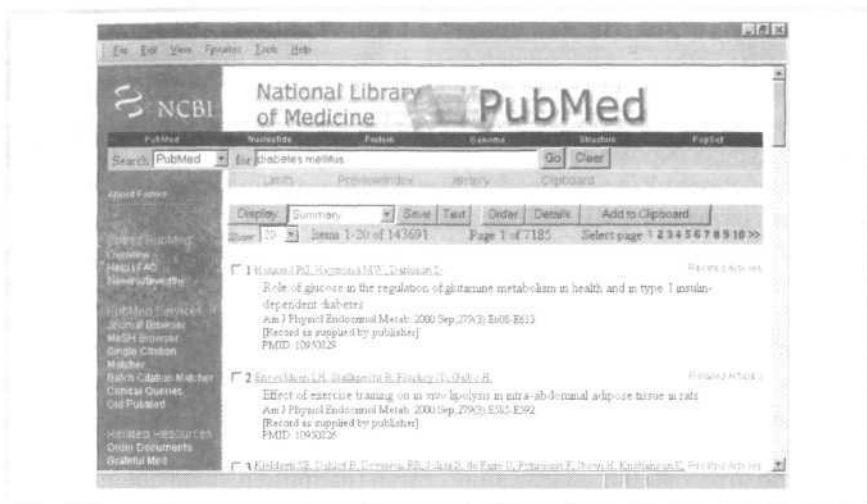


Рис. 20. База данных PubMed: результат поиска по термину "сахарный диабет". Воспроизводится с разрешения Национальной медицинской библиотеки США.

нем по кнопке "Go". В результате такого поиска появится список (рис. 20), включающий 143 691 ссылку на статьи, опубликованные с 1966 г. Следует отметить, что при использовании системы PubMed поиску данных в базах MEDLINE и PreMEDLINE предшествует этап обработки запроса с определением стратегии, подразумевающей поиск не только по введенному термину, но и по предметной рубрике, более всего соответствующей теме запроса. С целью увеличения числа выявленных ссылок автоматически включается расширение поиска. В окончательном виде запрос выглядит следующим образом:

diabetes mellitus ("сахарный диабет"; термин для поиска в тексте рефератов) OR ("ИЛИ"; логический оператор) explode diabetes mellitus (команда расширенного поиска по предметным рубрикам).

Логический оператор "или" применяется для объединения результатов поиска по двум запросам и более полного выявления публикаций, в которых обсуждаются какие-либо аспекты сахарного диабета.

После этого аналогичным образом осуществляем поиск по ключевым словам "hypertension" ("артериальная гипертензия"; 180 333 ссылки) и "mortality" ("смертность"; 320 133 ссылки). Для объединения результатов, полученных по всем трем запросам, следует нажать на кнопку "History" ("Все запросы") и ввести следующие символы:

"#1 AND #2 AND #3".

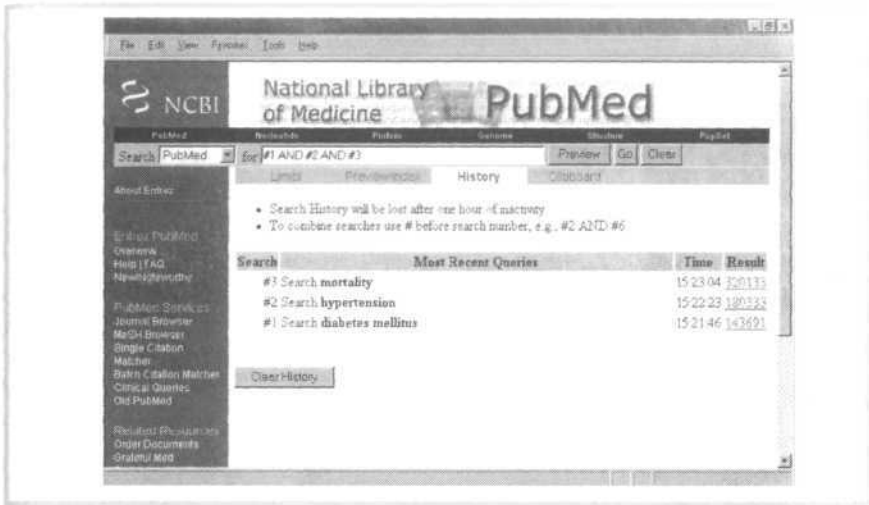


Рис. 21. База данных PubMed: объединение результатов поиска по трем терминам (начало).

Воспроизводится с разрешения Национальной медицинской библиотеки США.

В этом случае система PubMed представит список только тех статей, в которых присутствуют все три термина (рис. 21).

К сожалению, даже после этого список насчитывает 1965 ссылок (рис. 22), поэтому целесообразно несколько видоизменить параметры поиска, уточнив структуру интересующих нас исследований. При этом можно прибегнуть к апробированному систематическим образом "фильтрам", которые позволяют отобрать исследования высокого методологического качества, посвященные изучению эффективности и безопасности лечения, диагностике или прогнозу (рис. 23). Например, высокочувствительный поиск прогностических исследований может начинаться следующим образом:

"incidence ("заболеваемость"; термин для поиска по предметным рубрикам)

OR ("ИЛИ") explode mortality (команда расширенного поиска по предметным рубрикам)

OR follow-up studies ("когортные исследования"; термин для поиска по предметным рубрикам)

OR mortality ("смертность"; термин для поиска по подрубрикам)

OR prognos: (корень ключевого слова для поиска в тексте рефератов)

OR predict: (корень ключевого слова для поиска в тексте рефератов)

OR course ("течение"; термин для поиска в тексте рефератов).

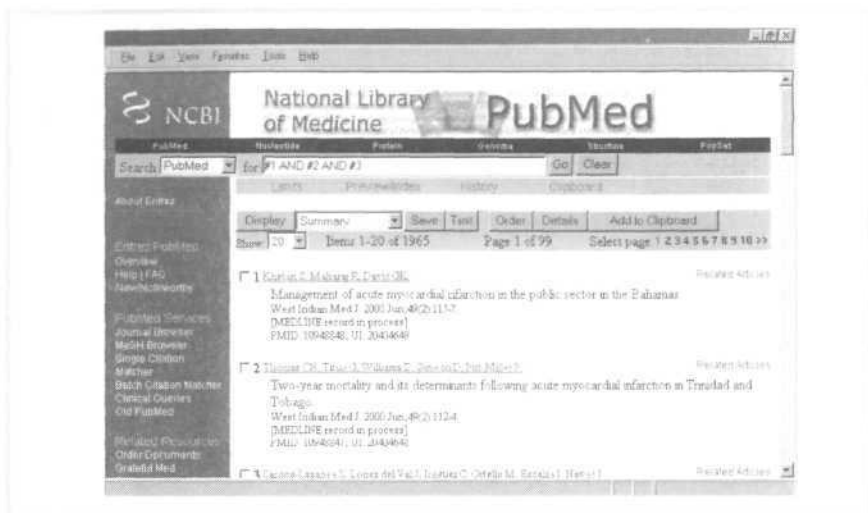


Рис. 22. База данных PubMed: объединение результатов поиска по трем терминам (окончание).

Воспроизводится с разрешения Национальной медицинской библиотеки США.

Выбранная затем стратегия высокоспецифичного поиска будет выглядеть так:

"prognosis ("прогноз"; термин для поиска по предметным рубрикам)
 OR survival analysis" ("анализ выживаемости"; термин для поиска по предметным рубрикам).

Высокая чувствительность поиска подразумевает большой объем найденной информации, часть которой не будет иметь отношения к интересующему нас вопросу, в то время как высокоспецифичный поиск заканчивается выявлением меньшего числа ссылок при более высокой их актуальности. Перечень различных стратегий поиска с указанием специфичности и чувствительности каждой из них можно найти в соответствующих публикациях [16—18]. Несмотря на то что выработка стратегии поиска достаточно сложна и многоступенчата, в большинство программ для поиска данных в MEDLINE встроены шаблоны, позволяющие автоматизировать этот процесс. Так, в левой части титульной страницы PubMed находится кнопка "Clinical queries" ("Клинические вопросы"), которая активизирует функцию, позволяющую выбрать различные подходы к поиску данных,

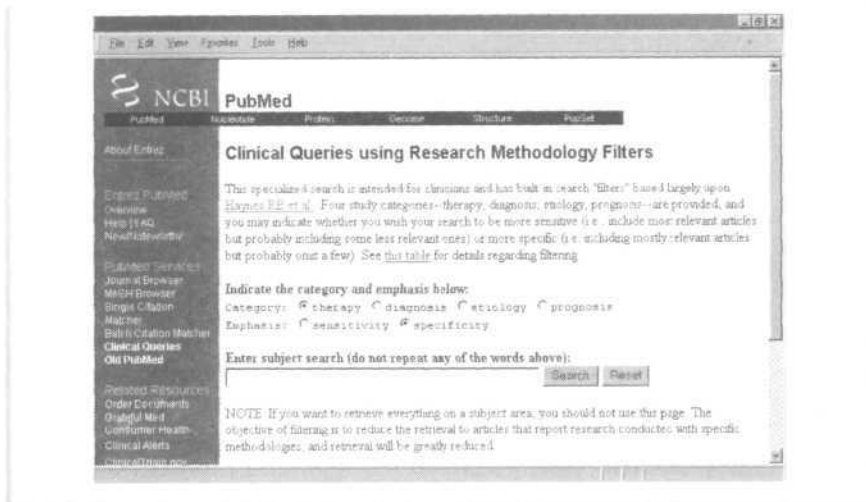


Рис. 23. База данных PubMed: поиск с использованием шаблонной стратегии. Воспроизводится с разрешения Национальной медицинской библиотеки США.

В качестве альтернативы можно использовать так называемые универсальные термины, в том числе "clinical trial" ("клиническое испытание"; структура исследования) при выявлении исследований, посвященных оценке эффективности методов лечения, "sensitivity" ("чувствительность"; термин для поиска в тексте рефератов) при выявлении исследований, посвященных оценке информативности диагностических методов, "explode cohort studies" (команда расширенного поиска по предметным рубрикам) при выявлении исследований, посвященных изучению прогноза, "harm" ("риск"; термин для поиска в тексте рефератов) при выявлении исследований, посвященных оценке факторов риска или безопасности методов лечения (см. табл. 2).

Если в предыдущем запросе обозначить интересующую структуру исследования, включив термин "clinical trial" ("клиническое испытание"), число найденных ссылок уменьшится до 117 (рис. 24). Среди них снова можно найти материалы испытаний UKPDS и HOT.

Другим достоинством поисковой системы PubMed можно считать простоту ее применения, благодаря которой даже пользователь, не искушенный в построении запроса, может ввести несколько ключевых слов или фраз в первом текстовом окне. В упоминавшейся выше клиниче-

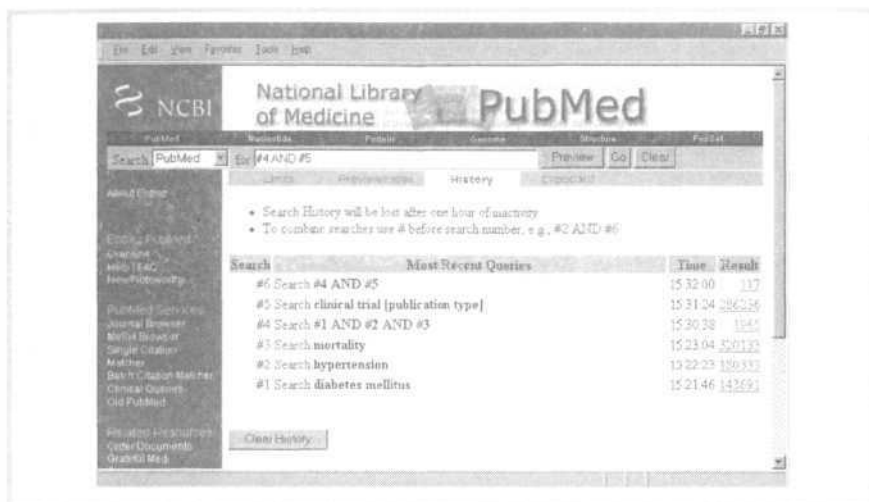


Рис. 24. База данных PubMed: алгоритм поиска с дополнительным использованием универсального термина.

Воспроизводится с разрешения Национальной медицинской библиотеки США.

ской ситуации (см. *Пример 2. Тактика ведения больных при предполагаемой нестабильной стенокардии*) запрос может выглядеть следующим образом:

"unstable angina ("нестабильная стенокардия") monitor*
(общий корень ключевых

слов "монитор", "мониторинг" и т.д.) discharg**

(общий корень ключевых слов "выписка", "выписывать" и т.д.).

Звездочка указывает на необходимость включения в список рефератов, в которых встречаются слова с указанным общим корнем. В результате такого поиска появится список из 26 ссылок. Под номером 11 в списке значится обзорная статья [19]. Щелчок на кнопке "See related" ("Смотри похожие ссылки") открывает окно с 238 ссылками, расположенными в порядке убывания соответствия теме запроса, которое автоматически оценивает поисковая программа.

Интернет

За короткое время Интернет ("всемирная паутина", World Wide Web) превратился в важный источник медицинских знаний. В нем содержится масса полезной информации, распространяемой бесплатно или по

подписке. Чтобы сделать эту информацию еще доступнее, многие сайты (например, Medical Matrix, Medscape, ScHARR и Medical World Search) специально организованы таким образом, чтобы с помощью гиперссылок обеспечить связь между разными базами данных и облегчить поиск медицинских данных. С помощью Интернета можно также знакомиться с содержанием многих медицинских журналов и клинических рекомендаций.

Однако необходимо помнить, что далеко не все из найденных рекомендаций будут соответствовать критериям доказательной медицины (см. главу 11 *От доказательств к практике*). Платный доступ ко многим ресурсам, в том числе журналам, учебникам и руководствам, можно получить на сайте MD Consult.

Наконец, многие сайты созданы и поддерживаются такими известными организациями, как, например, Американское общество по борьбе с раком (www.cancer.org) или Американская диабетическая ассоциация (www.diabetes.org). Обращение к этим специализированным сайтам представляет собой еще один источник необходимой информации.

ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ НАЙДЕННОЙ ДОКАЗАТЕЛЬНОЙ ИНФОРМАЦИИ

Число медицинских публикаций крайне велико и продолжает стремительно расти. Это свидетельствует о проведении все новых исследований и вселяет надежду на улучшение качества медицинской помощи, но, с другой стороны, существенно затрудняет поиск наиболее достоверных и современных данных. Поэтому можно только приветствовать появление новых информационных технологий, позволяющих быстро и с минимальными усилиями получить такие сведения.

Поиск статей, содержащих ответ на конкретный клинический вопрос, займет от 5 до 30 мин в зависимости от используемых источников доказательной информации или от опыта работы с такими источниками [3]. Еще полчаса может понадобиться для оценки достоверности и применимости выявленных данных. Так, в одной из рассмотренных выше клинических ситуаций (см. *Пример 1. Оптимальный уровень АД, который следует поддерживать у больных с артериальной гипертензией на фоне сахарного диабета*) в наибольшей степени применимы результаты испытаний UKPDS [10] и HOT [12]. Они свидетельствуют о том, что при сахарном Диабете II типа проведение гипотензивной терапии, направленной на поддержание более низкого уровня АД, способствует снижению смерт-

ности, связанной с осложнениями сахарного диабета. На основании этих данных врач принимает решение о том, что его больной требуется проводить более интенсивную гипотензивную терапию для поддержания систолического АД на уровне < 150 мм рт. ст., а диастолического АД на уровне < 80 мм рт. ст.

Литература

1. *Haynes R.B.* Loose connections between peer reviewed clinical journals and clinical practice. *Ann Intern Med* 1990;113:724-728.
2. *Oxman A.D., Sackett D.L., Guyatt G.H.* For the Evidence-based Medicine Working Group. Users' guides to the medical literature, I: How to get started. *JAMA* 1993;270:2093-2095.
3. *Richardson W.S., Wilson M.C., Nishikawa J., Hayward R.S.A.* The well-built clinical question: a key to evidence-based decisions. *ACP Journal Club*. Nov-Dec 1995; 123:A-12.
4. *McKibbin K.L., Richardson W.S., Walker Dilks C* Finding answers to well built clinical questions. *Evidence-Based Medicine* 1999;6:164-167.
5. *Sackett D.L., Straus S.E.* Finding and applying evidence during clinical rounds: the "evidence cart." *JAMA* 1998;280:1336-1338.
6. *Covell D.G., Uman G.C., Manning P.R.* Information needs in office practice: are they being met? *Ann Intern Med* 1985; 103:596-599.
7. *Osheroff J.A., Forsythe D.E., Buchanan B.G., Bankowitz R.A., Blumenfeld B.H., Miller R.A.* Physicians' information needs: analysis of questions posed during clinical teaching. *Ann Intern Med* 1991; 114:576-581.
8. *Barrie A.R., Ward A.M.* Questioning behaviour in general practice: a pragmatic study. *BMJ* 1997;315:1515.
9. *Ely J. W., Osheroff J.A., Ebell M.H., et al.* Analysis of questions asked by family doctors regarding patient care. *BMJ* 1999;319:358-361.
10. Tight blood pressure control and risk of macrovascular and microvascular complications in Type 2 diabetes mellitus: UKPDS 38. *BMJ* 1998;317:703-713.
11. *Fuller J., Stevens L.K., Chaturvedi N., Holloway J.F.* Antihypertensive therapy in diabetes mellitus (Cochrane Review). In: *The Cochrane Library*, Issue 4, 1999. Oxford: Update Software.
12. *Hansson L., Zanchetti A., Carruthers S.G., et al., for the HOT Study Group.* Effects of intensive blood-pressure lowering and low-dose aspirin in patients with hypertension: principal results of the Hypertension Optimal Treatment (HOT) randomized trial. *Lancet* 1998;351:1755-1762.
13. *The Canadian Lung Oncology Group.* Investigation for mediastinal disease in patients with apparently operable lung cancer. *Ann Thorac Surg* 1995;60:1382-1389.
14. *Greenhalgh T.* How to read a paper: the Medline database. *BMJ* 1997;315:180-183.
15. *McKibbin K.A.* Evidence-based Principles and Practice. Hamilton: Decker; 1999.
16. *Haynes R.B., Wilczynski N., McKibbin K.A., et al.* Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *J Am Med Informatics Assoc* 1994; 1:447-458.
17. *Wilczynski N.L., Walker C.J., McKibbin K.A., Haynes R.B.* Assessment of methodologic search filters in MEDLINE. *Proc Ann Symp Comp Appl Med Care* 1994; 17:601-605.
18. *Deville W.L., Bezemer P.D., Bouter L.M.* Publications on diagnostic test evaluation in family medicine journals: an optimal search strategy. *J Clin Epidemiol* 2000;53:65-69.
19. *Bankwala Z, Swenson L.J.* Unstable angina pectoris: what is the likelihood of further cardiac events? *Postgrad Med* 1995;98:155-158,161-162,164-165.

3**ЛЕЧЕНИЕ И РИСК:
ВВЕДЕНИЕ**

Г. Гайятт

В подготовке данной главы принимал участие П. Уай

В ЭТОМ РАЗДЕЛЕ

Три этапа оценки результатов исследований, опубликованных в **медицинской литературе**

Лечение и **риск**: структура исследований

Рандомизированные контролируемые испытания, посвященные оценке эффективности лечения

Обсервационные исследования, посвященные оценке влияния факторов риска

Правильный выбор критериев оценки

К наиболее сложным клиническим вопросам относится выбор оптимальной тактики лечения у конкретного больного. Например, врач должен хорошо представлять себе преимущества лекарственной терапии и специальных диет, применяемых с целью снижения артериального давления, концентрации холестерина в крови или массы тела больного; скринингового обследования женщин для выявления рака молочной «елезы; скринингового обследования мужчин с целью выявления рака предстательной железы; программы, направленной на отказ от курения. Он должен объяснять больным, как различные методы лечения повлияют на продолжительность жизни и выраженность симптомов хронической сердечной недостаточности, бронхиальной астмы или сахарного диабета. Для больных не менее важно иметь представление о ранних и отдаленных побочных эффектах, которые могут возникать в результате применения определенного лечебного вмешательства.

Любой из этих вопросов содержит две взаимосвязанные составляющие. Во-первых, следует знать, к каким неблагоприятным последствиям (если таковые вообще существуют) может привести нарушение функции сердца, курение, ожирение, повышение артериального давления, концентрации холестерина или глюкозы в крови? Предполагается, что наличие этих и некоторых других факторов риска способствует развитию заболеваний, т.е. эта составляющая клинического вопроса относится к оценке риска. Во-вторых, каковы возможные благоприятные последствия вмешательств, направленных на изменение поведения или стиля жизни, а также будет ли польза от этих вмешательств перевешивать связанные с их применением неблагоприятные последствия? Эта составляющая клинического вопроса относится к оценке эффективности лечения.

Когда врачу приходится одновременно решать вопросы эффективности лечебного вмешательства и оценить риск, возникает проблема причинно-следственных отношений. Можно привести массу примеров. Гак, надо знать, существует ли в определенных популяциях (здоровые мужчины или женщины, больные с сердечно-сосудистой недостаточностью или сахарным диабетом и т.д.) причинно-следственная связь между имеющимися факторами риска (курение, ожирение, высокое артериальное давление) или применяемыми вмешательствами (программы, направленные на отказ от курения или снижение избыточной массы тела, прием гипотензивных препаратов) и развитием ожидаемых (рак легких, инфаркт миокарда, инсульт) или неожиданных (например, прогрессирующее нарушение зрения в результате приема гипотензивного препарата [1]) клинических исходов.

Будь наши знания полными, мы смогли бы ответить на любой вопрос; однако гораздо чаще приходится строить предположения. Если они ложны, последствия могут оказаться катастрофическими. Трудно даже приблизительно подсчитать, сколько больных погибло за те несколько веков, когда врачи были убеждены, что кровопускание эффективно при самых разных болезнях. Напротив, учет клинических исходов в ходе множества РКИ с известным числом участников позволяет с большой долей уверенности предположить, что за время применения противоритмических препаратов I класса были потеряны тысячи жизней (по мнению врачей, использование этих препаратов должно было предотвращать возникновение смертельно опасных аритмий, но в действительности вызывало их) [2].

Почему медицинское сообщество совершало столь чудовищные ошибки, и как можно предотвратить их в дальнейшем? Ответ заключается в том, что врачи обязаны знать и применять в своей повседневной работе правила, позволяющие отличить ложные выводы и предположения от достоверных. В данной главе читатель может ознакомиться лишь с основными практическими аспектами применения этих правил. Для более глубокого понимания данной проблемы следует обратиться к главе 1 раздела II *Лечение и риск: причины недостоверности результатов — систематические и случайные ошибки*.

ТРИ ЭТАПА ОЦЕНКИ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ, ОПУБЛИКОВАННЫХ В МЕДИЦИНСКОЙ ЛИТЕРАТУРЕ

Материалы исследований, выявленные в медицинской литературе при поиске ответа на конкретный клинический вопрос, следует оценивать в три этапа.

На первом этапе нужно ответить на вопрос: "Достоверны ли результаты исследования?". Речь идет о том, можно ли вообще доверять полученным данным. При этом следует обратить внимание, насколько вывод об эффективности лечения отражает направление и величину истинного клинического эффекта. Этот вопрос можно сформулировать иначе: "Не были ли результаты подвержены влиянию какой-либо систематической ошибки, что привело к неверным выводам?". Если в исследовании использовались методы, позволяющие избежать возникновения систематической ошибки, его результаты скорее всего достоверны, и статья заслуживает внимательного изучения.

На втором этапе нужно ответить на вопрос: "Каковы результаты следования?". Здесь речь идет об оценке размера выявленного исследователями эффекта лечения и точности его расчета. Наилучшим образом размер эффекта отражают результаты испытания в том виде, в каком < были получены; точность расчетов обычно выше в более крупных следованиях.

И наконец, следует задать себе третий вопрос: "Каким образом я м использовать эти результаты в своей повседневной практике?". 3 вопрос можно разбить на две части. Во-первых, применимы ли результаты исследования (полностью или частично) в определенной клинической ситуации? Так, вряд ли целесообразно назначать лечебное вмешательство, эффективность которого изучалась в ходе РКИ, если характеристики конкретного больного и участников испытания существенно различаются. Во-вторых, если результаты исследования применимы данному случаю, каково будет общее влияние лечения? Оценивали ли исследователи все клинические исходы, важные для больных? Об влиянии лечебного вмешательства определяется его преимущества недостатками (включая побочные эффекты и токсическое действие также последствиями отмены лечения. Так, даже самое эффективное лечение иногда может быть отменено, если прогноз благоприятен и лечения (особенно в тех случаях, когда вмешательство характеризуется тяжелыми побочными эффектами или выраженной токсичностью).

ЛЕЧЕНИЕ И РИСК: СТРУКТУРА ИССЛЕДОВАНИЙ

РАНДОМИЗИРОВАННЫЕ КОНТРОЛИРУЕМЫЕ ИСПЫТАНИЯ, ПОСВЯЩЕННЫЕ ОЦЕНКЕ ЭФФЕКТИВНОСТИ ЛЕЧЕНИЯ

При оценке эффективности лечения контролю за условиями применения изучаемого вмешательства уделяется гораздо большее внимание, чем при оценке влияния факторов риска. Так, исследователи формируют группы вмешательства и контроля (например, с отсутствием лечения или применением плацебо). В идеале больных следует распределять по этим группам в случайном порядке, используя метод, принципиально не отличающийся от подбрасывания монеты. Такое распределение называется рандомизацией, а исследование — рандомизированным контролируемым испытанием (РКИ). Кроме того, можно организовать исследование таким образом, чтобы ни больные, ни медицинский персонал не знали, к какой группе относится конкретный больной.

ОБСЕРВАЦИОННЫЕ ИССЛЕДОВАНИЯ, ПОСВЯЩЕННЫЕ ОЦЕНКЕ ВЛИЯНИЯ ФАКТОРОВ РИСКА

Напротив, при оценке влияния факторов риска подобные методы контроля обычно неприменимы. Невозможно приказывать больным, чтобы они жили в местах с высокой или низкой загрязненностью окружающей среды, с низкой или высокой плотностью заселения и т.п. Исследователи не могут скрыть от участников характеристики среды, в которой они живут, или наличие тех или иных факторов риска (например, курение). В этих случаях организуется проведение исследований, которые называются обсервационными. В ходе таких исследований можно осуществлять наблюдение за участниками, которые подвержены каким-либо неблагоприятным воздействиям (вследствие собственного предпочтения или жизненных обстоятельств). Наблюдение продолжается в течение определенного времени, чтобы выяснить, развилось ли у участников изучаемое заболевание или состояние (эти исследования называются когортными). Можно использовать и другой подход: набрать группу лиц, у которых уже возник изучаемый клинический исход, и контрольную группу лиц без этого клинического исхода, а затем сравнить степень влияния предполагаемого неблагоприятного фактора в обеих группах (исследование случай—контроль; см. главу 5 *Оценка риска*).

ПРАВИЛЬНЫЙ ВЫБОР КРИТЕРИЕВ ОЦЕНКИ

Результаты исследований, в которых изучается влияние факторов риска, в целом менее достоверны, чем результаты испытаний методов лечения. Правильный подход к работе с медицинской литературой предусматривает использование разных критериев оценки применительно к результатам исследований этих двух типов. Поэтому в данной книге рекомендации по работе с соответствующими источниками излагаются отдельно (см. главы 4 *Исследования, посвященные эффективности лечения* и 5 *Оценка риска*).

Однако существуют исключения из этого общего правила. Иногда фактором риска может стать изучаемое лечебное вмешательство (например, применение лекарственного препарата), и неблагоприятные клинические исходы достаточно быстро развиваются у многих больных. В таком случае выявление причинно-следственной зависимости между применением препарата и возникновением токсической реакции может быть проведено в рамках того же исследования (хотя исследования подобной структуры обычно применяются для оценки эффективности лечения).

Кроме того, в медицинской литературе вообще могут отсутствовать материалы РКИ по определенным проблемам (порой такие РКИ в принципе неосуществимы). Структура РКИ не всегда применима при изучении редких заболеваний, оценке эффективности амбулаторных вмешательств, лечения в разных больницах или качества медицинской помощи в условиях одного лечебного учреждения. Использование рандомизации при сравнении эффективности первичной и специализированной медицинской помощи, оплаты за проведенные манипуляции и подушевого финансирования, оплаты медицинских услуг из государственного бюджета или кармана больного представляется невероятным, по меньшей мере в ближайшем будущем.

Если при поиске ответа на клинический вопрос врач не находит материалов РКИ, он должен опираться на результаты когортных исследований и исследований случай—контроль; в данной ситуации они будут источниками самой достоверной информации. При оценке методологического качества этих исследований следует применять соответствующие критерии оценки (т.е. критерии оценки качества исследований, посвященных изучению неблагоприятных воздействий). Однако необходимо помнить, что любые выводы о причинно-следственной связи между лечебным вмешательством и клиническим исходом, сделанные в ходе когортных исследований и исследований случай—контроль, существенно уступают в достоверности результатам РКИ.

Литература

1. *Wright P.* Untoward effects associated with practolol administration: oculomucocutaneous syndrome. *BMJ* 1975;1:595—598.
2. *Moore T.J.* *Deadly Medicine*. New York, NY: Simon & Schuster; 1995.

4**ИССЛЕДОВАНИЯ,
ПОСВЯЩЕННЫЕ
ЭФФЕКТИВНОСТИ ЛЕЧЕНИЯ**

Г. Гайятт, Д. Кук, П. Дж. Деверо, М. Мид,
Ш. Строе

В подготовке данной главы принимали также участие члены Рабочей группы по разработке и внедрению принципов доказательной медицины П. Уайер, Р. Йешке, Д. Хейланд, А. Холбрук и Л.М. Летельер

В ЭТОМ РАЗДЕЛЕ

Поиск доказательной информации

Достоверны ли результаты исследования?

Проводилась ли рандомизация?

Использовалось ли сокрытие порядка рандомизации?

Учитывалось ли при анализе данных первоначальное распределение больных по группам?

Были ли группы сходны по всем известным факторам, влияющим на прогноз?

Знали ли больные о том, к какой группе они были отнесены в процессе рандомизации?

Знали ли врачи о том, к какой группе тот или иной больной был отнесен в процессе рандомизации?

Знали ли исследователи, оценивающие клинические исходы, о том, к какой группе тот или иной больной был отнесен в процессе рандомизации?

Все ли больные наблюдались до конца исследования?

Каковы **результаты** исследования?

Каков размер эффекта изучаемого вмешательства?

Насколько точным был расчет размера эффекта?

Как оценивать результаты исследования, если в отчете не приведены значения доверительного интервала?

Можно ли применить результаты исследования при лечении конкретного больного?

Совпадают ли характеристики участников исследования и конкретного больного?

Оценивались ли все клинически значимые исходы?

Перевешивают ли предполагаемые преимущества лечения все возможные побочные эффекты и затраты, связанные с его проведением?

Принятие предварительного решения

ВОЗМОЖНАЯ КЛИНИЧЕСКАЯ СИТУАЦИЯ

Из Интернета я узнал, что прием спиронолактона может увеличить продолжительность моей жизни.

Доктор, стоит ли мне начать прием этого препарата?

Представьте себя врачом общей практики. К вам на прием пришел 66-летний мужчина с идиопатической дилатационной кардиомиопатией, которого вы наблюдаете в течение 3 лет. Больной очень заботится о своем здоровье и активно участвует в обсуждении любых клинических решений. В этот раз он показывает вам сообщение из Интернета, озаглавленное "Спиронолактон спасает жизнь при сердечной недостаточности". Больного очень обнадежило сообщение о результатах нового исследования; он верит, что прием спиронолактона поможет продлить ему жизнь.

В предшествующие 18 мес состояние больного было стабильным, имеющиеся симптомы сердечной недостаточности слабо выражены и соответствуют II функциональному классу (ФК) по классификации Нью-Йоркской кардиологической ассоциации (New York Heart Association — NYHA). При электрокардиографии (ЭКГ), выполнявшейся 3 мес назад, подтверждено наличие общего нарушения функции левого желудочка, но его прогрессирования не выявлено; фракция выброса составила 30%. В настоящее время больной принимает эналаприл (по 10 мг 2 раза в сутки), метопролол (по 50 мг 2 раза в сутки) и фуросемид (по 20 мг 1 раз в сутки). Артериальное давление составляет 110/70 мм рт. ст., частота сердечных сокращений — 60 уд/мин. На предыдущей неделе концентрация креатинина в крови равнялась 100 мкмоль/л, концентрация калия — 4,1 мммоль/л. Зная, что эналаприл подавляет действие альдостерона, вы заинтересовались, может ли антагонист альдостерона спиронолактон усилить этот эффект. В *Настольном справочнике врача* сказано, что одновременное применение эналаприла и спиронолактона не всегда показано из-за риска развития гиперкалиемии [1]. Вы делитесь с больным своими сомнениями, заверив его, что тщательно проанализируете всю имеющуюся информацию и через неделю дадите ему исчерпывающий ответ.

Сначала необходимо сформулировать клинический вопрос, который в данном случае будет звучать следующим образом:

"Как влияет прием спиронолактона на смертность и качество жизни при сердечной недостаточности II ФК по NYHA со снижением фракции выброса?"

Поскольку искомый отчет об исследовании был опубликован не более 2 мес назад, его еще не включили в базу данных Best Evidence, с которой вы обычно начинаете поиск. Поэтому вы проведете поиск в базе данных MEDLINE с помощью поисковой системы OVID, а также ключевых слов "heart failure, congestive" ("сердечная недостаточность, застойная"; термин для поиска по предметным рубрикам MeSH) и "spironolactone" ("спиронолактон"; термин для поиска по предметным рубрикам MeSH), ограничив поиск определенной структурой исследований — "clinical trials" ("клинические испытания") и определенным годом ("1999"). В результате такого поиска вы находите только 4 ссылки; одна из них явно имеет отношение к интересующему вас вопросу [2].

В этой статье сообщается о проведении РКИ, в ходе которого 1663 больным с сердечной недостаточностью III или IV ФК по NYHA рандомизированно предписывался прием спиронолактона (по 25 мг 1 раз в сутки). Продолжительность наблюдения составила в среднем 2 года. Поскольку у вашего больного сердечная недостаточность соответствует II ФК, его вряд ли включили бы в это исследование. Однако вы по-прежнему предполагаете, что в отчете об РКИ может содержаться полезная информация, и решаете внимательно изучить его.

Как уже отмечалось, в данной книге рассматриваются способы оценки методологического качества исследований, посвященных лечению; однако термин "лечение" авторы понимают достаточно широко. Под ним подразумеваются вмешательства, направленные на снижение выраженности симптомов, предупреждение развития осложнений и наступления смерти при острых и хронических заболеваниях (включая профилактический прием спиронолактона при сердечной недостаточности), на снижение частоты смертности или развития осложнений в отдаленном периоде после перенесенного заболевания (например, применение Р-блокаторов после инфаркта миокарда — ИМ), на снижение смертности и частоты развития заболеваний при исходном наличии факторов риска (например, высокого артериального давления), на повышение качества медицинской помощи для улучшения клинических исходов (см.

главу 2В1 на компакт-диске). Кроме того, под такими вмешательствами подразумевается применение диагностических тестов, способствующее снижению смертности и частоты развития тяжелых осложнений (например, выполнение эзофагогастродуоденоскопии при желудочно-кишечном кровотечении), а также сочетание диагностических тестов и последующего лечения в рамках скрининговых программ (например, выявление скрытой крови в кале; см. главу 2F на компакт-диске). В каждой из этих ситуаций не исключено, что вред от вмешательства может перевесить пользу. Поэтому прежде чем действовать, необходимо сопоставить все преимущества и недостатки вмешательства, а также убедиться в том, что общая стоимость лечения не будет чрезмерной.

ДОСТОВЕРНЫ ЛИ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ?

При оценке материалов исследований, выявленных в медицинской литературе, можно применить трехэтапный подход (см. главу 3 *Лечение и риск: введение*). Сначала следует определить достоверность результатов исследования, затем провести их оценку и лишь потом решить, можно ли использовать эти результаты при лечении конкретного больного (табл. 1).

Таблица 1. Весь комплекс вопросов, на которые следует ответить при оценке результатов испытания лечебных вмешательств

Достоверны ли результаты исследования?

Был ли прогноз в основной и контрольной группах одинаковым перед началом исследования?

- Проводилась ли рандомизация?
- Использовалось ли сокрытие порядка рандомизации?
- Учитывалось ли при анализе данных первоначальное распределение больных по группам?
- Были ли группы сходны по всем известным факторам, влияющим на прогноз?

Оставался ли прогноз в основной и контрольной группах одинаковым на протяжении исследования?

- Знали ли больные о том, к какой группе они были отнесены в процессе рандомизации?
- Знали ли врачи о том, к какой группе тот или иной больной был отнесен в процессе рандомизации?
- Знали ли исследователи, оценивающие клинические исходы, о том, к какой группе тот или иной больной был отнесен в процессе рандомизации?
- Все ли больные наблюдались до конца исследования?

Продолжение табл. 1 на с. 76

Каковы результаты исследования?

- Каков размер эффекта изучаемого вмешательства⁹
- Насколько точным был расчет размера эффекта⁰
- Как оценивать результаты исследования, если в отчете не приведены значения доверительного интервала?

Можно ли применить результаты исследования при лечении конкретного больного?

- Совпадают ли характеристики участников исследования и конкретного больного?
 - Оценивались ли все клинически значимые исходы⁹
 - Перевешивают ли предполагаемые преимущества лечения все возможные побочные эффекты и затраты, связанные с его проведением?
-

Достоверность результатов исследования зависит от того, насколько его структура и особенности проведения позволяли выявить преимущества или недостатки изучаемого вмешательства. При оценке методологического качества исследования следует ответить на вопросы, объединенные в два блока; первый из них позволяет определить, был ли прогноз в основной и контрольной группах одинаковым перед началом исследования, а второй позволяет убедиться, оставался ли он таковым на протяжении исследования. Каждый из этих блоков состоит в свою очередь из 4 вопросов.

ПРОВОДИЛАСЬ ЛИ РАНДОМИЗАЦИЯ?

В качестве примера рассмотрим следующий клинический вопрос: "Может ли лечение в условиях стационара продлить жизнь тяжело больных?". По имеющимся данным, смертность среди таких больных в условиях стационара выше, чем в амбулаторных условиях. Однако наивным был бы вывод о том, что лечение в условиях стационара повышает смертность; интуитивно все понимают, что состояние госпитализированных больных в целом тяжелее, чем состояние больных, получающих лечение в амбулаторных условиях. Если не учитывать это различие, высока вероятность возникновения систематической ошибки и общей недооценки преимуществ госпитализации. В идеале для проведения качественного сравнения необходимо, чтобы больным с одинаково тяжелым состоянием рандомизированно предписывалось лечение либо в условиях стационара, либо в амбулаторных условиях; вряд ли в каком-либо медицинском учреждении комитет по этике разрешит проведение подобного исследования.

Приведем еще один пример. В 70—80-х годах XX века хирурги часто прибегали к наложению анастомоза между поверхностной височной артерией (ветвью наружной сонной артерии) и средней мозговой артерией (ветвью внутренней сонной артерии). Считалось, что это вмешательство предотвращает развитие инсульта при стенозе сонных артерий с клиническими проявлениями в тех случаях, когда невозможно выполнить каротидную эндартерэктомию. Сравнение клинических исходов в нерандомизированных когортах больных (по тем или иным причинам одним больным выполняли эту операцию, а другим нет) подтверждало это мнение. Результаты подобных исследований указывали на большую эффективность хирургического вмешательства по сравнению с его отсутствием. Однако к немалому удивлению врачей при проведении крупного многоцентрового РКИ, в ходе которого больных распределяли по группам хирургического или медикаментозного лечения рандомизированно (т.е. с использованием метода, принципиально не отличающегося от подбрасывания монеты), единственным подтвержденным эффектом хирургического вмешательства оказалось повышение частоты развития осложнений в раннем послеоперационном периоде [3].

Можно привести другие примеры неожиданных результатов, которые были получены в ходе РКИ и противоречат результатам ранних исследований более низкого методологического качества. Показано, например, что инъекции кортикостероидов не уменьшают выраженность боли, связанной с поражением межпозвоночных суставов [4]; что плазмаферез неэффективен при полимиозите [5] и что применение при сердечной недостаточности некоторых "многообещающих" препаратов способствует повышению смертности [6—10]. Такие сюрпризы нередки при сравнении результатов действительно рандомизированных испытаний и исследований, в которых распределение по группам зависит от решения врача или самого больного (см. главу 2В1 на компакт-диске).

В исследованиях последнего типа (*наблюдательные исследования*) зачастую не учитывается тот факт, что на клинические исходы (такие, как развитие осложнений и смерть больного), помимо изучаемого вмешательства, влияют и другие факторы; это способствует возникновению систематической ошибки. В исследованиях, посвященных оценке эффективности лечебного вмешательства, изучается его влияние на развитие таких событий, как инсульт, ИМ или смерть — их называют *оцениваемыми клиническими исходами*. Возраст больного, тяжесть основного заболевания, наличие сопутствующих заболеваний, а также ряд других обстоятельств (*называемых прогностическими факторами или факторами, влияющими на прогноз*) в значительной степени определяют частоту

развития оцениваемых клинических исходов. При разном распределении прогностических факторов (известных или неизвестных) в основной и контрольной группах сравнительная оценка клинических исходов будет ошибочной, а выявленный эффект лечения окажется завышенным или заниженным. Поскольку наличие известных прогностических факторов часто влияет на принятие решений врачом или больным об использовании того или иного вмешательства, достоверность результатов наблюдений исследований, как правило, невелика. В таких исследованиях размер благоприятного эффекта вмешательства обычно завышен по сравнению с РКИ [11—14], хотя может наблюдаться и его систематическая недооценка [15]. Теоретически даже в наблюдательных исследованиях можно учитывать известные прогностические факторы (при отборе больных или при последующем статистическом анализе) и добиваться определенного соответствия характеристик больных в группах сравнения (см. главы 5 *Оценка риска* и 1 раздела II *Лечение и риск: причины недостоверности результатов — систематические и случайные ошибки*). Однако проведение рандомизации существенно повышает вероятность того, что основная и контрольная группы будут сходны по всем известным и неизвестным факторам, способным повлиять на развитие определенного клинического исхода.

Следует заметить, что даже при использовании рандомизации не всегда удается добиться однородности групп сравнения. Возможны ошибочные действия, которые могут нивелировать преимущества рандомизации: например, не применяется сокрытие порядка рандомизации (исследователи, оценивающие соответствие больного критериям включения, знают, в какую группу он будет распределен) или при окончательном анализе данных не учитывается первоначальное отнесение больных с той или иной группе. Кроме того, исследователи просто могут столкнуться с элементарным невезением.

ИСПОЛЬЗОВАЛОСЬ ЛИ СОКРЫТИЕ ПОРЯДКА РАНДОМИЗАЦИИ?

Несколько лет назад группа австралийских исследователей провела РКИ, в котором сравнивалась эффективность выполнения аппендэктомии открытым и лапароскопическим способом [16]. В течение дня у исследователей никаких проблем не возникало. Ночью ситуация была [ной, поскольку во время проведения лапароскопического вмешательства в операционной обязательно должен присутствовать старший врач, при выполнении аппендэктомии открытым способом его присутствие [е требуется; из-за меньшего числа операционных, действующих в ночное время, выполнение более длительного лапароскопического вмеша-

тельства становится неудобным для персонала. Не желая беспокоить старших врачей, особенно некоторых из них, ординаторы иногда прибегали к "оригинальному" решению этой проблемы. При поступлении больного, отвечающего критериям включения в испытание, они сначала узнавали, кто из старших врачей дежурит и сколько больных готовится к экстренной операции. Если характер старшего врача и очередь в операционную внушали ординаторам опасения, они просвечивали изготовленные из недостаточно плотной бумаги конверты с указанием об отнесении участника в ту или иную группу и затем вскрывали тот, в котором содержалась инструкция выполнить аппендэктомию открытым способом. Утром при поступлении первого больного, отвечающего критериям включения, вскрывали отложенный ночью конверт; именно таким образом некоторые больные попадали в группу лапароскопического вмешательства [личное сообщение D. Wall, 2000 г.]. Если состояние больных, доставляемых в ночное время, было более тяжелым, чем у поступавших днем, подобное поведение ординаторов могло привести к возникновению систематической ошибки при оценке клинических исходов и недостоверным результатам испытания.

Из этого примера видно, что в отсутствие сокрытия порядка рандомизации лица, принимающие решение о соответствии больных критериям включения, могут систематически включать более тяжелых больных в группу вмешательства, а менее тяжелых больных — в контрольную группу. Это сведет на нет преимущества рандомизации; в результате влияния систематической ошибки полученные данные будут недостоверными [17, 18]. В хорошо организованном исследовании порядок рандомизации должен быть скрыт. Для этого могут применяться такие способы, как одинаковое внешнее оформление самих сравниваемых препаратов и их упаковки; применение "рандомизации по телефону", при которой лица, ответственные за включение больных в исследование, связываются с методическим центром, чтобы узнать, к какой группе следует отнести нового участника; использование запечатанных конвертов с соответствующими указаниями (на взгляд авторов, наименее надежный подход).

УЧИТЫВАЛОСЬ ЛИ ПРИ АНАЛИЗЕ ДАННЫХ ПЕРВОНАЧАЛЬНОЕ РАСПРЕДЕЛЕНИЕ БОЛЬНЫХ ПО ГРУППАМ?

Исследователи могут также уменьшить преимущества рандомизации, исключая из окончательного анализа данные о больных, которые не получали предписанное вмешательство. На первый взгляд это может показаться правильным, однако подобная тактика приведет к возник-

новению систематической ошибки и снизит достоверность результатов исследования.

Часто причины, по которым больные не принимают назначенные препараты, так или иначе связаны с прогнозом. В ряде РКИ больные, не выполнявшие предписанную схему лечения, чувствовали себя хуже, чем больные, строго следовавшие рекомендациям (это соотношение не изменялось даже при учете всех известных прогностических факторов и применения плацебо) [19—24]. Исключение из анализа данных о больных, не выполняющих предписаний врача, приводит к утрате информации о тех, у кого клинический исход мог быть лучше, и нарушает однородность групп, которая была достигнута в результате рандомизации.

Сходная ситуация наблюдается и при испытании хирургических методов лечения. Некоторым больным, отнесенным при рандомизации к группе хирургического вмешательства, операцию не выполняют из-за крайне тяжелого состояния или развития изучаемого клинического исхода (например, инсульта или ИМ) до перевода в операционную. Если при окончательном анализе данные о таких больных объединяют с данными об участниках из контрольной группы, даже бесполезная операция может показаться эффективной. Однако эта мнимая эффективность будет обусловлена не улучшением клинических исходов в группе хирургического вмешательства, а систематическим исключением данных о больных с наихудшим прогнозом.

Принцип рассмотрения данных обо всех больных с учетом их отнесения к той или иной группе при рандомизации (а не того вмешательства, которое применялось в действительности) лежит в основе *анализа данных в зависимости от назначенного вмешательства (анализ, проводимый исходя из допущения, что все больные получили предписанное вмешательство)*. Подобная тактика позволяет воспользоваться преимуществами рандомизации, поскольку распределение факторов, влияющих на прогноз (как известных, так и неизвестных), будет примерно одинаковым в двух группах, а наблюдаемый эффект будет связан только с назначенным лечением.

Таким образом, читая отчет об РКИ, следует обратить внимание на то, проводился ли анализ данных о больных с учетом их отнесения к той или иной группе при рандомизации (см. главу 2В1 на компакт-диске).

Были ли группы сходны по всем известным факторам, влияющим на прогноз?

Задача рандомизации заключается в формировании групп, в которых прогноз в отношении изучаемого исхода будет одинаковым. Ино-

гда в силу ряда причин эта цель не может быть достигнута при рандомизации. Чем меньше размер выборки, тем более вероятно, что прогноз в сравниваемых группах будет различаться.

Представим себе РКИ нового лекарственного препарата, в котором участвовали больные с сердечной недостаточностью III и IV ФК по NYHA. При сердечной недостаточности IV ФК прогноз значительно хуже, чем при сердечной недостаточности III ФК. В испытание были включены всего 8 участников. Не исключено, что все 4 больных с сердечной недостаточностью III ФК были распределены в основную группу, а 4 больных с сердечной недостаточностью IV ФК — в контрольную группу. Такое распределение может существенно повлиять на результаты исследования и привести к переоценке эффективности изучаемого препарата. Если бы в РКИ участвовали 800 больных, вряд ли все 400 больных с сердечной недостаточностью III ФК при рандомизации были отнесены к группе лечения. Чем больше размер выборки, тем более вероятно, что после рандомизации прогноз в сравниваемых группах будет одинаковым.

Проверить успешность рандомизации можно путем оценки распределения в основной и контрольной группах всех исходных прогностических характеристик. Хотя мы не можем знать, имеются ли между группами различия в распределении неизвестных прогностических факторов, мы будем уверены в сходном влиянии всех известных прогностических факторов.

Статистическая значимость различий в распределении известных прогностических факторов между группами менее важна (поскольку в РКИ любые возникающие различия случайны, что делает бесполезным часто применяемый расчет величины p), чем абсолютное значение этих различий. Если они велики, достоверность результатов исследования сомнительна. Чем сильнее связь между прогностическими факторами и изучаемым клиническим исходом, чем более выражены исходные различия между группами, тем слабее будут любые выводы о влиянии лечения на клинический исход (иными словами, результаты окажутся менее достоверными).

Если сравниваемые группы неоднородны в начале исследования, это совсем не означает, что результаты обязательно будут недостоверными. Существуют статистические методы, позволяющие провести коррекцию результатов с учетом исходных различий. Поэтому необходимо убедиться, что исследователи оценивали сходство групп по исходным характеристикам; если группы были неоднородными, надо убедиться в использовании специальных методов коррекции результатов. В том случае, ко-

гда до и после такого анализа выводы останутся **одинаковыми, можно** быть уверенным в их обоснованности.

ЗНАЛИ ЛИ БОЛЬНЫЕ О ТОМ, К КАКОЙ ГРУППЕ ОНИ БЫЛИ ОТНЕСЕНЫ В ПРОЦЕССЕ РАНДОМИЗАЦИИ?

Если больные считают лечение эффективным, их самочувствие и состояние может быть лучше, чем у больных, не верящих в эффективность того же самого вмешательства (даже когда это вмешательство не оказывает биологического действия). Хотя мы не знаем точно, насколько выражен и постоянен *эффект плацебо* [25], его наличие может ввести в заблуждение исследователей относительно биологического действия лекарственной терапии. Даже в отсутствие истинного эффекта плацебо результаты анкетирования или функциональных проб могут оказаться различными у больных, которые уверены или не уверены в получении настоящего лечебного вмешательства.

Самый лучший способ избежать этих проблем заключается в сокрытии от больных информации о принадлежности к той или иной группе. Например, в контрольной группе больные могут получать таблетки или капсулы с нейтральными компонентами, по цвету, вкусу и консистенции не отличающиеся от таблеток или капсул с активным лекарственным препаратом. Соблюдение этих условий позволяет уравновесить выраженность эффекта плацебо в основной и контрольной группах.

ЗНАЛИ ЛИ ВРАЧИ О ТОМ, К КАКОЙ ГРУППЕ ТОТ ИЛИ ИНОЙ БОЛЬНОЙ БЫЛ ОТНЕСЕН В ПРОЦЕССЕ РАНДОМИЗАЦИИ?

Если рандомизация прошла успешно, то исходный прогноз в основной и контрольной группах одинаков. Однако сама процедура рандомизации еще не гарантирует сходства прогностических характеристик в группах сравнения на протяжении всего исследования. Различия в применении каких-либо медицинских вмешательств, не относящихся к изучаемому методу лечения, могут способствовать возникновению систематической ошибки. Вернемся к рассматривавшемуся ранее примеру с оценкой эффективности спиронолактона. Если в ходе РКИ больные основной группы дополнительно получали ингибиторы ангиотензин-превращающего фермента или р-блокаторы, которые в контрольной группе не применялись, то конечная оценка эффективности изучаемого вмешательства окажется завышенной, поскольку прием указанных препаратов увеличивает продолжительность жизни при сердечной недостаточности.

Большее доверие вызывают те отчеты об исследованиях, в которых сообщается о более или менее равномерном применении в группах сравнения всех сопутствующих вмешательств, способных повлиять на клинический исход. Отсутствие такой информации допустимо лишь в том случае, когда врачи не знали, к какой группе тот или иной больной был отнесен в процессе рандомизации. Применение слепого метода позволяет исключить возможность различного подхода к дополнительному лечению в основной и контрольной группах (как сознательного, так и неосознанного).

ЗНАЛИ ЛИ ИССЛЕДОВАТЕЛИ, ОЦЕНИВАЮЩИЕ КЛИНИЧЕСКИЕ ИСХОДЫ, О ТОМ, К КАКОЙ ГРУППЕ ТОТ ИЛИ ИНОЙ БОЛЬНОЙ БЫЛ ОТНЕСЕН В ПРОЦЕССЕ РАНДОМИЗАЦИИ?

При более тщательном наблюдении чаще сообщается о возникновении изучаемых исходов. Кроме того, если лица, регистрирующие физиологические показатели или клинические исходы (например, при проведении функциональных проб, оценке общего состояния или качества жизни), знают о том, к какой группе тот или иной больной был отнесен в процессе рандомизации, они могут либо по-разному интерпретировать сомнительные результаты, либо по-разному выполнять обследование [26]. В отношении исследователей, оценивающих клинические исходы, почти всегда можно применить слепой метод, даже если от самих больных и лечащих врачей скрыть отнесение к той или иной группе нереально (например, при сравнении эффективности различных хирургических вмешательств или работы разных медицинских служб). Можно принять дополнительные меры предосторожности путем создания экспертной комиссии, члены которой, не зная результатов рандомизации, будут оценивать клинические данные, решать вопросы диагностики (например, действительно ли в ходе исследования у больного развился инсульт или ИМ), определять причину смерти (например, смерть от рака или ишемической болезни сердца). Чем большее значение имеет мнение исследователя при оценке изучаемого клинического исхода, тем важнее становится применение слепого метода (это менее важно, если в качестве изучаемого клинического исхода рассматривается смерть от любых причин).

ВСЕ ЛИ БОЛЬНЫЕ НАБЛЮДАЛИСЬ ДО КОНЦА ИССЛЕДОВАНИЯ?

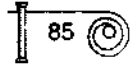
В идеале по завершении исследования должно быть известно о наличии или отсутствии изучаемого клинического исхода у каждого участника. Однако иногда больные выбывают из исследования до его завер-

шения. Чем больше доля таких больных, тем менее достоверными будут результаты, поскольку прогноз у больных, выбывших из исследования, часто отличается от прогноза у тех, кто наблюдался до его завершения. Больные могут выбывать из исследования в связи со смертью или развитием побочных эффектов либо, напротив, в связи с существенным улучшением состояния (после чего они просто не приходят на повторное обследование). Выше уже указывалась причина, по которой необходимо проводить анализ данных в зависимости от назначенного вмешательства. В этом случае ситуация аналогична: у выбывших больных вероятность возникновения изучаемого неблагоприятного исхода может быть меньшей или (как правило) большей, чем среди наблюдавшихся до конца РКИ.

Когда выбывание больных существенно влияет на достоверность результатов? В этом случае относительные показатели неприменимы (в литературе иногда рекомендуют использовать точку разделения, равную 20%). В качестве примера представим два РКИ, каждое из которых включает по 1000 больных в основной и контрольной группах. Из каждой группы выбыли по 30 больных (3%; табл. 2). В испытании А в основной группе умерли вдвое меньше больных, чем в контрольной (200 и 400 больных соответственно); снижение относительного риска (СОР) составило 50%. В какой мере выбывание больных из исследования может повлиять на вывод о том, что лечение вдвое снижает смертность? Если предполо-

Таблица 2. В каких случаях выбывание больных из исследования существенно снижает достоверность результатов?

	Испытание А		Испытание В	
	основная группа	контрольная группа	основная группа	контрольная группа
Число больных	1000	1000	1000	1000
Число больных, выбывших из исследования (%)	30 (3)	30 (3)	30 (3)	30 (3)
Число умерших больных (%)	200 (20)	400 (40)	30 (3)	60 (6)
Значение СОР без учета числа выбывших больных	0,2/0,4 = 0,50		0,03/0,06 = 0,50	
Значение СОР с учетом наихудшего предположения о смерти всех выбывших больных	0,17/0,4 = 0,43		0,00/0,06 = 0	



жить самое худшее, т.е. что в группе лечения все выбывшие из исследования больные умерли, число случаев смерти составит 230 (23%). Если же в контрольной группе среди выбывших больных не было ни одного случая смерти, значение COP будет иным; этот показатель изменится с 200/400, или 50%, до 170/400, или 43%. Оба предположения мало повлияли на оценку эффекта лечения, поэтому первоначальные выводы можно считать достоверными.

Теперь рассмотрим результаты испытания В. В нем COP смерти также составляет 50%, однако число умерших больных меньше — 30 в группе лечения и 60 в контрольной группе. Если предположить, что в группе лечения все выбывшие больные умерли, результаты существенно изменятся: число случаев смерти в группе лечения возрастает с 30 до 60 и становится равным числу случаев смерти в контрольной группе. Если в каждой группе умерли по 60 больных, эффективность лечения снижается до 0. Таким образом, выбывание из исследования 3% больных влияет на достоверность оценки эффективности лечения ($\text{COP}=50\%$ без учета клинических исходов у выбывших больных и $\text{COP}=0\%$ при наихудшем предположении о том, что все выбывшие больные относились к группе лечения и все они умерли).

Конечно, наихудшее предположение вряд ли отражает истинное положение вещей. Однако, если при его рассмотрении существенно изменяются конечные результаты, необходимо оценить вероятность наличия выраженных различий в частоте возникновения изучаемого клинического исхода у тех больных в основной и контрольной группах, которые не наблюдались до конца исследования. Вероятность таких различий уменьшается, но не становится равной нулю, если в статье представлена информация, подтверждающая одинаковое исходное распределение важных прогностических факторов (таких как возраст и тяжесть состояния) у больных, выбывших из обеих групп.

Таким образом, наблюдение не за всеми участниками исследования может снизить достоверность полученных результатов. Если наихудшее предположение о клинических исходах у выбывших больных не влияет на окончательные результаты, доля таких больных не имеет решающего значения; в противном случае достоверность результатов снижается. Степень снижения достоверности зависит от вероятности развития неблагоприятного клинического исхода у больных, выбывших из основной группы, и от вероятности развития благоприятного клинического исхода у больных, выбывших из контрольной группы.

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

Вернемся к нашему первоначальному примеру. Были ли в ходе испытания спиронолактона сформированы группы, однородные по влиянию прогностических факторов? Исследователи указывают, что испытание было рандомизированным, однако не сообщают о сокрытии порядка рандомизации. Из 822 больных основной группы в связи с отсутствием ожидаемого эффекта, развития побочных реакций или по иным причинам лечение прекратили 214 больных, а в контрольной группе это соотношение составило 841 и 200 больных. По-видимому, данные о всех выбывших больных исследователи включили в окончательный анализ, который проводился исходя из допущения, что все больные получали назначенное лечение. Сообщается о сходстве групп по таким показателям, как возраст, пол, раса, величина артериально-го давления, частота сердечных сокращений, величина фракции выброса, причина сердечной недостаточности и применение дополнительных лекарственных препаратов. Единственным показателем, неравномерно представленным в обеих группах, была выраженность сердечной недостаточности: признаки сердечной недостаточности IV ФК по NYHA имелись у 31 % больных в контрольной группе и у 27% больных в основной группе. Это могло привести к возникновению систематической ошибки и завышению эффективности лечения. Однако влияние такой неоднородности групп скорее всего было незначительным; кроме того, исследователи сообщили о том, что при проведении окончательного анализа данных учитывались исходные различия по известным прогностическим показателям.

Как и во многих отчетах о РКИ, авторы называют свое исследование "двойное слепое". К сожалению, ни эксперты по клинической эпидемиологии, ни читатели не знают точно, что понимается под этим термином. Поэтому следует просто уточнить, по отношению к кому применялся слепой метод. В статье сообщается, что о результатах рандомизации не знали ни больные, ни медицинский персонал, ни лица, оценивавшие клинические исходы; это подтверждают и редакторы справочника Best Evidence, контактировавшие с авторами [28].

В отчете не приведены точные сведения о доле выбывших больных, но представленные данные указывают на то, что наблюдение было полным. Следует заметить, что при выборе смерти в качестве изучаемого клинического исхода такая ситуация маловероятна, хотя она вполне возможна при изучении других клинических исходов.

Конечная оценка достоверности результатов никогда не подразумевает простого ответа ("да" или "нет"). Разумнее представить достоверность конкретного испытания в виде точки на непрерывной шкале; на одном конце этой шкалы можно расположить РКИ высокого методологического качества с надежными и точными результатами, а на другом — исследования низкого методологического качества с результатами, почти наверняка искаженными за счет наличия систематических ошибок. Увы, определить расположение конкретного исследования на этой шкале невозможно без некоторого субъективизма. В данном случае, несмотря на неясность по поводу доли больных, выбывших из исследования, его методологическое качество и достоверность результатов можно в целом расценить как достаточно высокие. Вероятность возникновения систематических ошибок при оценке эффективности спиронолактона у больных с сердечной недостаточностью была минимальной; полученные данные могут быть использованы при решении вопроса о применении спиронолактона в конкретном случае.

КАКОВЫ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ?

КАКОВ РАЗМЕР ЭФФЕКТА ИЗУЧАЕМОГО ВМЕШАТЕЛЬСТВА?

В ходе РКИ наиболее тщательно оценивается частота развития неблагоприятных клинических исходов. Примеры *дихотомических исходов* (изучаемое явление возникает или не возникает, при его оценке ответ будет "да" или "нет") включают в себя рецидивирование рака, развитие ИМ и смерть. В отчете об исследовании сообщается о доле больных, у которых развился данный исход. Например, в ходе исследования смертность среди больных контрольной группы составила 20%, а среди больных основной группы — 15%. Как можно выразить эти результаты?

Во-первых, их можно представить в виде абсолютного различия (*снижение абсолютного риска*, *САР*) между долей больных, умерших в контрольной (x) и основной (y) группах: $x - y = 0,20 - 0,15 = 0,05$. Во-вторых, результаты можно представить в виде отношения между риском развития этого неблагоприятного клинического исхода в основной и контрольной группе (*относительный риск*, *ОР*): $x/y = 0,15/0,20 = 0,75$

Чаше всего данные о дихотомических исходах представляют в виде такого дополнительного показателя, как *снижение относительного риска* (*СОР*). Этот показатель выражается в процентах: $(1 - y/x) \cdot 100 = (1 - 0,75) \cdot 100 = 25\%$. Величина СОР, равная 25%, означает, что примене-

ние нового лечебного вмешательства снижает риск смерти на 25% по сравнению с контрольным вмешательством; чем больше величина СОР, тем более эффективно лечение. Исследователи могут подсчитать ОР за определенный период времени, как при *анализе выживаемости*; этот показатель называется *отношением риска* (см. главу 2В2 на компакт-диске). В тех случаях, когда не указано, имеется ли в виду САР или СОР (например, говорится, что "лечение препаратом X на 30% эффективнее снижало риск смерти" или "эффективность вакцинации составила 92%"), почти всегда речь идет о СОР. В рекламе лекарственных препаратов, даже если это не оговорено особо, почти всегда имеется в виду ОР. По сравнению с другими способами представления результатов показатель СОР всегда формирует субъективное впечатление о более выраженном положительном эффекте лечения; подробнее об этом написано в главе 2В2 на компакт-диске.

НАСКОЛЬКО ТОЧНЫМ БЫЛ РАСЧЕТ РАЗМЕРА ЭФФЕКТА?

В действительности истинное снижение риска никогда нельзя установить. Самое лучшее, что мы имеем, это результат тщательно контролируемых испытаний; лучший расчет основан на истинном эффекте вмешательства, наблюдавшемся в ходе такого испытания. Полученное значение называется *точечным*. Оно рассчитано на основании наблюдений за ограниченной выборкой, и по нему судят о значении изучаемого показателя в популяции. Такая *точечная оценка* почти точна, но истинное значение в большей или меньшей степени отличается от нее. Исследователи часто сообщают о диапазоне, в котором, вероятнее всего, находится истинное значение для популяции. Этот диапазон называют *доверительным интервалом* (ДИ). [29].

Обычно используется произвольно выбранный 95% ДИ (см. главу 2В2 на компакт-диске), который указывает на то, что в 95% случаев истинная величина СОР будет находиться в приводимом диапазоне значений. Она редко приближается к границам этого диапазона, а за его пределы выходит только в 5% случаев. Такие характеристики ДИ тесно связаны с общепринятым уровнем статистической значимости, т.е. соответствуют значению $p < 0,05$ (см. главу 2В2 на компакт-диске). Приведем несколько примеров использования ДИ.

Пример 1. В начале испытания в основную и контрольную группу рандомизированно были распределены по 100 больных, затем в контрольной группе зарегистрированы 20 случаев смерти, а в основной группе —

15. При расчете СОР исследователями его точечное значение составило 25% ($x = 20 / 100$, или 0,20; $y = 15 / 100$, или 0,15; $1 - y/x = [1-0,75] \cdot 100=25\%$). Однако можно предположить, что истинное значение СОР окажется меньше или больше 25%, поскольку расчет СОР основывался на небольшом абсолютном различии, составлявшем всего 5 случаев смерти. Фактически не исключено, что лечение может оказаться неэффективным (СОР=0%) или даже принести вред (отрицательное значение СОР). Действительно, полученное точечное значение СОР находится в интервале от -38% (в основной группе риск смерти на 38% выше, чем в контрольной) до $+59\%$ (в основной группе риск смерти почти на 60% ниже, чем в контрольной). Другими словами, результаты этого испытания (СОР=25% при 95% ДИ от -38 до +59%) не помогут решить вопрос, стоит ли применять новый препарат.

Пример 2. Допустим, что в каждой группе было не по 100, а по 1000 больных при тех же значениях смертности: 200 случаев смерти в контрольной группе ($x = 200/1000 = 0,20$) и 150 — в основной группе ($y = 150 / 1000 = 0,15$). В таком испытании точечное значение СОР также составит 25% ($x = 200 / 1000$, или 0,20; $y = 150 / 1000$, или 0,15; $1 - y/x = [1-0,75] \cdot 100=25\%$). Однако при рассмотрении результатов более крупного исследования можно предположить, что истинное СОР гораздо ближе к 25%; скорее всего, это будет именно так. В данном случае значения всего диапазона 95% ДИ будут выше 0 (СОР=25% при 95% ДИ от 9 до 41%).

Таким образом, чем больше размер выборки, тем выше частота возникновения изучаемого клинического исхода и уверенность в том, что истинное значение СОР (или любого другого показателя эффективности) будет близким к расчетному. Во втором из приведенных примеров при оценке СОР нижняя граница ДИ составляла 9%, а верхняя — 41%. Точечное значение (в данном случае 25%) наиболее близко к истинному значению СОР. Чем больше любое другое значение СОР будет отклоняться от этой точечной оценки, тем ниже будет вероятность того, что оно соответствует истинному значению СОР. В момент пересечения верхней или нижней границы ДИ эта вероятность минимальна.

На рисунке представлены диапазоны ДИ при определении СОР в двух вышеуказанных примерах (СОР=0 означает отсутствие лечебного эффекта). В обоих испытаниях точечное значение СОР одинаково (25%), однако во втором случае ДИ гораздо более узок.

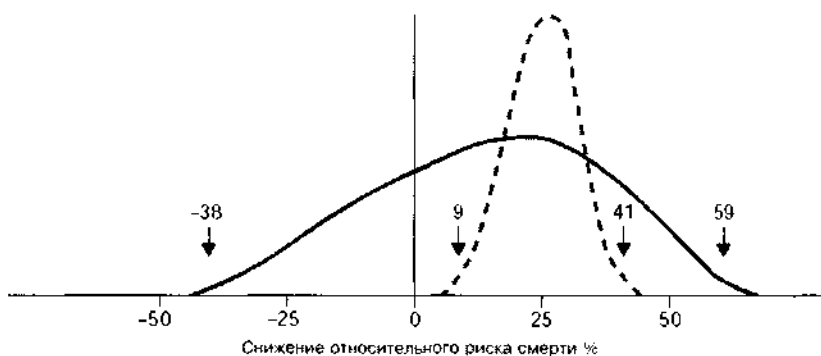


Рисунок. Доверительные интервалы при расчете снижения относительного риска.

Результаты 2 испытаний с одинаковым точечным значением изучаемого показателя ($SOR=25\%$), но с разным размером выборки и, следовательно, с разным диапазоном ДИ. Сплошной линией обозначены значения ДИ в первом испытании (по 100 больных в основной и контрольной группе; 15 и 20 случаев смерти соотв.), пунктирной — значения ДИ во втором испытании (по 1000 больных в основной и контрольной группах; 150 и 200 случаев смерти соотв.).

Таким образом, чем больше размер выборки, тем уже диапазон ДИ. Какой размер выборки можно считать достаточно большим [30] (см. главу 2В2 на компакт-диске)? При чтении материалов *исследования с положительным результатом* (т.е. исследования, в котором сделан вывод об эффективности изучаемого лечения) следует обращать внимание на нижнюю границу ДИ. Во втором из приведенных примеров нижняя граница ДИ составляла 9%. Если вы считаете это значение SOR (минимальное в рамках данного исследования) клинически значимым (т.е. позволяющим рекомендовать такое лечение), то размер выборки будет достаточным. Если же это минимальное значение SOR (9%) расценивается как клинически незначимое, результаты исследования нельзя считать бесспорными, хотя они статистически значимы (диапазон значений ДИ не включает 0). Следует помнить о том, что вероятность выхода истинного значения SOR за нижнюю границу ДИ составляет лишь 2,5%. Кроме того, использование иного ДИ (например, 90% ДИ) может быть столь же или даже более приемлемым.

Анализ значений ДИ позволяет также интерпретировать результаты исследований с отрицательным результатом (т.е. исследований, в которых был сделан вывод об отсутствии преимуществ изучаемого вмешательства по сравнению с контрольным). В отчетах о таких исследованиях нужно обратить внимание на верхнюю границу ДИ. Если при соответствующем ей значении СОР результаты клинически значимы, то размер выборки скорее всего был недостаточным для исключения клинически значимого эффекта лечения. В первом из рассмотренных выше примеров (испытание, включавшее по 100 больных в основной и контрольной группах) нельзя исключить возможность вреда от лечения (повышение ОР смерти на 38%); такому результату должно соответствовать значение $p > 0,05$. Поскольку убедительно продемонстрировать преимущества лечения не удалось, результаты исследования можно было бы расценить как отрицательные (см. рис. 1). Однако верхняя граница ДИ составляла 59%. Истинное значение СОР может находиться неподалеку от верхней границы ДИ; в этом случае польза от лечения будет существенной. Поэтому можно сделать вывод о том, что исследователи не доказали наличие преимуществ изучаемого вмешательства по сравнению с плацебо, но не доказали и их отсутствие; полученные данные не позволяют исключить возможность существенного положительного эффекта лечения. В заключение хочется напомнить, что выбор 95% ДИ произволен. Вполне приемлемо использовать более узкий 90% ДИ.

КАК ОЦЕНИВАТЬ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ, ЕСЛИ В ОТЧЕТЕ НЕ ПРИВЕДЕНЫ ЗНАЧЕНИЯ ДОВЕРИТЕЛЬНОГО ИНТЕРВАЛА?

Что можно сделать, если в публикации приводится значение СОР, но не сообщается о диапазоне его ДИ? Проще всего проанализировать значения p . Если $p = 0,05$, нижняя граница 95% ДИ значений СОР располагается на нулевой отметке ($ОР = 1$), а потому невозможно исключить вероятность того, что лечение неэффективно. Если $p < 0,05$, нижняя граница 95% ДИ всегда выше 0.

Можно также рассчитать ДИ самостоятельно [31] или обратиться за помощью к кому-либо еще (например, к статистику). После расчета ДИ вы будете знать, насколько высоким и низким может быть значение СОР (т.е. оценить точность определения размера эффекта), и сможете интерпретировать результаты по вышеописанной методике.

Не во всех РКИ оцениваются дихотомические клинические исходы, порой это и не требуется. Так, авторы публикации об испытании спиронолактона могли сообщить только об изменениях физической активности или качества жизни. В исследовании, посвященном оценке эффек-

тивности тренировки дыхательных мышц при хронических обструктивных заболеваниях легких, в качестве одного из основных клинических исходов было выбрано расстояние, которое больные могли пройти за 6 мин по коридору [32]. В основной группе это расстояние увеличилось с 406 до 416 м (т.е. на 10 м), а в контрольной группе — с 409 до 429 м (т.е. на 20 м). Точечное значение различия между этими показателями оказалось отрицательным (—10 м) и свидетельствовало не в пользу изучаемого вмешательства.

Однако в данном случае тоже полезно проанализировать диапазон 95% ДИ. В статье указано, что нижняя граница 95% ДИ составила -26 м (т.е. больные из контрольной группы могли пройти на 26 м больше), а верхняя граница — +5 м. Даже при наилучшем стечении обстоятельств увеличение исходного расстояния, которое больной может пройти за 6 мин, в основной группе составило бы всего 5 м; такое улучшение клинического исхода несущественно для больных, поэтому полученный результат позволяет полностью исключить наличие клинически значимых преимуществ тренировки дыхательных мышц (по крайней мере такой, какая применялась в данном исследовании).

Оценив размер эффекта и точность его расчета, врач может рассмотреть последнюю группу вопросов о том, как применить результаты исследования в собственной практике.

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

За время наблюдения (в среднем 2 года) в группе лечения умерли 284 (35%) из 822 больных, в контрольной группе — 386 (46%) из 841 больного. Исследователи провели анализ выживаемости, при котором учитывается не только число больных, умерших к концу испытания, но также и сроки наступления смерти у этих больных. На основании общих данных об уровне смертности можно рассчитать значения ОР (76%) и СОР (24%). Поскольку в контрольной группе больные умирали не только чаще, но и раньше больных из основной группы, при анализе выживаемости значение СОР (или отношение риска) составило 35% при 95% ДИ от 18 до 48%. Кривые выживаемости стали расходиться через 6 мес после начала исследования, и далее эта тенденция нарастала.

Число больных, госпитализированных по поводу заболеваний сердца, в основной и контрольной группах составило 336 и 260 соответственно

(OR=0,70 при 95% ДИ от 0,59 до 0,82). В группе плацебо примерно у 33% больных было отмечено улучшение состояния (уменьшение выраженности сердечной недостаточности по классификации NYHA на 1 ФК и более), у 18% больных состояние не изменилось, а у 48% больных оно ухудшилось. В группе спиронолактона эти показатели составили 41, 21 и 38% соответственно. Маловероятно, что различия в изменении состояния, оцениваемого по классификации NYHA, были случайными ($p < 0,001$).

У 9 (1%) больных в контрольной группе и у 61 (10%) больного в основной группе отмечалось развитие гинекомастии или возникновение боли в молочной железе ($p < 0,001$). Выраженная гиперкалиемия возникла у 10 (1%) и 14 (2%) больных соответственно; различие в частоте развития данного побочного эффекта было статистически незначимым.

МОЖНО ЛИ ПРИМЕНИТЬ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ ПРИ ЛЕЧЕНИИ КОНКРЕТНОГО БОЛЬНОГО?

СОВПАДАЮТ ЛИ ХАРАКТЕРИСТИКИ УЧАСТНИКОВ ИССЛЕДОВАНИЯ
И КОНКРЕТНОГО БОЛЬНОГО?

Часто характеристики больного, которого вам предстоит лечить, и участников исследования различаются. Ваш больной может быть старше, его состояние — более или менее тяжелым по сравнению с состоянием участников исследования, он может иметь сопутствующие заболевания, из-за которых его не включили бы в исследование. Если характеристики конкретного больного соответствуют критериям включения и не отвечают ни одному из критериев исключения, результаты исследования можно уверенно применять на практике.

Однако и тут есть ограничения. Эффективность лечебного вмешательства не всегда одинакова у отдельных больных. Обычно у одних больных наблюдается существенное улучшение, а у других оно вообще отсутствует. В ходе РКИ, как правило, оценивается средний эффект лечения. Руководствуясь средними результатами, врач с большей или меньшей долей вероятности подвергает некоторых больных риску неэффективного лечения с сопутствующими побочными эффектами и затратами.

Эту проблему можно решить, проводя РКИ с участием отдельно взятого больного (РКИ "n=1"). Чтобы свести к минимуму вероятность возникновения систематических ошибок, в этих исследованиях использу-

ются те же методы, что и в обычных РКИ с большими выборками [33]. В ходе такого РКИ каждый больной в течение определенного времени получает изучаемое лечебное вмешательство, а затем на протяжении такого же срока — плацебо или какое-либо другое вмешательство (см. главу 2В1 на компакт-диске). При этом ни врачи, ни больные не знают о характере вмешательства, последовательность чередования таких периодов назначается рандомизированно; в течение каждого из них больные количественно оценивают выраженность симптомов. Исследование продолжается до тех пор, пока больной и врач не сделают вывод о наличии (или отсутствии) преимуществ изучаемого вмешательства. При выполнении всех необходимых условий подобные РКИ характеризуются целым рядом преимуществ: а) они достаточно легко осуществимы [34, 35]; б) позволяют получить точные данные об эффективности вмешательства у отдельных больных; в) на основе результатов этих РКИ в ряде случаев может быть пересмотрена стандартная схема лечения [36].

С другой стороны, РКИ с участием отдельно взятого больного непригодны для оценки эффективности лечения кратковременных острых заболеваний или эффективности вмешательств, приводящих к полному излечению, а также для оценки отдаленных или редких клинических исходов. Кроме того, их проведение возможно только в том случае, если больные и врачи в этом заинтересованы и располагают достаточным временем. В большинстве случаев врач вынужден применять при лечении конкретного больного результаты традиционных РКИ.

Как быть, если больной, которого вам предстоит лечить, не отвечает критериям включения в исследование? Не исключено, что результаты будут применимы, даже если больной на 2 года старше участников этого исследования, если его состояние тяжелее, если ранее он лечился аналогичными препаратами или страдает сопутствующими заболеваниями. Вместо выяснения того, в какой степени отдельные характеристики больного не соответствуют критериям включения и исключения, лучше сразу задаться вопросом, есть ли какие-то важные обстоятельства, мешающие применить результаты исследования. Обычно такие непреодолимые обстоятельства отсутствуют, и чаще всего результаты исследования можно с достаточной долей уверенности применить в конкретном случае (см. главу 2В3 на компакт-диске).

С этим связана еще одна проблема: до какой степени обобщаемы результаты исследования в том случае, если собираются применять не тот препарат, который испытывали в РКИ, а его аналог или препарат того же фармакологического класса? Ответ на этот вопрос зависит от того, оказывают ли все препараты этого фармакологического класса одина-

ковый лечебный эффект; на него редко можно дать однозначный ответ (см. главу 2В3 на компакт-диске).

И наконец, дилемма возникает в том случае, когда характеристики больного соответствуют характеристикам участников в одной из подгрупп. В отчетах об РКИ (особенно если лечение оказалось неэффективным для среднестатистического больного) часто сообщается о многочисленных подгруппах больных разного возраста, с разными стадиями основного заболевания, различными сопутствующими заболеваниями и т.п. Очень часто проведение анализа в подгруппах не планируется заранее, а потому данные просто загружают в компьютер в поисках существующих и несуществующих связей между явлениями. В результате таких манипуляций исследователи порой делают выводы о том, что размер эффекта лечения различен в разных подгруппах. Например, может быть показано, что клинические исходы в подгруппах пожилых больных или больных с более тяжелым состоянием лучше (или, напротив, хуже), чем в других подгруппах.

К результатам анализа в подгруппах необходимо относиться с осторожностью [37]. Вероятность того, что лечение в данной подгруппе более или менее эффективно по сравнению с другими подгруппами, высока только при наличии выраженных различий, которые вряд ли могли быть случайными. Но даже в таком случае выводы могут быть ошибочными, если исследователи не сформулировали эту гипотезу до начала исследования, если исходных гипотез было много, если в других исследованиях не удастся подтвердить полученные результаты.

ОЦЕНИВАЛИСЬ ЛИ ВСЕ КЛИНИЧЕСКИ ЗНАЧИМЫЕ ИСХОДЫ?

Лечение показано лишь в тех случаях, когда оно действительно приводит к существенным положительным результатам. Данные о том, что применение бронходилататора несколько повышает объем форсированного выдоха при хронических обструктивных заболеваниях легких, что применение вазодилататора повышает величину сердечного выброса при сердечной недостаточности или что прием гиполипидемического препарата улучшает липидный профиль крови, не всегда можно считать достаточной причиной для назначения указанных лекарственных средств. Необходимы данные об улучшении клинических исходов, важных для больных, например об уменьшении одышки при обычной физической активности, о снижении частоты госпитализаций по поводу сердечной недостаточности или о снижении риска развития ИМ. Достижение определенных значений объема форсированного выдоха за 1-ю секунду, сердечного выброса или уровня липидных фракций можно ис-

тользовать в качестве *косвенного критерия оценки* или *промежуточного клинического исхода* (см. главу 2В3 на компакт-диске). Такие критерии зценки обычно используют вместо истинных клинических исходов, имеющих значение для больного; обычно это связано с тем, что для под-гверждения положительного влияния лечебного вмешательства на ис-гинные клинические исходы пришлось бы включать в исследование го-заздо больше участников и наблюдать за ними на протяжении более дли-гельного времени.

В качестве драматического примера, свидетельствующего об опас-ности использования косвенных критериев оценки, можно вспомнить, 4ем закончились испытания противоаритмических препаратов после перенесенного ИМ. На основе данных о том, что кратковременное при-менение этих препаратов снижает частоту возникновения несвоевремен-ной деполяризации желудочков (косвенный клинический исход), было выдвинуто предположение о возможном снижении частоты развития угрожающих жизни аритмий в отдаленном периоде. Группа исследова-гелей организовала проведение нескольких рандомизированных испы-таний энкаинида, флекаинида и морицизина (ранее эти три препарата были признаны эффективными средствами, предотвращающими раз-витие косвенного клинического исхода — несвоевременной деполяри-зации желудочков), чтобы определить, снижает ли применение этих пре-паратов смертность среди больных, перенесших ИМ с нарушениями сердечного ритма (в отсутствие симптоматики или при незначительных клинических проявлениях). Исследователи были вынуждены прекратить эти испытания гораздо раньше намеченного срока, поскольку смерт-ность среди принимающих антиаритмические препараты оказалась су-щественно выше, чем в группе плацебо [38, 39]. Доверяя результатам исследований с использованием косвенных критериев оценки, врачи могли бы и дальше назначать такое лечение в ущерб своим больным.

Даже если авторы сообщают о положительном влиянии лечения на один клинически значимый исход, необходимо выяснить, нет ли отри-цательного влияния на другие исходы. Например, применение гиполи-пидемических препаратов снижает смертность от сердечно-сосудистых заболеваний, но может повышать общую смертность [40]. Проведение химиотерапии при раке может удлинить жизнь, но снижает ее качество (см. главу 2В2 на компакт-диске). В исследованиях по оценке эффек-тивности хирургических вмешательств часто сообщается об увеличении продолжительности жизни больных, благополучно перенесших опера-цию (например, в группе хирургического вмешательства уровень 3-лет-ней выживаемости выше, чем в контрольной группе); в этой группе по-

вышен риск смерти во время операции или в раннем послеоперационном периоде. Поэтому при изучении отчета об испытании хирургических методов лечения необходимо обратить внимание не только на информацию об отдаленных результатах, но и на данные о периоперационной смертности, которая обычно выше в основной группе. К важным клиническим исходам, о которых "забывают" сообщать авторы статей, относится частота развития токсических реакций и побочных эффектов. \ Другим важным критерием оценки, на который в течение длительного времени не обращали должного внимания, следует считать эффективность затрат. В ходе немногих РКИ изучаются разные виды затрат: прямые (средства, которые расходуются на лекарственные препараты или перечисляются медицинскому персоналу в качестве заработной платы) либо непрямые (например, вызванные снижением работоспособности больных). Более жесткое ограничение ресурсов здравоохранения диктует необходимость проведения тщательного экономического анализа, особенно в случаях, когда вмешательство требует больших затрат (см. главу 2F на компакт-диске).

ПЕРЕВЕШИВАЮТ ЛИ ПРЕДПОЛАГАЕМЫЕ ПРЕИМУЩЕСТВА ЛЕЧЕНИЯ
ВСЕ ВОЗМОЖНЫЕ ПОБОЧНЫЕ ЭФФЕКТЫ И ЗАТРАТЫ, СВЯЗАННЫЕ
С ЕГО ПРОВЕДЕНИЕМ?

Если результаты исследования можно применить при лечении данного больного, а оцениваемые исходы клинически значимы, следует сопоставить предполагаемые преимущества лечения со всеми возможными побочными эффектами и затратами. Выявленное в ходе исследования снижение относительного риска смерти на 25% впечатляет, однако в конкретной ситуации эффект лечения может оказаться минимальным. Здесь стоит обратиться к такому показателю, как *число больных, которых необходимо лечить* (ЧБНЛ; это число больных, которых надо лечить определенным методом в течение определенного времени, чтобы предотвратить развитие одного неблагоприятного исхода или достичь одного благоприятного исхода [41]).

Эффекты лечения заключаются не только в снижении риска развития неблагоприятных исходов, но и в повышении риска развития побочных эффектов. В одном крупном РКИ было показано, что при остром ИМ применение тканевого активатора плазминогена по сравнению со стрептокиназой снижает относительный риск смерти от ИМ примерно на 12% [42]. В табл. 3 сопоставляются данные о двух больных с острым ИМ, сопровождающимся подъемом сегмента ST.

Таблица 3. Данные, имеющие значение для принятия решения о применении тканевого активатора плазминогена или стрептокиназы у двух больных с ИМ

	Риск смерти в течение 1 года после ИМ при лечении стрептокиназой, %	Риск смерти в течение 1 года после ИМ при лечении тканевым активатором плазминогена, %	САР, % (абсолютное значение)	ЧБНЛ
40-летний больной с мелкоочаговым ИМ	2	1,76	0,24(0,0024)	417
70-летний больной с обширным ИМ и сердечной недостаточностью	40	35,2%	4,8 (0,048)	21

В первом случае у 40-летнего мужчины на ЭКГ выявлены изменения, указывающие на наличие ИМ в области нижней стенки левого желудочка. У больного нет признаков сердечной недостаточности, отмечается нормальный синусовый ритм с частотой сердечных сокращений 90 уд/мин. В данном случае на фоне применения стрептокиназы абсолютный риск смерти в течение 1 года после ИМ составляет 2%. Применение тканевого активатора плазминогена снизит этот показатель на 12% (с 2 до 1,76%), т.е. САР составит 0,24% (или 0,0024). Величина, обратная САР ($1 / \text{САР}$), соответствует значению ЧБНЛ для интересующего нас клинического исхода (здесь речь идет о предотвращении 1 случая смерти у больных с невысоким риском ее наступления). Чтобы спасти одну жизнь, придется лечить примерно 417 таких больных ($1 / 0,0024 = 416,7$). Учитывая некоторое повышение риска возникновения внутрисердечного кровоизлияния, связанное с применением тканевого активатора плазминогена, а также более высокую стоимость этого препарата, многие врачи могут предпочесть в данном случае применение стрептокиназы.

Второй больной в возрасте 70 лет поступил в стационар с признаками обширного ИМ в области передней стенки левого желудочка на ЭКГ и отеком легких. У такого больного риск смерти в течение последующего года высок и составляет примерно 40%. При этом снижение риска смерти на 12% за счет применения тканевого активатора плазминогена будет соответствовать значению САР, равному 4,8% (или 0,048). Чтобы спасти одну жизнь, придется лечить только 21 такого больного ($1 / 0,048 = 20,8$). В

данной ситуации многие врачи предпочтут применение тканевого активатора плазминогена.

Таким образом, при принятии решения о начале применения определенного лечебного вмешательства крайне важно оценить риск развития неблагоприятного исхода в отсутствие этого вмешательства. Чем выше вероятность развития неблагоприятного исхода в отсутствие лечения, тем выше вероятность того, что лечение окажется эффективным, и для предупреждения одного неблагоприятного исхода необходимо будет лечить меньшее число больных; это положение справедливо при любом значении СОР (см. главу 2В2 на компакт-диске). Зная о величине ЧБНЛ, врач лучше сможет сопоставлять преимущества и недостатки разных вмешательств (см. главу 2В3 на компакт-диске). В главе 2В3 на компакт-диске представлены значения этого показателя при использовании некоторых распространенных методов лечения в определенных группах риска.

При сопоставлении преимуществ и недостатков лечения требуется точная оценка побочных эффектов. РКИ с относительно небольшой выборкой непригодны для выявления редких, но опасных для жизни побочных эффектов лечения. В ходе РКИ можно оценить частоту развития распространенных побочных эффектов, однако в отчетах сообщения о таких клинических исходах зачастую отсутствуют. Чтобы оценить вероятность развития различных побочных эффектов лечения, врачам нередко приходится обращаться к другим источникам информации (как правило, характеризующимся более низким методологическим качеством).

В конечном итоге выбор метода лечения зависит от предпочтений и жизненных ценностей самого больного. В последнее время появляется все больше интересных статей об отношении больных к своей болезни (см. главу 2В3 на компакт-диске). Пока еще нет единого мнения о том, как наилучшим образом сообщить больному необходимую для принятия решения информацию, и как учитывать их предпочтения. Однако в настоящее время проводится исследование, посвященное этому вопросу (см. главу 11 *От доказательств к практике*).

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ; ПРИНЯТИЕ ПРЕДВАРИТЕЛЬНОГО РЕШЕНИЯ

В испытании спиринолактона оценивались многие клинические исходы, включая смерть, частоту госпитализаций и повседневную актив-

ность. Кроме того, в отчете сообщается о существенном повышении частоты развития гинекомастии и боли в молочной железе у мужчин, принимавших этот препарат, а также о небольшом и статистически незначимом повышении частоты возникновения эпизодов выраженной гиперкалиемии.

Для группы лечения в целом при снижении риска смерти с 46 до 35% (САР=11%, или 0,11) расчет величины ЧБНЛ будет следующим: ЧБНЛ = $1/0,11 = 9$. Однако не у всех больных с сердечной недостаточностью прогноз одинаков. При сердечной недостаточности IV ФК смертность в течение 2 лет достигает 60%, а при сердечной недостаточности III ФК — 40%. Можно предположить, что при сердечной недостаточности II ФК этот показатель составит примерно 20%.

В табл. 4 приведены некоторые возможные преимущества и недостатки применения спиронолактона при сердечной недостаточности. Если использовать точечную оценку СОР, значение ЧБНЛ при сердечной недостаточности IV ФК и высоком исходном риске смерти будет равно 6, а при сердечной недостаточности III и II ФК и более низком исходном риске смерти — 9 и 17 соответственно. В этой таблице также представлено наименьшее значение СОР (18%), которое не противоречит результатам РКИ (нижняя граница 95% ДИ) и соответствует наибольшему значению ЧБНЛ (для больных с сердечной недостаточностью IV, III и II ФК это значение ЧБНЛ составляет 9, 14 и 27 соответственно). Риск возникновения гинекомастии и боли в молочной железе скорее всего не зависит от выраженности сердечной недостаточности. Поэтому при общей частоте развития этих побочных эффектов, равной 9%,

Таблица 4. Сравнительная характеристика преимуществ и недостатков применения спиронолактона при разной выраженности сердечной недостаточности

ФК по классификации NYHA	Риск смерти в течение 2 лет без лечения, %	Возможное значение САР, % (ЧБНЛ) на фоне лечения при СОР=30%	Наименьшее возможное значение САР, % (ЧБНЛ) при СОР=18%*	Риск развития гинекомастии или боли в молочной железе у мужчин, % (ИПВ)
IV	60	18(6)	11 (9)	9(11)
III	40	12(9)	7(14)	9(11)
II	60	6(17)	3,6 (27)	9(11)

Примечание * — соответствует нижней границе 95% ДИ.

значение *индекса потенциального вреда* (ИПВ — показатель, близкий по смыслу к ЧБНЛ; это число больных, которые должны получить изучаемое вмешательство, чтобы у одного дополнительного больного развился неблагоприятный исход) составит 11 для всех трех подгрупп. Расчет ИПВ проводят так же, как расчет ЧБНЛ, однако ИПВ — величина, обратная повышению абсолютного риска (т.е. $1 / 0,09$). В завершение можно сказать, что спиронолактон относится к недорогим препаратам, и затраты на лечение в течение года составят приблизительно 25 долларов США.

Можно предположить, что большинство больных, зная об этих преимуществах и недостатках, выберут лечение спиронолактоном. При возникновении боли в молочной железе или гинекомастии мужчины всегда смогут прекратить прием препарата. Однако в исследовании не участвовали больные с сердечной недостаточностью II ФК. Можно ли считать, что у этих больных значение СОР будет таким же, как у больных с более тяжелой сердечной недостаточностью?

По многим причинам на этот вопрос можно ответить положительно. Биологические механизмы сердечной недостаточности на всем протяжении ее развития не изменяются. Авторы исследования сообщают, что применение спиронолактона предотвращает прогрессирование сердечной недостаточности (за счет уменьшения задержки натрия и замедления развития кардиосклероза) и наступление внезапной смерти (за счет предупреждения потери калия и увеличения захвата норадреналина кардиомиоцитами). Кроме того, применение спиронолактона может предотвращать развитие кардиосклероза, подавляя влияние альдостерона на синтез коллагена. Вряд ли эти механизмы действия препарата (если они действительно объясняют полученные результаты) не будут важны для больных с сердечной недостаточностью II ФК по NYHA.

Можно привести и другие доводы. Так, при оценке риска смерти значение СОР на фоне применения спиронолактона оказалось одинаковым в разных подгруппах больных с сердечной недостаточностью; этот показатель не зависел от наличия ишемии миокарда и дополнительного применения р-блокаторов. И наконец, известно, что при использовании других препаратов, снижающих смертность при сердечной недостаточности (ингибиторы ангиотензинпревращающего фермента и β -блокаторы), значение СОР одинаково при сердечной недостаточности II, III или IV ФК по NYHA [43].

Ваш больной всегда активно участвует в процессе принятия решений и выборе медицинского вмешательства. Поэтому во время обсуждения вопроса о недостатках и преимуществах приема спиронолактона

большому прежде всего необходимо сообщить о риске развития боли в молочной железе или гинекомастии, составляющем 9% в течение 2 лет, и о снижении риска смерти с 20 до 14%. Следует также разъяснить некоторую условность этих расчетов, которая связана как с диапазоном ДИ значений СОР в группе спиронолактона (риск смерти может снизиться с 20 до 16%, а не до 14%), так и с исключением из исследования больных с сердечной недостаточностью II ФК по НУНА. Убедившись в том, что больной понял все правильно, вы можете быть уверены, что его решение будет осознанным, отражающим индивидуальные предпочтения и жизненные ценности.

Литература

- 1 Vasotec tablets enalapril maleate In Physician's Desk Reference 52nd ed Montvale, NJ Medical Economics, 1998 1771—1774
- 2 Pitt B, Zannad F, Remme WJ, et al The effect of spironolactone on morbidity and mortality in patients with severe heart failure N Engl J Med 1999;341 709—717
- 3 Haynes R B, Mukherjee J, Sackett D L, et al Functional status changes following medical or surgical treatment for cerebral ischemia results in the EC/IC Bypass Study JAMA 1987 257 2043—2046
- 4 Carette S, Martoux S, Truchon R, et al A controlled trial of corticosteroid injections into facet joints for chronic low back pain N Engl J Med 1991;325 1002-1007
- 5 Miller FW, Leitman S F, Cromm M E, et al Controlled trial of plasma exchange and leukapheresis in polymyositis and dermatomyositis N Engl J Med 1992;326 1380—1384
- 6 The Xamoterol in Severe Heart Failure Group Xamoterol in severe heart failure Lancet 1990;336 1—6
- 7 Packer M, Carver J R, Rodeheffer R J, et al, for the PROMISE Study Research Group Effects of oral milrinone on mortality in severe chronic heart failure N Engl J Med 1991;325 1468—1475
- 8 Packer M, Rouleau J L, Svedberg K, Pitt B, Fisher L, and the Profile investigators Effect of flosequinan on survival in chronic heart failure preliminary results of the PROFILE study [abstract] Circulation 1993;88(suppl 1) 1-301
- 9 Hampton J R, van Veldhuisen D J, Kleber F X, et al, for the Second Prospective Randomized Study of Ibopamine on Mortality and Efficacy (PRIME II) Investigators Randomised study of effect of Ibopamine on survival in patients with advanced severe heart failure Lancet 1997;349 971—977
- 10 Cahiff R M, Adams K F, McKenna W J, et al A randomized controlled trial of epoprostenol therapy for severe congestive heart failure the Flolan International Randomized Survival Trial (FIRST) Am Heart J 1997;134 44-54
- 11 Sacks H S, Chalmers TC, Smith H Jr Sensitivity and specificity of clinical trials randomized \ historical controls Arch Intern Med 1983;143 753—755
- 12 Chalmers TC, Celano P, Sacks H S, Smith H Jr Bias in treatment assignment in controlled clinical trials N Engl J Med 1983;309 1358—1361
- 13 Colditz G A, Miller J N, Mosteller F How study design affects outcomes in comparisons of therapy, 1 medical Stat Med 1989 8 441-454
- 14 Emerson J D, Burdick E, Hoaghn DC, et al An empirical study of the possible relation of treatment differences to quality scores in controlled randomized clinical trials Controlled Clin Trials 1990;11 339—352

- 15 *Kunz R , Oxman A D* The unpredictability paradox review of empirical comparisons of randomised and non-randomised clinical trials **BMJ** 1998,317 1185—1190
- 16 *Hansen J B , Smithers B M , Schache D , Wall D R , Miller B J , Menves B L* Laparoscopic versus open appendectomy prospective randomized trial *World J Surg* 1996,20 17—20
- 17 *Schuli K F , Chalmers I , Hayes R J , Altman D G* Empirical evidence of bias dimensions of methodological quality associated with estimates of treatment effects in controlled trials *JAMA* 1995,273 408—412
- 18 *Moher D , Jones A , Cook D J , et al* Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses¹ *Lancet* 1998,352 609—613
- 19 *Coronary Drug Project Research Group* Influence of adherence treatment and response of cholesterol on mortality in the Coronary Drug Project *N Engl J Med* 1980,303 1038—1041
- 20 *Asher W L , Harper H W* Effect of human chorionic gonadotropin on weight loss, hunger, and feeling of well-being *Am J Clin Nutr* 1973,26 211-218
- 21 *Hogarty G E , Goldberg S C* Drug and sociotherapy in the aftercare of schizophrenic patients *Arch Gen Psychiatry* 1973,28 54-64
- 22 *Fuller R , Roth H , Long S* Compliance with disulfiram treatment of alcoholism *J Chronic Dis* 1983,36 161-170
- 23 *Pizzo PA , Robtchaud K J , Edwards B K , Schumaker C , Kramer B S , Johnson A* Oral antibiotic prophylaxis in patients with cancer a double-blind randomized placebo-controlled trial *J Pediatr* 1983,102 125—133
- 24 *Horwitz R / , Viscoli C M , Berkman L , et al* Treatment adherence and risk of death after myocardial infarction *Lancet* 1990,336 542—545
- 25 *Kaptchuk T J* Powerful placebo the dark side of the randomised controlled trial *Lancet* 1998,351 1722—1725
- 26 *Guyatt G H , Pugsley S O , Sullivan M J , et al* Effect of encouragement on walking test performance *Thorax* 1984,39 818-822
- 27 *Devereaux P J , Manns B J , Ghah WA , et al* In the dark physician interpretation of blinding terminology in randomized controlled trials *JAMA* In press
- 28 *Henderson M , Mulrow C D* Commentary on "The effect of spironolactone on morbidity and mortality *m* patients with severe heart failure " *ACP J Club* 2000 Jan-Feb, 132(1) 2
- 29 *Altman D G , Gore S M , Gardner M J , Pocock S J* Statistical guidelines for contributors to medical journals In *Gardner M J , Altman D G , eds Statistics With Confidence Confidence Intervals and Statistical Guidelines* London British Medical Journal, 1989 83-100
- 30 *Detsky A S , Sackett D L* When was a "negative" trial big enough? How many patients you needed depends on what you found *Arch Intern Med* 1985,145 709—7)5
- 31 *Sackett D L , Haynes R B , Guyatt G H , Tugwell P* *Clinical Epidemiology A Basic Science for Clinical Medicine* 2nd ed Boston Little Brown & Co Inc. 1991 218
- 32 *Guyatt G H , Keller J , Singer J , Halcrow S , Newhouse M* Controlled trial of respiratory muscle training in chronic airflow limitation *Thorax* 1992,47 598-602
- 33 *Guyatt G H , Sackett D L , Taylor D W , et al* Determining optimal therapy randomized trials in individual patients *N Engl J Med* 1986,314 889-892
- 34 *Guyatt G H , Keller J L , Jaeschke R , et al* The n-of-1 randomized control trial clinical usefulness Our three year experience *Ann Intern Med* 1990,112 293—299
- 35 *Larson E B , Ellsworth A J , Oas J* Randomized clinical trials in single patients during a 2-year period *JAMA* 1993,270 2708-2712
- 36 *Mahon J , Laupaas A , Donner A , Wood T* Randomised study of n of 1 trials versus standard practice **BMJ** 1996,312 1069-1074
- 37 *Oxman A D , Guyatt G H* A consumer's guide to subgroup analysis *Ann Intern Med* 1992,116 78—84

38. *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—788.
39. *The Cardiac Arrhythmia Suppression Trial II Investigators.* Effect of the antiarrhythmic agent moricizine on survival after myocardial infarction. *N Engl J Med* 1992;327:227—233.
40. *Muldoon M.F., Manuck S.B., Matthews K.A.* Lowering cholesterol concentrations and mortality: a quantitative review of primary prevention trials. *BMJ* 1990;301:309—314.
41. *Laupacis A., Sackett D.L., Roberts R.S.* An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988;318:1728—1733.
42. *The GUSTO Investigators.* An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction. *N Engl J Med* 1993;329:673-682.
43. *Garg R., Yusuf S.* Overview of randomized trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure. *JAMA* 1995;273:1450—1456.

5**ОЦЕНКА РИСКА**

М. Левайн, Д. Хэслам, Ш. Уолтер,
Р. Камминг, Х. Ли, Т. Хэйнс, А. Холбрук,
В. Мойер, Г. Гайятт

В подготовке данной главы принимали также участие члены Рабочей группы по разработке и внедрению принципов доказательной медицины П. Проновост и Ш. Строе

В ЭТОМ РАЗДЕЛЕ

Поиск доказательной информации

Достоверны ли результаты исследования?

Продemonстрировали ли исследователи сходство обеих групп по всем известным факторам риска, влияющим на изучаемый клинический исход? Были ли учтены имеющиеся различия при анализе данных?

Использовались ли в основной и контрольной группах одинаковые методы для оценки влияния изучаемого фактора риска?

Использовались ли в основной и контрольной группах одинаковые методы для оценки клинических исходов?

Какова доля больных, наблюдавшихся до конца исследования?

Каковы результаты исследования?

Насколько сильна связь между воздействием фактора риска и клиническим исходом?

Насколько точно рассчитан риск развития неблагоприятного клинического исхода?

Как можно применить результаты исследования при лечении конкретного больного?

Совпадают ли характеристики больного и участников исследования?

Было ли наблюдение за участниками исследования достаточно продолжительным?

Каков риск развития неблагоприятного клинического исхода?

Следует ли принимать меры, направленные на прекращение неблагоприятного воздействия?

Принятие клинического решения

ВОЗМОЖНАЯ КЛИНИЧЕСКАЯ СИТУАЦИЯ

Может ли прием селективных ингибиторов обратного захвата серотонина приводить к возникновению желудочно-кишечного кровотечения?

Представьте, что вы врач общей практики, решающий вопрос о выборе оптимального препарата для лечения депрессии. К вам на прием пришел 55-летний мужчина, который до недавнего времени был вполне жизнерадостным и уравновешенным; однако в течение последних двух месяцев он испытывает подавленность и тоску с нарушениями концентрации внимания и ранним пробуждением по утрам, но без суицидальных мыслей. Вы наблюдали больного в течение 20 лет, хорошо его знаете и расцениваете это состояние как эпизод большой депрессии, полагая, что в данном случае показаны антидепрессанты.

В последнее время в качестве препарата первого ряда для лечения депрессии вы назначали селективный ингибитор обратного захвата серотонина (СИОЗС) пароксетин. Однако в недавно опубликованных обзорах вы прочитали, что препараты этой группы не более эффективны, чем трициклические антидепрессанты (ТЦА) [1—3], и при терапии препаратами обеих групп доля больных, которые прекращают лечение, одинакова [1—4]. Поэтому в некоторых случаях вы опять стали начинать терапию с нортриптилина. Ваши больные обычно соглашались на то или иное лечение лишь после обсуждения его возможных побочных эффектов; многие предпочитают принимать СИОЗС именно из-за большей безопасности препаратов этой группы.

Кроме того, на протяжении предшествующих 5 лет для снижения интенсивности болей, вызванных остеоартрозом тазобедренного сустава, данный больной принимает кетопрофен (препарат, относящийся к классу нестероидных противовоспалительных средств, — НПВС) в дозе 50 мг 3 раза в сутки. Вы вспоминаете об одном обзоре [6], в котором говорилось, что терапия СИОЗС может повышать риск развития кровотечений. Поскольку больной уже принимает НПВС, вы начинаете сомневаться в целесообразности сочетанного применения этих препаратов из-за возможного возникновения желудочно-кишечного кровотечения. К сожалению, в реферате из базы данных Evidence Based Mental Health [5], откуда вы обычно черпаете данные о побочных эффектах антидепрессантов, информация по этому вопросу отсутствует.

Вы находите текст обзора [6] среди своих документов и после беглого изучения понимаете, что он не поможет вам по трем причинам: четкие критерии включения и исключения не приведены, поиск данных не был систематическим и всеобъемлющим, методологическое качество представленных исследований не оценивалось (см. главу 10 *Обобщение доказательств*). Кроме того, в обзоре нет ссылок на исследования, в которых оценивалась бы взаимосвязь между приемом СИОЗСи риском развития желудочно-кишечного кровотечения.

Прежде чем давать какие-либо рекомендации, вы должны сами обратиться в этом вопросе. Поэтому вы сообщаете больному, что ему показан прием антидепрессантов, но из-за опасений по поводу возможного возникновения желудочно-кишечного кровотечения необходимо проанализировать все имеющиеся данные. Вы назначаете следующий прием через 2 дня и обязуетесь за это время выработать тактику лечения.

ПОИСК ДОКАЗАТЕЛЬНОЙ ИНФОРМАЦИИ

Вы формулируете интересующий вас клинический вопрос следующим образом:

"Повышен ли риск развития тяжелого кровотечения из верхних отделов желудочно-кишечного тракта (ЖКТ) у взрослых больных, страдающих депрессией и принимающих СИОЗС, по сравнению с больными, не принимающими антидепрессанты?"

Вечером того же дня вы начинаете поиск с использованием термина "serotonin reuptake inhibitor" ("ингибитор обратного захвата серотонина") в различных источниках доказательной медицинской информации: обновляемых базах данных Evidence Based Mental Health, Best Evidence⁴, Clinical Evidence, а также в Cochrane Library. В базе данных Evidence Based Mental Health вы находите 8 обзоров — в томах 1 (1998 г.) и 2 (1999 г.). В четырех из этих обзоров речь идет о побочных эффектах СИОЗС, но ни в одном не оценивается частота возникновения желудочно-кишечного кровотечения. В базе данных Best Evidence⁴ вы находите 17 статей, столь же бесполезных для решения вашего вопроса. В Clinical Evidence есть только обзор, посвященный лечению депрессивных расстройств у взрослых. В Cochrane Library помещены 4 полных систематических обзора и 2 реферата, но ни в одном из них не упоминается о возникновении желудочно-кишечного кровотечения на фоне приема СИОЗС.

Тогда вы решаете использовать поисковые системы MEDLINE: PubMed и PreMedline (www.ncbi.nlm.nih.gov/entrez/query.fcgi). Чтобы сделать поиск максимально эффективным, вы можете либо щелкнуть по кнопке "Clinical queries" ("Клинические вопросы") в разделе "PubMed Services", что позволяет воспользоваться прошедшими систематическую оценку подходами к поиску данных, либо прибегнуть к функции "Search hedges" ("Границы поиска"), что поможет выявить высококачественные исследования по оценке риска (см. главу 2 *Поиск доказательной информации*). Запрос может выглядеть следующим образом:

"serotonin reuptake inhibitor" ("ингибитор обратного захвата серотонина"; термин для поиска по предметным рубрикам) AND ("И"; логический оператор) "bleeding" ("кровотечение"; термин для поиска по предметным рубрикам).

При этом в разделе "Study category" ("Категория исследований") вы выбираете термин "Etiology" ("Этиология"), а в разделе "Emphasis" ("Особое выделение") — термин "Specificity" ("Специфичность"). При поиске в базе данных MEDLINE материалов, опубликованных в период с 1966 по 2000 г., вы находите ссылку на эпидемиологическое исследование, в котором оценивалась взаимосвязь между приемом СИОЗС и частотой возникновения кровотечений из верхних отделов ЖКТ [7]. Полученные данные свидетельствуют о трехкратном повышении риска возникновения кровотечения на фоне приема СИОЗС. Полагая, что в материалах этого исследования можно найти ответ на интересующий вас вопрос, вы копируете полный текст статьи с открытого сайта British Medical Journal (www.bmj.com) в формате pdf, позволяющем изучать документы в том виде, в котором они были напечатаны.

ДОСТОВЕРНЫ ЛИ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ?

Врачам часто приходится иметь дело с лицами, которые подвергаются потенциально опасным воздействиям (речь может идти о медицинских вмешательствах или о факторах окружающей среды), в связи с чем возникает немало важных вопросов. Повышается ли риск выкидыша у беременных, работающих на компьютере? Повышается ли риск развития рака предстательной железы после выполнения вазэктомии? Может ли применение на производстве программ по выявлению и лечению артериальной гипертонии повышать число прогулов? При поиске ответа на подобные вопросы, врач должен оценить достоверность имеющихся данных, выраженность корреляции между предполагаемым фактором

риска и неблагоприятным клиническим исходом, а также применимость результатов исследования в условиях собственной практики.

Как и при ответе на любой клинический вопрос, следует прежде всего искать систематические обзоры по интересующей теме, в которых за счет применения объективных методов обобщаются все имеющиеся данные (см. главу 10 *Обобщение доказательств*). Однако для правильной интерпретации результатов систематического обзора нужно знать правила оценки доказательности данных, получаемых в ходе обсервационных (нерандомизированных) исследований. Применение таких правил (как и при рассмотрении результатов РКИ) поможет определить, был ли прогноз одинаковым в основной и контрольной группах не только в начале исследования, но и на всем его протяжении (табл. 1).

Таблица 1. Рекомендуемый комплекс вопросов, на которые следует ответить при оценке результатов исследований, посвященных изучению риска

Достоверны ли результаты исследования?

Был ли прогноз одинаковым в основной и контрольной группах в начале исследования?

- Продемонстрировали ли исследователи сходство обеих групп по всем известным факторам риска, влияющим на изучаемый клинический исход? Были ли учтены имеющиеся различия при анализе данных⁹
- Использовались ли в основной и контрольной группах одинаковые методы для оценки влияния изучаемого фактора риска?

Был ли прогноз одинаковым в основной и контрольной группах на протяжении всего исследования?

- Использовались ли в основной и контрольной группах одинаковые методы для оценки клинических исходов⁹
- Какова доля больных, наблюдавшихся до конца исследования?

Каковы результаты исследования?

- Насколько сильна связь между воздействием фактора риска и клиническим исходом?
- Насколько точно рассчитан риск развития неблагоприятного клинического исхода⁹

Как можно применить результаты исследования у конкретного больного?

- Совпадают ли характеристики данного больного и участников исследования?
 - Было ли наблюдение за участниками исследования достаточно продолжительным?
 - Каков риск развития неблагоприятного клинического исхода?
 - Следует ли предпринимать меры, направленные на прекращение неблагоприятного воздействия?
-

ПРОДЕМОНСТРИРОВАЛИ ЛИ ИССЛЕДОВАТЕЛИ СХОДСТВО ОБЕИХ ГРУПП ПО ВСЕМ ИЗВЕСТНЫМ ФАКТОРАМ РИСКА, ВЛИЯЮЩИМ НА ИЗУЧАЕМЫЙ КЛИНИЧЕСКИЙ ИСХОД? БЫЛИ ЛИ УЧТЕНЫ ИМЕЮЩИЕСЯ РАЗЛИЧИЯ ПРИ АНАЛИЗЕ ДАННЫХ?

Результаты исследований, посвященных изучению факторов риска, могут быть ошибочными, если прогноз в основной группе (участники, подвергающиеся воздействию предполагаемого фактора риска) и в контрольной группе с самого начала был неодинаковым. Например, нас интересует, как госпитализация влияет на смертность. Если просто сравнить уровень смертности среди госпитализированных больных и среди лиц того же возраста и пола в общей популяции, можно было бы сделать вывод о том, что сама по себе госпитализация смертельно опасна; однако вряд ли кто-либо воспримет это серьезно. Понятно, что госпитализируют только больных людей, а у них риск смерти выше. При использовании упрощенного подхода может сложиться ложное (т.е. не основанное на причинно-следственных отношениях) впечатление о наличии связи между воздействием (в данном случае госпитализацией) и клиническим исходом (в данном случае смертью). В целом состояние лиц, обращающихся за медицинской помощью или принимающих лекарственные препараты, хуже, чем у тех, кто в этом не нуждается. Без учета этого обстоятельства повышается вероятность недостоверных выводов о причинно-следственных связях между применением медицинского вмешательства и развитием неблагоприятных эффектов.

Как убедиться в том, что в начале исследования риск развития того или иного неблагоприятного клинического исхода в сравниваемых группах был одинаковым? При оценке влияния факторов риска в ходе РКИ вероятность возникновения систематической ошибки гораздо ниже, чем в ходе исследований любого другого типа, поскольку рандомизация является наилучшим методом формирования групп, однородных по всем известным и неизвестным характеристикам (см. главу 4 *Исследования, посвященные эффективности лечения*). Хотя РКИ обычно проводят с целью оценки эффективности вмешательства, полученные данные позволяют судить и о его безопасности. Неожиданные результаты некоторых РКИ (например, в тех случаях, когда применение препарата, который считали более эффективным, приводит к повышению смертности) демонстрируют возможность использования исследований с такой структурой для оценки риска (см. главу 2В1 на компакт-диске).

Существуют две причины, по которым обычно не удается найти РКИ, помогающее ответить на вопрос, действительно ли предполагаемое вред-

Таблица 2. Принципы организации, достоинства и недостатки исследований с различной структурой

Структура исследования	Показатели, оцениваемые в начале исследования	Показатели, оцениваемые в конце исследования	Достоинства	Недостатки
Когортное исследование	Влияние изучаемого фактора	Клинические исходы	Допустимо при невозможности проведения рандомизации	Высока вероятность возникновения систематических ошибок, достоверность результатов ограничена
Исследование случай-контроль	Клинические исходы	Влияние изучаемого фактора	Не требует больших затрат времени и крупной выборки	Высока вероятность возникновения систематических ошибок, достоверность результатов ограничена
РКИ	Влияние изучаемого фактора	Неблагоприятные клинические исходы, побочные эффекты	Вероятность возникновения систематических ошибок низкая	Не всегда осуществимо; часто результаты неприменимы к различным популяциям больных

ное воздействие оказывает неблагоприятный эффект. Во-первых, неэтично проводить такую рандомизацию, в результате которой кто-то из участников будет подвергаться воздействию опасных для здоровья факторов (в том числе неэффективному лечению). Даже если поступиться этим принципом, информированные больные не согласятся на подобный эксперимент.

Во-вторых, нас часто интересуют редкие и тяжелые побочные эффекты, которые могут быть зарегистрированы только при длительном наблюдении, после лечения десятков тысяч больных. Например, даже в очень крупном РКИ [8] не удалось выявить взаимосвязь между приемом клопидогрела и развитием тромботической тромбоцитопенической пурпуры, однако такая связь была обнаружена в более позднем обсервационном исследовании [9]. В ходе РКИ, которые проводятся специально для оценки побочных эффектов, эти эффекты можно выявить даже

при низкой (1%) частоте их возникновения [10, 11]; в таких случаях целесообразно также выполнение мета-анализа [12]. Если частота развития неблагоприятного клинического исхода меньше 1 на 100 участников, следует проводить **РКИ** с очень большой выборкой и продолжительным наблюдением; это проблематично по организационным и финансовым соображениям.

Поскольку при поиске ответа на большинство вопросов о факторах риска врач не найдет подходящих **РКИ**, он должен знать иные методы, способные уравновесить прогноз в сравниваемых группах. Для этого следует ознакомиться со структурой обсервационных исследований (табл. 2).

Когортные исследования

При проведении когортного исследования формируют две группы (когорты), лишь в одной из которых участники подвержены изучаемому воздействию; в ходе последующего наблюдения в обеих когортах регистрируют возникновение изучаемого клинического исхода. Так, в одном подобном исследовании оценивалась частота развития неблагоприятных клинических исходов в перинатальном периоде у детей грудного возраста, отцы которых были заняты в печатном производстве и подвергались воздействию свинца и органических растворителей. В сравниваемые когорты включили всех мужчин, которые состояли в профессиональных союзах печатников в Осло [13]. Сортируя участников по роду профессиональной деятельности, исследователи определяли, кто подвергался влиянию вышеуказанного фактора риска. Оказалось, что в соответствующей когорте в 8 раз повысился риск преждевременных родов у жен участников, но не изменилась частота возникновения врожденных уродств у детей.

Результатам когортных исследований можно доверять, если неблагоприятные клинические исходы развиваются редко. Так, частота возникновения желудочно-кишечного кровотечения с клиническими проявлениями на фоне приема НПВС составляет примерно 1,5 на 1000 человеко-лет лечения; у лиц, не принимающих НПВС, этот показатель составляет 1,0 на 1000 человеко-лет [14]. Поскольку частота возникновения желудочно-кишечного кровотечения вне воздействия указанного фактора риска очень мала (0,1%), то в **РКИ**, способное выявить повышение риска на 50%, пришлось бы включить огромное число участников (около 75 000 в каждой группе); только в этом случае статистическая мощность (чувствительность) исследования будет достаточной для проверки гипотезы о том, что прием НПВС повышает частоту кровотечений [15]. Такое **РКИ** провести невозможно; а вот проведение когорт-

ного исследования (с использованием информации из крупных баз данных медицинских учреждений) вполне реально

Достоверность оценки влияния потенциально вредных факторов в обсервационном исследовании может быть снижена из-за различий в исходном риске развития изучаемого клинического исхода в сравниваемых группах. Например, при оценке влияния приема НПВС на частоту возникновения желудочно-кишечного кровотечения возраст больных может коррелировать как с приемом НПВС, так и с частотой развития кровотечения. Другими словами, поскольку больные, принимающие НПВС, будут старше, а риск кровотечения также повышен у пожилых больных, то при анализе результатов обязательно следует учитывать такой дополнительный фактор, как возраст (*вмешивающийся фактор*); в противном случае обоснованность вывода о взаимосвязи между повышенным риском возникновения кровотечения и приемом НПВС будет сомнительной.

Поскольку в группах участников, подвергающихся (добровольно или в процессе лечения) или не подвергающихся потенциально неблагоприятным воздействиям, возможно наличие дополнительных неучтенных факторов, никогда нельзя быть уверенным, что характеристики участников в группах сравнения сопоставимы. Более того, следует предполагать, что эти характеристики по многим причинам различаются. Врачи редко рекомендуют препараты, применение которых, по их мнению, может подвергнуть больных дополнительному риску, и стараются назначать лекарственные средства с менее выраженными побочными эффектами. Так, в одном из первых испытаний кетопрофена доля больных, лечившихся по поводу язвенной болезни в предшествующие 2 года, в основной группе составила 24,1%, а в контрольной группе — только 15,7% [16]. По-видимому, информационные материалы фирмы-производителя убедили врачей, что новый препарат реже вызывает желудочно-кишечное кровотечение, чем другие НПВС. Последующее сравнение кетопрофена с другими препаратами этого фармакологического класса может создать ложное впечатление о том, что частота кровотечений в основной группе повышается, поскольку именно кетопрофен будут назначать больным, у которых риск развития желудочно-кишечного кровотечения и так высок.

Применение бензодиазепинов у пожилых больных можно рассматривать в качестве еще одного примера того, как избирательное назначение препаратов с более редкими и менее выраженными побочными эффектами способно вызвать перераспределение риска. Этот феномен называют *систематической ошибкой, связанной с назначением препаратов с учетом прогноза* [17, 18]. При анализе данных, полученных в период с

1977 по 1979 г., W. Ray et al. [19] выявили статистически значимую зависимость между применением бензодиазепинов длительного действия и повышенным риском падений (ОР=2,0 при 95% ДИ от 1,6 до 2,5); однако при анализе более поздних данных (с 1984 по 1985 г.) эта связь уже отсутствовала (ОР=1,3 при 95% ДИ от 0,9 до 1,8). Скорее всего, это объясняется тем, что в ранних исследованиях бензодиазепины избирательно назначали больным с деменцией и тревожностью или психомоторным возбуждением, у которых риск падений был исходно высоким. После опубликования сообщений о взаимосвязи между применением бензодиазепинов и падениями врачи начали проявлять особую осторожность, избегая назначать эти препараты больным, у которых риск падений и так высок; в итоге отмеченная ранее взаимосвязь перестала выявляться.

Вот почему исследователи должны документировать характеристики участников в обеих группах и либо представлять доказательства их сопоставимости, либо применять статистические методы коррекции результатов с учетом имеющихся различий (т.е. рассчитывать стандартизованные показатели). Поскольку невозможно сформировать группы, идеально совпадающие по возрасту участников, при анализе результатов следует вносить соответствующие поправки.

Эффективная коррекция результатов с учетом различий в прогностических факторах возможна лишь при точном измерении последних. В крупных базах данных официальных учреждений (позволяющих сформировать выборку, достаточно большую для оценки частоты развития редких клинических исходов) характеристики конкретных больных могут быть неполными или неточными. Изучая эту проблему, J.G. Jollis et al. [20] сравнили точность информации об одних и тех же больных, которая содержалась в базе данных страховой компании и в личной базе данных кардиолога, занимавшегося научными исследованиями. Показано, что коэффициент согласия (κ) между сравниваемыми базами данных был высоким для наличия сахарного диабета ($\kappa=0,83$; см. главу 2С на компакт-диске) и инфаркта миокарда ($\kappa=0,76$), умеренным для артериальной гипертонии ($\kappa=0,56$), низким для сердечной недостаточности ($\kappa=0,39$) и очень низким для курения ($\kappa=0,19$). Более подробно ограничения баз данных официальных учреждений рассматриваются в главе 2В на компакт-диске.

Даже в тех случаях, когда исследователи сообщают о сопоставимости факторов, которые могут повлиять на прогноз в сравниваемых когортах, и применяют статистические методы коррекции, распределение важных прогностических факторов, оставшихся неизвестными или не-

измеренными, может оказаться разным и по-разному повлиять на изучаемый клинический исход. Возвращаясь к нашему примеру, отметим, что повышенный риск кровотечения может быть связан не только с приемом НПВС, но и с особенностями самой болезни, по поводу которой эти препараты необходимо применять. Поэтому выводы, полученные в когортном исследовании, всегда менее обоснованны по сравнению с результатами хорошо спланированного РКИ

Исследования случай—контроль

Проведение когортного исследования может оказаться нецелесообразным, если изучаемый клинический исход развивается редко или очень медленно. В подобных ситуациях можно сначала сформировать группу *случаев* (включающую больных, у которых уже наступил изучаемый клинический исход), а затем подобрать контрольную группу из лиц без данного заболевания или состояния, схожую по важным прогностическим характеристикам (возраст, пол, наличие сопутствующих заболеваний и др.). В исследовании *случай—контроль* можно рассчитать в обеих группах долю больных, подвергавшихся тем или иным неблагоприятным воздействиям (в том числе множественным), и провести коррекцию результатов с учетом известных и измеряемых прогностических факторов.

Подобная структура исследования была, например, использована для демонстрации существования взаимосвязи между приемом диэтилstilbэстрола беременными и развитием аденокарциномы влагалища у их дочерей спустя многие годы [21]. При проведении РКИ или проспективного когортного исследования для выявления такой причинно-следственной связи потребовалось бы как минимум двадцатилетнее наблюдение и, учитывая редкость заболевания, включение сотен тысяч участниц. В рамках исследования случай—контроль были сформированы две группы. Больные с аденокарциномой влагалища были отнесены к основной группе ($n=8$), молодые женщины без аденокарциномы — к контрольной группе ($n=32$), после чего в обеих группах ретроспективно оценивалась частота воздействия предполагаемого неблагоприятного фактора (прием диэтилstilbэстрола матерью участницы во время беременности). Была выявлена сильная взаимосвязь между воздействием данного препарата на плод и последующим развитием аденокарциномы влагалища; при этом вероятность влияния случая оказалась минимальной ($p<0,00001$). Таким образом, ответ на поставленный вопрос был получен в весьма сжатые сроки (без многолетнего наблюдения) и с использованием небольшой выборки (всего 40 женщин).

В качестве другого примера можно рассмотреть исследование случай—контроль, проведенное для выявления возможной зависимости между применением β -агонистов при бронхиальной астме и смертностью; с этой целью была использована компьютерная база данных, в которой содержались сведения о получателях медицинских страховых услуг (95% населения провинции Саскачеван на западе Канады) и назначении лекарственных препаратов [22]. В основную группу включили 129 больных со смертельным (или угрожавшим жизни) приступом бронхиальной астмы, а в контрольную группу — 655 больных с более легкой формой этого заболевания.

Тенденция к более частому применению β -агонистов при более тяжелой форме бронхиальной астмы может привести к ложному выводу о наличии прямой связи между этим вмешательством и риском смерти. Однако исследователи попытались учесть возможное влияние такого фактора, как тяжесть заболевания, путем определения числа госпитализаций в течение 24 мес до смерти (в основной группе) или до включения в исследование (в контрольной группе), а также индекса общей дозы лекарственных препаратов. В итоге была выявлена связь между регулярным применением высоких доз (β -агонистов с помощью ингалятора-дозатора и повышенным риском смерти от бронхиальной астмы (0111=2,6 в пересчете на 1 флакон в месяц; 95% ДИ от 1,7 до 3,9); статистическая значимость полученных данных сохранялась и после их коррекции с учетом тяжести заболевания.

На результаты исследований случай—контроль, как и на результаты когортных исследований, могут влиять неучтенные факторы, особенно если изучаемое воздействие изменяется с течением времени. Например, число госпитализаций и общая доза препаратов вряд ли полностью отражают тяжесть БА. Кроме того, усилить выявленную связь могло наличие других вредных привычек у больных, злоупотребляющих β -агонистами. Неправильный отбор больных в контрольную группу также может способствовать незаметному для авторов искажению результатов. Например, при изучении взаимосвязи между потреблением кофе и развитием рака поджелудочной железы обе группы были сформированы из больных, наблюдавшихся у одних и тех же врачей [23]. В контрольной группе больные страдали различными желудочно-кишечными заболеваниями, обострявшимися при употреблении кофе; поэтому многие больные исключали этот продукт из своего рациона. В итоге исследователи выявили связь между употреблением кофе (которое у больных основной группы было таким же, как в общей популяции) и развитием рака поджелудочной железы. В дальнейших исследованиях с правильно

подобранной контрольной группой наличие этой связи подтвердить не удалось [24, 25]. Таким образом, достоверность результатов обсервационных исследований всегда ограничена, даже если была проведена коррекция результатов с учетом всех известных факторов, влияющих на прогноз.

Серии случаев и описания случаев

В исследованиях, получивших название *серии случаев* (описание определенной серии больных) и *описания случаев* (описание отдельных больных), группа сравнения отсутствует, поэтому рассуждения об одинаковом или различном прогнозе теряют всякий смысл. Хотя в ходе такого описательного исследования иногда можно прийти к выводу, диктующему незамедлительное изменение тактики врача (как, например, произошло при выявлении зависимости между приемом талидомида во время беременности и появлением на свет детей с врожденными аномалиями [26]), следует помнить, что при низкой достоверности результатов исследования последствия таких изменений могут оказаться нежелательными. Так, после появления сообщений о тератогенном действии препарата бендектина (комбинация доксиламина, пиридоксина и дицикломина, применяемая в качестве противорвотного средства при беременности) фирма-производитель отозвала его с фармацевтического рынка [27]. Впоследствии в ряде сравнительных испытаний была подтверждена относительная безопасность этого препарата, однако из-за распространившегося отрицательного к нему отношения возобновить продажи не удалось. Таким образом, многие беременные, которым по тем или иным причинам показан прием бендектина, фактически лишились возможности получать симптоматическую терапию.

В целом на основании описания клинических случаев врачи не должны делать выводы о причинно-следственных связях; такие данные скорее интересны для контролирующих органов и исследователей, поскольку могут выявить вопросы, требующие дальнейшего изучения.

Структура исследований: заключение

При решении вопроса о степени риска (как и при решении вопроса об эффективности лечебного вмешательства) врач должен сначала искать материалы РКИ. Часто такой поиск оказывается безрезультатным, и врач вынужден обращаться к исследованиям с иной структурой. Однако, какой бы ни была эта структура, следует внимательно изучить информацию о контрольной группе, прежде чем делать окончательный вывод о предполагаемом вредном воздействии. В РКИ и когортных ис-

следованиях исходный риск развития изучаемого клинического исхода в группах сравнения должен быть одинаковым; в противном случае должны быть использованы соответствующие статистические методы коррекции данных. В исследованиях случай—контроль основная и контрольная группы также должны быть сходными по характеристикам, которые могут повлиять на клинический исход (за исключением изучаемого фактора). Если же полного сходства между группами добиться не удалось, исследователям следует применять специальные статистические методы коррекции результатов. Даже в том случае, когда исследователи приняли все меры, чтобы свести к минимуму вероятность возникновения систематических ошибок, врач должен помнить, что в наблюдательных исследованиях какие-то неучтенные различия между группами всегда могут снизить достоверность результатов [29]. Поскольку в реальной клинической практике применение того или иного метода лечения основывается не только на доказательной информации, но также на представлениях врача, жизненных ценностях и предпочтениях больного, в нерандомизированных исследованиях выраженность нежелательных воздействий в разных группах скорее всего будет различаться (т.е. не исключена вероятность систематической ошибки, связанной с назначением препаратов с учетом прогноза).

ИСПОЛЬЗОВАЛИСЬ ЛИ В ОСНОВНОЙ И КОНТРОЛЬНОЙ ГРУППАХ ОДИНАКОВЫЕ МЕТОДЫ ДЛЯ ОЦЕНКИ ВЛИЯНИЯ ИЗУЧАЕМОГО ФАКТОРА РИСКА?

При проведении исследований случай—контроль главная задача заключается в подтверждении воздействия изучаемого фактора риска. Например, вероятность положительного ответа на вопрос о предшествующем контакте с растворителями у больных лейкозом выше, чем у лиц в контрольной группе, отчасти из-за повышенной заинтересованности опрашиваемого (*систематическая ошибка, связанная с влиянием клинического исхода на воспоминания о воздействии*), отчасти из-за большей настойчивости лица, ведущего опрос (*систематическая ошибка, связанная с болеетцательным вопросом участников одной из сравниваемых групп*). Анализируя результаты исследования случай—контроль, необходимо обращать внимание на то, пытались ли исследователи свести к минимуму вероятность возникновения таких систематических ошибок (например, применялся ли слепой метод, т.е. проводилось ли сокрытие гипотезы исследования от участников и лиц, проводящих опрос). Так, в одном из исследований случай—контроль было выявлено двукратное повышение риска перелома шейки бедренной кости при употреблении психотропных препаратов. При этом информация об употреблении пре-

паратов была получена из компьютерной базы данных страховой программы Michigan Medicaid program, что позволило избежать вышеуказанных систематических ошибок [30]. При проведении исследования, выявившего связь между применением Р-агонистов и смертностью от бронхиальной астмы, также использовалась информация из административной базы данных, позволявшая объективно оценить изучаемое воздействие [9]. В обоих случаях наша уверенность в отсутствии систематических ошибок при оценке влияния изучаемого фактора риска повышает доверие к выводам авторов.

Использовались ли в основной и контрольной группах одинаковые методы для оценки клинических исходов?

При проведении РКИ и когортных исследований главная задача заключается в том, чтобы установить факт наступления изучаемого клинического исхода. Наблюдаемое трехкратное повышение риска возникновения злокачественной меланомы у лиц, работающих с радиоактивными веществами, можно объяснить тем, что врачи, которые знают о неблагоприятном воздействии и потому ожидают наступления клинического исхода, более тщательно обследуют больных. При этом выявляются болезни или состояния, которые в другом случае могли остаться незамеченными или быть диагностированы в более поздних сроках. Это в свою очередь может привести к завышенной оценке риска в когорте лиц, подверженных радиоактивному излучению (*систематическая ошибка, связанная с влиянием ожидаемого клинического исхода на особенности обследования*) [31].

КАКОВА ДОЛЯ БОЛЬНЫХ, НАБЛЮДАВШИХСЯ ДО КОНЦА ИССЛЕДОВАНИЯ?

Ранее уже отмечалось (см. главу 4 *Исследования, посвященные эффективности лечения*), что неполное наблюдение также способствует возникновению систематической ошибки, поскольку клинические исходы у выбывших и оставшихся участников могли существенно различаться. Чем больше требуемая продолжительность наблюдения, тем выше вероятность того, что оно будет неполным.

Например, в хорошо спланированном исследовании была проведена оценка состояния здоровья у 1235 (98%) из 1261 мужчин европеоидной расы, занятых на производстве хризотил-асбестового текстиля в период с 1940 по 1975 г. [32]. При продолжительности контакта с этим материалом ≥ 15 лет относительный риск смерти от рака легких повышался с 1,4 до 18,2 прямо пропорционально общей продолжительности и интенсивности контакта. Поскольку отсутствие данных о 2% участников вряд ли

могло повлиять на результаты, частота выбывания не отразилась на достоверности вывода о том, что контакт с асбестом повышает смертность от рака легких.

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

Вернемся к обсуждению материалов исследования случай—контроль [6], которое мы обнаружили, пытаясь оценить взаимосвязь между приемом СИОЗС и риском развития желудочно-кишечного кровотечения. Информация для этого исследования была получена из британской электронной базы данных врачей общей практики, содержащей сведения более чем о 3 000 000 человек; большая часть сведений включалась в эту базу данных на протяжении предшествующих 5 лет [33—35]. За период с 1993 по 1997 г. исследователи выявили все случаи возникновения кровотечения из верхних отделов ЖКТ ($n=1651$) и перфорации язвы ($n=248$) у больных в возрасте от 40 до 79 лет. Затем в контрольную группу были рандомизированно отобраны 10 000 участников с таким же исходным риском развития вышеуказанных осложнений, что и у больных в основной группе. Отбор производился таким образом, чтобы группы были сходными по следующим характеристикам: возраст, пол и год, в котором сведения о больных включались в базу данных.

При анализе учитывался ряд прогностических факторов: наличие в анамнезе диспепсии, гастрита, кровотечения из верхних отделов ЖКТ, язвы или ее перфорации, курение, использование НПВС, антикоагулянтов, кортикостероидных препаратов и аспирина (информация о применении препаратов, отпускаемых без рецепта, отсутствовала). После учета прогностических факторов исследователи сравнили в основной и контрольной группах долю больных, которые принимали СИОЗС в течение 30 сут, предшествовавших определенной дате (в основной группе это был день, когда возникло кровотечение или перфорация, а в контрольной группе — день, выбранный случайным образом).

Хотя исследователи учли ряд прогностических факторов, вполне возможно существование других важных факторов, также влияющих на риск развития кровотечения, но оставшихся неучтенными. Например, у больных, получающих лечение по поводу депрессии или тревожности, болевые синдромы различной природы могут наблюдаться чаще, чем в отсутствие названных состояний. Не исключено, что та-

кие больные принимали НПВС, отпускаемые без рецепта, что не было отражено в использованной базе данных.

Другим фактором, который мог повлиять на результаты исследования, является злоупотребление алкоголем. Хотя данные о больных с подтвержденным алкоголизмом не включали в исследование, во многих случаях подобные проблемы могли остаться нераспознанными врачом общей практики. Известно, что алкоголизм повышает риск развития депрессии и тревожности, в связи с чем больным могли назначать СИОЗС. При алкоголизме повышен риск развития желудочно-кишечного кровотечения; этот прогностический фактор, будучи неучтенным, также мог повлиять на результаты исследования. И наконец, у больных, которые приходили на повторный прием за новым рецептом для приобретения СИОЗС, вероятность выявления желудочно-кишечного кровотечения могла быть выше, чем у лиц, находящихся под менее пристальным наблюдением (*систематическая ошибка, связанная выявлением определенного клинического исхода*).

В ходе этого исследования все указанные систематические ошибки могли повлиять на оценку безопасности любых антидепрессантов (т.е. СИОЗС, неселективных ингибиторов обратного захвата серотонина и препаратов других фармакологических групп). Полученные данные, которые мы обсудим ниже, указывают на наличие взаимосвязи между развитием желудочно-кишечного кровотечения и приемом только препаратов группы СИОЗС; в отношении других антидепрессантов такая связь не выявлена. Влияние возможных систематических ошибок должно быть одинаковым при оценке безопасности любых антидепрессантов. Тот факт, что риск возникновения кровотечения оказался повышенным только на фоне приема СИОЗС, вселяет уверенность в том, что достоверность результатов не была снижена из-за возможных различий в прогностических факторах в основной и контрольной группах.

В то же время большинство врачей основывают свое решение о применении СИОЗС или трициклических антидепрессантов на характеристиках конкретного больного. Не исключено, что некоторые из этих характеристик могут быть так или иначе связаны с повышенным риском возникновения желудочно-кишечного кровотечения (например, в гипотетической ситуации, когда врач предпочтет назначить СИОЗС больному, злоупотребляющему алкоголем).

Основным преимуществом использования крупной базы данных в этом исследовании можно считать то, что в обеих группах была исключена возможность возникновения систематической ошибки, связанной с влиянием клинического исхода на воспоминания об изучаемом воздействии (приеме СИОЗС). Клинические исходы и воздейст-

вия, по всей видимости, в обеих группах оценивались одинаково, поскольку большинство лечащих врачей не знали о том, что кровотечение из верхних отделов ЖКТ может быть связано с приемом СИОЗС. Однако информация о числе выбывших больных отсутствует. В анализ включили данные только о больных, которые наблюдались у определенных врачей общей практики от начала до конца исследования; при этом остались неясными некоторые детали (например, сколько больных, сведения о которых содержались в базе данных, начали принимать СИОЗС, но впоследствии отказались от такого лечения).

Таким образом, для рассматриваемого исследования характерен недостаток, присущий любому наблюдению: основная и контрольная группы могут различаться по исходным прогностическим характеристикам. В данном случае по крайней мере два неучтенных фактора (применение отпускаемых без рецепта НПВС и алкоголизм) могли привести к ошибочному выводу о наличии взаимосвязи между приемом СИОЗС и риском возникновения желудочно-кишечного кровотечения. Другим существенным недостатком этого следует считать отсутствие информации о полноте наблюдения. Хотя эти недостатки несколько снижают обоснованность выводов, можно утверждать, что исследование проведено на достаточно высоком методологическом уровне и заслуживает ознакомления с его результатами.

КАКОВЫ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ?

НАСКОЛЬКО СИЛЬНА СВЯЗЬ МЕЖДУ ВОЗДЕЙСТВИЕМ ФАКТОРА РИСКА И КЛИНИЧЕСКИМ ИСХОДОМ?

Взаимосвязь между воздействием и клиническим исходом можно представить в виде таких показателей, как *относительный риск* (ОР) и *отношение шансов* (ОШ), которые описаны в других главах (см. главы 4 *Исследования, посвященные эффективности лечения* и 2В2 на компакт-диске). В когортном исследовании, посвященном оценке больничной смертности среди мужчин-ветеранов после некардиологических операций, из 289 больных с артериальной гипертонией в анамнезе умерли 23 (7,9%), из 185 больных без этого заболевания — 3 (1,6%). Таким образом, ОР смерти при наличии артериальной гипертонии составил $7,9/1,6 = 4,9$ [36]. Полученное значение ОР указывает на то, что после выполнения некардиологических операций больные с артериальной гипертонией в анамнезе умирают почти в 5 раз чаще, чем больные с нормальным артериальным давлением.

Расчет ОР можно провести, когда в группах известного размера у определенного числа больных наступил изучаемый клинический исход. Поэтому данный показатель неприменим при проведении исследований случай—контроль, в которых число участников в группах сравнения (т.е. доля больных с изучаемым клиническим исходом) выбирается самими исследователями. При этом используется другой показатель — ОШ (отношение шансов наличия изучаемого фактора риска у больных основной и контрольной группы; см. главу 2В2 на компакт-диске).

Доказательством вредного воздействия можно считать даже небольшое повышение риска, если оно было продемонстрировано в исследовании высокого методологического качества (РКИ). В исследованиях с иной структурой (когортные исследования и исследования случай—контроль) требуется выявить гораздо большее повышение риска, поскольку выше вероятность того, что незначительные его изменения обусловлены влиянием систематической ошибки. Высокие значения ОР или ОШ указывают на сильную корреляцию, которую труднее объяснить влиянием неучтенных факторов.

Кроме высокого значения ОР или ОШ есть еще один признак, указывающий на истинную вредность изучаемого воздействия. Если при увеличении интенсивности или продолжительности воздействия риск оцениваемого клинического исхода повышается (т.е. данные указывают на зависимость эффекта от дозы), причинно-следственная связь между воздействием и клиническим исходом скорее всего существует. В одном из исследований, включавших курящих мужчин-врачей, было показано, что риск смерти от рака легких повышается на 50, 132 и 220% у выкуривающих по 1—14, 15—24 и ≥ 25 сигарет в сутки соответственно; этот факт лишний раз подтверждает вывод о повышенном риске возникновения рака легких у курильщиков [37].

НАСКОЛЬКО ТОЧНО РАССЧИТАН РИСК РАЗВИТИЯ НЕБЛАГОПРИЯТНОГО КЛИНИЧЕСКОГО ИСХОДА?

Точность расчета любого показателя риска можно оценить, ознакомившись с диапазоном значений его доверительного интервала (ДИ; см. главы 4 *Исследования, посвященные эффективности лечения* и 2В2 на компакт-диске). Если в исследовании было продемонстрировано наличие связи между воздействием и неблагоприятным клиническим исходом, нижняя граница ДИ (наименьшее значение ОР) соответствует минимальной силе выявленной связи. Если наличие связи продемонстрировать

не удалось (исследования с отрицательным результатом), верхняя граница ДИ указывает на максимальную выраженность неблагоприятного влияния, которое может быть клинически значимым даже в отсутствие статистически значимой связи (см. главу 2В2 на компакт-диске).

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

Снова вернемся к обсуждению материалов исследования, в котором изучалось влияние применения СИОЗС на частоту возникновения желудочно-кишечного кровотечения. Исследователи рассчитали ОШ возникновения кровотечения в основной и контрольной группах, однако результаты были представлены в виде ОР. Увы, такая практика нередка. К счастью, при низкой частоте развития изучаемого клинического значения ОР и ОШ очень близки (см. главу 2В2 на компакт-диске). Связь между развитием кровотечения из верхних отделов ЖКТ и вмешательством оказалась сильной на фоне продолжающегося приема СИОЗС (стандартизованное ОШ=3,0 при 95% ДИ от 2,1 до 4,4), слабой при использовании неселективных ингибиторов обратного захвата серотонина (стандартизованное ОШ=1,4 при 95% ДИ от 1,1 до 1,9) и отсутствовала при использовании антидепрессантов, не влияющих на обмен серотонина. По силе и направленности связь между вмешательством и возникновением кровотечений на фоне приема СИОЗС была практически такой же, как на фоне приема НПВС (стандартизованное ОШ=3,7 при 95% ДИ от 3,2 до 4,4). Одновременный прием СИОЗС и НПВС, отпускаемых по рецепту, в еще большей степени повышал риск возникновения кровотечений из верхних отделов ЖКТ (стандартизованное ОШ=15,6 при 95% ДИ от 6,6 до 36,6). Доза и длительность приема СИОЗС практически не влияли на риск развития этого неблагоприятного клинического исхода.

КАК МОЖНО ПРИМЕНИТЬ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ НА ПРАКТИКЕ?

Когда возможные систематические ошибки в рассматриваемом исследовании не настолько серьезны, чтобы прекратить рассмотрение его результатов, следует решить, в какой мере они применимы при лечении конкретного больного. Совпадают ли характеристики этого больного с

характеристиками участников исследования, включая тяжесть заболевания, возраст, пол, расовую принадлежность и другие важные факторы? Если они не совпадают, будут ли в данном случае отличаться и биологические характеристики вредного воздействия (см. главу 2В3 на компакт-диске)? Насколько различается воздействие (в том числе и медицинское вмешательство), которому были подвержены участники исследования и конкретный больной? Например, полученные в 70-х годах XX века данные о том, что прием пероральных контрацептивов повышает риск развития тромбофлебита, могли оказаться неприменимыми в 90-е годы, поскольку применявшиеся позднее контрацептивы содержали меньше эстрогенов. Данные о повышении заболеваемости раком эндометрия на фоне заместительной эстрогенотерапии в постменопаузе нельзя распространять на женщин, одновременно принимающих прогестин по схеме, составленной таким образом, чтобы при продолжительном нециклическом применении возникали ежемесячные маточные кровотечения.

Вернемся к исследованию, в котором у лиц, занятых в производстве кризотил-асбестового текстиля в 1940—1975 гг. и контактировавших с этим материалом в течение ≥ 15 лет, было выявлено повышение ОР смерти от рака легких с 1,4 до 18,2 в прямой зависимости от общей продолжительности и интенсивности контакта [32]. Наблюдение было достаточно продолжительным, чтобы выявить большинство случаев заболевания, поэтому результаты исследования скорее всего применимы в повседневной практике врача. И наоборот, при слишком коротком наблюдении могут остаться незамеченными неблагоприятные эффекты, которые были бы выявлены при длительном наблюдении.

Значения ОР и ОШ не указывают на реальную частоту возникновения изучаемого клинического исхода; они указывают только на то, что в одной группе он наблюдается чаще, чем в другой. Поэтому необходимо оценивать клиническую значимость таких результатов. В главах, посвященных оценке эффективности лечебных вмешательств (см. главы 4 *Исследования, посвященные эффективности лечения* и 2В2 на компакт-диске), излагается способ расчета такого показателя, как ЧБНЛ (число Зольных, которых нужно лечить определенным методом в течение определенного времени, чтобы достичь определенного благоприятного исхода или предотвратить неблагоприятный клинический исход у одного больного). Когда речь идет об оценке риска, можно использовать ценные РКИ или когортного исследования (к результатам исследования случай—контроль это не относится) и с помощью аналогичного подхода рассчитать такой показатель, как ИПВ (индекс потенциального

вреда; число больных, которые в течение определенного времени должны быть подвержены действию фактора риска, чтобы у одного больного развился неблагоприятный клинический исход).

Например, в испытании CAST (Cardiac Arrhythmia Suppression Trial), посвященном оценке эффективности противоаритмических препаратов [38, 39], в среднем через 10 мес наблюдения смертность в группе плацебо составила 3,0%, а в группе больных, принимавших энкаинид или флекаинид, — 7,7%. Повышение абсолютного риска (ПАР) смерти составило 4,7%. Величина, обратная этому показателю (ИПВ = 1 / ПАР = 21), указывает на то, что при лечении 21 больного энкаинидом или флекаинидом в течение примерно 1 года один из этих больных может умереть. В примере, иллюстрирующем взаимосвязь между приемом НПВС и возникновением желудочно-кишечного кровотечения, значение ИПВ будет совершенно иным. Из 2000 больных, не принимающих НПВС, кровотечение в течение года возникает у двоих; из 2000 больных, принимающих НПВС, — у троих. Таким образом, если мы будем назначать НПВС 2000 больных, то в течение года можно ожидать только одно дополнительное кровотечение [7].

СЛЕДУЕТ ЛИ ПРИНИМАТЬ МЕРЫ, НАПРАВЛЕННЫЕ НА ПРЕКРАЩЕНИЕ НЕБЛАГОПРИЯТНОГО ВОЗДЕЙСТВИЯ?

После того как врач оценил данные о неблагоприятных эффектах лечебного вмешательства (или другого воздействия), а также удостоверился в том, что результаты исследования применимы к конкретному больному, избрать последующую тактику порой бывает нелегко. Принимая клиническое решение, нужно иметь в виду по крайней мере три аспекта.

Во-первых, следует ответить на вопрос, достаточно ли высоким было методологическое качество исследований, в которых продемонстрирован вред изучаемого воздействия. Во вторых, нужно составить представление о том, насколько велик риск при продолжении использования лечебного вмешательства (или влияния другого неблагоприятного фактора). В-третьих, необходимо отдавать себе отчет в нежелательных последствиях прекращения лечения, с которым связаны те или иные неблагоприятные эффекты.

Принятие клинических решений оказывается легким при высокой вероятности развития неблагоприятного исхода и существенном размере этого эффекта. Поскольку данные о повышении смертности в результате приема энкаинида и флекаинида получены в ходе РКИ [38], можно быть уверенным в наличии причинно-следственной связи. Так как один дополнительный случай смерти регистрировался при лечении всего лишь

!1 больного, нет ничего удивительного в том, что врачи немедленно пре-
ратили применение этих антиаритмических препаратов после опубли-
ования результатов исследования.

Решение не очень сложно принять и в тех случаях, когда существует
фиемлемая альтернатива применяемому лечебному вмешательству. Так,
Использование Р-блокаторов для лечения артериальной гипертензии мо-
кет ухудшить дыхательную функцию у больных с сопутствующей брон-
лиальной астмой или хроническими обструктивными заболеваниями
[егких; поэтому у больных данной категории вместо (3-блокаторов сле-
(ует применять другие гипотензивные препараты (например, тиазидные
(иуретики) [40].

Наличие альтернативы облегчает решение даже тогда, когда досто-
(ерность информации о побочных эффектах недостаточна. Например,
(первые взаимосвязь между приемом аспирина и развитием синдрома
³ейе была продемонстрирована в исследованиях случай—контроль весь-
ia низкого методологического качества, что ставило их выводы под со-
мнение. Хотя достоверность данных была невысока, доступность безо-
исного, недорогого и хорошо переносимого альтернативного препара-
a ацетаминофена предопределила его применение у детей с риском раз-
штия синдрома Рейе [41].

Если взаимосвязь между приемом аспирина и развитием синдрома
³ейе, выявленная в первых исследованиях случай—контроль, была со-
мнительной, то корреляция между приемом НПВС и возникновением
сривотечений из верхних отделов ЖКТ убедительно продемонстриро-
шна во многих хорошо спланированных когортных исследованиях и
исследованиях случай—контроль, выводы которых о побочных эффек-
тах препаратов этого класса более обоснованы. Однако риск возникно-
зения кровотечений очень низок, и до последнего времени в арсенале
фаца не было столь же эффективных, но более безопасных противвос-
галительных препаратов. Поэтому продолжение использования НПВС
три наличии показаний было оправданным. Возможно, в ближайшие
юды приемлемой альтернативой традиционным НПВС могут стать дру-
гие препараты этого класса — ингибиторы циклоксигеназы-2, если при
эолее длительном опыте применения подтвердятся их сравнительная
Зезопасность и эффективность затрат.

ПРИНЯТИЕ КЛИНИЧЕСКОГО РЕШЕНИЯ

Перед выработкой клинического решения вы должны пройти все три
этапа изучения медицинской литературы. Во-первых, следует оценить

достоверность результатов исследования, посвященного выявлению взаимосвязи между приемом различных антидепрессантов и риском возникновения кровотечений из верхних отделов ЖКТ или перфорации язвы. Вы решаете, что ограничения, присущие исследованию случай-контроль, а также отсутствие информации о числе больных, которые наблюдались до конца лечения, не позволяют окончательно утверждать, что между приемом СИОЗС и развитием желудочно-кишечного кровотечения существует причинно-следственная связь. Более того, это единственная публикация по данному вопросу, а примеры из других областей медицины свидетельствуют о том, что многие выводы о наличии связи между воздействием и неблагоприятным клиническим исходом, казавшиеся очевидными в ходе первых исследований [10, 46, 47], впоследствии не подтверждались [11, 12, 42—45].

При изучении результатов найденного исследования вы отмечаете очень сильную корреляцию между применением СИОЗС в сочетании с НПВС и риском возникновения кровотечения. Несмотря на методологические ограничения данного исследования, вы решаете, что эта взаимосвязь слишком сильна, чтобы ее игнорировать. Поэтому вы переходите к третьему этапу оценки результатов и пытаетесь определить их применимость в конкретном случае.

При отборе участников исследователи пользовались базой данных о больных, которые наблюдались и получали лечение в условиях первичной медицинской помощи; это указывает на то, что результаты исследования вполне применимы при лечении вашего больного. Вы определяете степень риска, которому можете подвергнуть больного, если при назначении СИОЗС у него действительно возникнет желудочно-кишечное кровотечение. Пользуясь данными об исходном риске в аналогичной популяции [14], вы подсчитываете, что 1 дополнительный случай кровотечения возникает, если в течение года проводить монотерапию СИОЗС примерно 625 больным или комбинированную терапию СИОЗС и НПВС 55 больным.

Из опыта общения с этим больным вы знаете, что он не склонен рисковать. Во время повторного приема вы сообщаете больному об одинаковой эффективности СИОЗС и трициклических антидепрессантов, а также даете сравнительную характеристику их безопасности. Вы также упоминаете о возможном повышении риска возникновения желудочно-кишечного кровотечения на фоне приема СИОЗС. Выслушав вас, больной решает, что при прочих равных условиях ему больше подойдет трициклический антидепрессант, и покидает ваш кабинет с рецептом нортриптилина.

Литература

- 1 } *Geddes J Y , Freemantle N , Mason J , Eccles M P , Boynton J* SSRIs versus other antidepressants for depressive disorder *Cochrane Database of Systematic Reviews* 2000 1—26
- 2 *Tyndale E, Menon D* Selective Serotonin Reuptake Inhibitors (SSRIs) for Major Depression Ottawa Canadian Coordinating Office for Health Technology Assessment, 1997
- 3 *Mulrow C D , Williams J W Jr, Tnvedi M* Treatment of Depression Newer Pharmacotherapies San Antonio, Texas San Antonio Evidence-based Practice Centre, 1999
- 4 *HoioffM, Hardy R , Lewis G* Discontinuation rates of SSRIs and tncych antidepressants a meta-analysis and investigation of heterogeneity *Br J Psychiatry* 1997,170 120—127
- 5 *Tyndale E, Menon D* Review selective serotonin reuptake inhibitors differ from tncvcllc antidepressants in adverse events *Evidence Based Mental Health* 1998,1 50
- 6 *Goldberg R J* Selective serotonin reuptake inhibitors infrequent medical adverse effects *Arch Fam Med* 1998,7 78-84
- 7 *deAbajoFJ, Rodriguez L A , Montero D* Association between selective serotonin reuptake inhibitors and upper gastrointestinal bleeding population based case-control studv *BMJ* 1999 319 1106—1109
- 8 *CAPRIE Steering Committee* A randomised, blinded trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE) *Lancet* 1996,348 1329-1339
- 9 *Bennett C L , Connors J M , Carwile J M , et al* Thrombotic thrombocytopenic purpura associated with clopidogrel *N Engl J Med* 2000,342 1773-1777
- 10 *Siherslem F E, Graham D Y, Senior J R , et al* Misoprostol reduces serious gastrointestinal complications in patients with rheumatoid arthritis receiving nonsteroidal anti-inflammatory drugs a randomized double-blind, placebo-controlled trial *Ann Intern Med* 1995,123 241-249
- 11 *Merck and Co* VIGOR Study Summary Paper presented at Digestive Disease Week Congress, May 24, 2000, San Diego
- 12 *hangman M J, Jensen D M, Watson D J, et al* Adverse upper gastrointestinal effects of rofecoxib compared with NSAIDs *JAMA* 1999,282 1929-1933
- 13 *Knstensen P, Irgens L M, Dalveit A K, Andersen A* Perinatal outcome among children of men exposed to lead and organic solvents in the printing industry *Am J Epidemiol* 1993 137 134—143
- 14 *Carson J L , Strom B L , Soper K A , et al* The association of nonsteroidal anti-inflammatory drugs with upper gastrointestinal tract bleeding *Arch Intern Med* 1987,147 85—88
- 15 *Walter S D* Determination of significant relative risks and optimal sampling procedures in prospective and retrospective comparative studies of various sizes *Am J Epidemiol* 1977,105 387—397
- 16 *Leujkens H G , Urquhart J, Strieker B H, Bakker A , Pern H* Channelling of controlled release formulation of ketoprofen (Oscorel) in patients with history of gastrointestinal problems *J Epidemiol Community Health* 1992,46 428-432
- 17 *Joseph K S* The evolution of clinical practice and time trends in drug effects *J Clin Epidemiol* 1994,47 593-598
- 18 *LeufkemHG Urquhart J* Variability in patterns of drug usage *J Pharm Pharmacol* 1994.46(suppl 1) 433—437
- 19 *Ray W A , Griffin M R , Downey W* Benzodiazepines of long and short elimination half-life and risk of hip fracture *JAMA* 1989,262 3303-3307
- 20 *JollisJ G , Ancukiewicz M , DeLong E R , Pryor D B , Muhlbaier L H , Mark D B* Discordance of databases designed for claims payment versus clinical information systems implications for outcomes research *Ann Intern Med* 1993,119 844—850
- 21 *HerbstA L , Uffelder H , Poskanzer D C* Adenocarcinoma of the vagina association of maternal stilbestrol therapy with tumor appearance in young women *N Engl J Med* 1971 284 878—881
- 22 *Spitzer W O , Suissa S Ernst P, et al* The use of beta-agonists and the risk of death and near death from asthma *N Engl J Med* 1992,326 501-506

- 23 *MacMahon B, Yen S, Tnchopoulos D, Warren K, Nardi G* Coffee and cancer of the pancreas *N Engl J Med* 1981,304 630-633
- 24 *Baghurst PA, McMwhael A J, Slavotmeck A H, Baghurst KI, Boyle P, Walker A M* A case-control study of diet and cancer of the pancreas *Am J Epidemiol* 1991,134 167-179
- 25 *Zheng W, McLaughlm J K, Gndley G, et al* A cohort study of smoking alcohol consumption and dietary factors for pancreatic cancer *Cancer Causes Control* 1993,4 477-482
- 26 *Lenz W* Epidemiology of congenital malformations *Ann NY Acad Sci* 1965,123 228-236
- 27 *Soverchia G, Pern P F* Two cases of malformation of a limb in infants of mothers treated with an antiemetic in a very early phase of pregnancy *Pediatr Med Chir* 1981 3 97-99
- 28 *Holmes L B* Teratogen update Bendectin *Teratology* 1983,27 277-281
- 29 *Kellermann A L, Rivara F P, Rushforth N B, Banton J G, et al* Gun ownership as a risk factor for homicide in the home *N Engl J Med* 1993,329 1084-1091
- 30 *Ray WA, Griffin M R, Schaffner W, et al* Psychotropic drug use and the risk of hip fracture *N Engl J Med* 1987,316 363-369
- 31 *Hmtt RA, Fireman B* The possible effect of increased surveillance on the incidence of malignant melanoma *Prev Med* 1986,15 652-660
- 32 *Dement J M, Hams R L Jr, Symons M J, Shy C M* Exposures and mortality among chrysotile asbestos workers Part II mortality *Am J Ind Med* 1983,4 421-433
- 33 *Jick H, Jick S S, Derby L E* Validation of information recorded on general practitioner based computerised resource in the United Kingdom *BMJ* 1991,302 766-768
- 34 *Garcia Rodriguez L A, Perez Gutthann S* Use of the UK general practice research database for pharmacoepidemiology *BrJ Clin Pharmacol* 1998,45 419-425
- 35 *Jick H, Terns B, Derby L E, Jick S S* Further validation of information recorded on a general practitioner database resource in the United Kingdom *Pharmacoepidemol Drug Saf* 1992,1 347-349
- 36 *Browner W S, Li J, Mangano D T* In-hospital and long-term mortality in male veteran following noncardiac surgery *JAMA* 1992,268 228-232
- 37 *Doll R, Hill A B* Mortality in relation to smoking ten years' observation of British doctors *BMJ* 1964,1 1399-1410, 1460-1467
- 38 *Eiht 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 324 781-788
- 39 *The Cardiac Arrhythmia Suppression Trial II Investigators* Effect of the antiarrhythmic agent monicizine on survival after myocardial infarction *N Engl J Med* 1992 327 227-233
- 40 *Ogilvie R I, Burgess E D, Cusson J R, Feldman R D, letter L A, Myers M G* Report of the Canadian Hypertension Society Consensus Conference, 3 pharmacologic treatment of essential hypertension *Can Med Assoc J* 1993 149 575-584
- 41 *Soumerai S B, Ross-Degnan D, Kahn J S* Effects of professional and media warnings about the association between aspirin use in children and Reye's syndrome *Milbank Q* 1992,70 155-182
- 42 *Danesh J, Appleby P* Coronary heart disease and iron status meta-analyses of prospective studies *Circulation* 1999,99 852-854
- 43 *Klebanoff M A, Read J S, Mills J L, Shiono P H* The risk of childhood cancer after neonatal exposure to vitamin K *N Engl J Med* 1993,329 905-908
- 44 *Passmore S J, Draper G, Brownbill P, Kroll M* Case-control studies of relation between childhood cancer and neonatal vitamin K administration retrospective case-control study *BMJ* 1998 316 178-184
- 45 *Parker L, Cole M, Craft A W, Hey E N* Neonatal vitamin K administration and childhood cancer in the north of England *BMJ* 1998,316 189-193
- 46 *Salonen J T, Nyyssonen K, Korpela H, Tuomilehto J, Seppanen R, Satonen R* High stored iron levels are associated with excess risk of myocardial infarction in eastern Finnish men *Circulation* 1992 86 803-811
- 47 *GoldmJ, Greenwood R, Birmingham K, Mott M* Childhood cancer, intramuscular vitamin K, and pethidine given during labour *BMJ* 1992,305 341-346

6

ДИАГНОСТИКА

У.С. Ричардсон, М. Уилсон, Г. Гайятт

В подготовке данной главы принимали также участие члены Рабочей группы по разработке и внедрению принципов доказательной медицины П. Уайер, Дж. Крэйг, Р. Йешке, Д. Лиймер, Л.М. Летельер, В. Мойер, К.Д. Нэйлор и Д. Кук

В ЭТОМ РАЗДЕЛЕ

Общие принципы диагностики

Рассмотрение альтернативных вариантов при дифференциальной диагностике

Признаки, указывающие на необходимость дополнительного обследования или немедленного начала лечения

Применение результатов систематических исследований в процессе диагностики

ВОЗМОЖНАЯ КЛИНИЧЕСКАЯ СИТУАЦИЯ

Проведение дифференциальной диагностики

Представьте себе обычный суматошный день в отделении неотложной медицинской помощи. Одна из медицинских сестер сообщает вам, что поступил мужчина в возрасте 60 лет с жалобами на сильный кашель, продолжающийся в течение суток. Еще не видя больного, вы предполагаете, что такой кашель может быть обусловлен инфекцией верхних дыхательных путей и, возможно, пневмонией. Входя в палату, вы сразу отмечаете у больного выраженную одышку и то, что его состояние явно более тяжелое, чем обычно наблюдается при указанных заболеваниях. В памяти всплывают другие возможные диагнозы: острый приступ обструкции дыхательных путей, инфаркт миокарда с отеком легких, пневмоторакс или тромбоэмболия легочной артерии (ТЭЛА). Вы начинаете собирать анамнез, предварительно распорядившись начать постоянную регистрацию гемодинамических показателей и газового состава крови, установить внутривенный катетер, выполнить ЭКГ в 12 отведениях и рентгенографию грудной клетки с помощью передвижного рентгеновского аппарата.

У больного умеренно увеличена частота дыхательных движений (24/мин), однако он способен произносить законченные предложения. Артериальное давление составляет 140/90 мм рт. ст., частота сердечных сокращений — 96 уд/мин, аритмии нет. Учитывая наличие одышки, вы просите медицинскую сестру измерить у больного ректальную температуру. По данным пульсоксиметрии, насыщение гемоглобина крови кислородом (SO_2) составляет 93%, поэтому вы сразу назначаете кислородотерапию (через носовые канюли со скоростью 4 л/мин).

Больной говорит, что еще сутки назад он прекрасно себя чувствовал, однако затем начался сильный кашель без предшествовавших или сопутствовавших лихорадки, насморка, боли в горле, головной или мышечной боли. В течение нескольких первых часов его беспокоила тупая боль за грудиной, которая самопроизвольно исчезла. При кашле отходила скудная прозрачная мокрота, в которой на протяжении последних 2 ч больной отмечает прожилки алой крови. В последние 12 ч одышка нарастала: сначала она появлялась при незначительной физической нагрузке, затем стала отмечаться и в покое

При аускультации сердца дополнительные тоны и шумы не выслушиваются. При общем осмотре выявляются только ослабленное дыхание и сухие хрипы в нижних отделах левого легкого. На ЭКГ отмечает-

—

ся незначительная синусовая тахикардия; других отклонений нет. На рентгенограмме грудной клетки выявляются незначительный выпот в левой плевральной полости и слабо выраженное затемнение в нижней доле левого легкого. Медицинская сестра сообщает, что ректальная температура составляет 38,1 °С. Вы назначаете вентиляционно-перфузионную сцинтиграфию легких и просите сообщить результаты анализа газового состава артериальной крови. По данным этого анализа, при дыхании атмосферным воздухом напряжение углекислого газа (PCO_2) нормальное, а напряжение кислорода (PO_2) снижено до 70 мм рт. ст. при SO_2 93%.

Ожидая результатов вентиляционно-перфузионной сцинтиграфии легких, вы пытаетесь оценить вероятность наличия у больного ТЭЛА с учетом имеющихся данных. С одной стороны, у него отсутствуют факторы риска развития данного заболевания, равно как и характерные клинические проявления (боль при дыхании и изменения на ЭКГ); в то же время имеется сильный кашель. С другой стороны, вы уверены, что исключили иные возможные диагнозы, в том числе бронхиальную астму, отек легких и пневмоторакс, а имеющаяся у больного клиническая картина не характерна для пневмонии. В итоге вы приходите к выводу, что вероятность наличия ТЭЛА в данном случае составляет примерно 30%. После этого становится известно, что при вентиляционно-перфузионной сцинтиграфии в одном из сегментов легкого было выявлено несоответствие вентиляции и перфузии. Насколько вы знаете, при таком результате отношение правдоподобия наличия ТЭЛА составляет 18% [1]. С помощью специальной таблицы (см. главу 8 *Диагностические методы*) вы определяете, что постлетестовая вероятность наличия ТЭЛА равна 90%, и начинаете терапию антикоагулянтами.

ОБЩИЕ ПРИНЦИПЫ ДИАГНОСТИКИ

Установление диагноза представляет собой сложный мыслительный процесс, включающий в себя как логические построения, так и учет особенностей клинической картины [2, 3]. Хотя данный процесс во многом происходит на уровне подсознания, в нем можно выделить 2 основных этапа.

Этап 1. В первую очередь вы перечисляете возможные диагнозы и оцениваете их относительное правдоподобие [4]. Опытные врачи часто объединяют наблюдающиеся изменения, описывая их несколькими сло-

вами, например, "генерализованный зуд", "безболевая желтуха", "общие симптомы". Эти комплексные формулировки отражают клинические проблемы, которые могут иметь биологическую, психологическую или социальную природу; именно на их основании проводится дифференциальная диагностика. В вышеуказанной ситуации у 60-летнего мужчины, прежде не жаловавшегося на здоровье, в течение предшествующих суток появились кашель и одышка. Дифференциальную диагностику при этом следует проводить между инфекцией дыхательных путей, острым приступом обструкции дыхательных путей, инфарктом миокарда с отеком легких, пневмотораксом и ТЭЛА.

Этап 2. На основании результатов обследования вы уточняете вероятность наличия тех или иных заболеваний или состояний, исключаете некоторые из них и в итоге останавливаетесь на самом правдоподобном варианте. Дополнительная информация повышает или снижает вероятность каждого из возможных диагнозов. В данном случае отсутствие клинических проявлений, характерных для инфекционного процесса, снижает вероятность наличия у больного инфекции верхних дыхательных путей или пневмонии. Непродолжительная тупая боль за грудиной позволяет предположить наличие инфаркта миокарда, протекающего нетипично, и своевременно выполнить ЭКГ. После общего осмотра сердечная недостаточность кажется менее вероятной причиной имеющихся клинических проявлений, и двумя возможными диагнозами остаются пневмония и ТЭЛА. После рентгенографии грудной клетки, не выявившей характерных признаков пневмонии, возникает необходимость в проведении дополнительного диагностического исследования — вентилационно-перфузионной сцинтиграфии легких.

Таким образом, каждый вновь обнаруживаемый факт позволяет продвигаться от оценки претестовой вероятности наличия того или иного заболевания к оценке послетестовой вероятности, хотя этот путь во многом интуитивен. Некоторые данные, такие как отсутствие признаков пневмоторакса на рентгенограммах грудной клетки, позволяют исключить определенный диагноз (его послетестовая вероятность становится равной 0). До проведения последнего исследования предпочтение диагнозу ТЭЛА было отдано исключительно в результате количественного анализа (претестовая вероятность наличия ТЭЛА равнялась 30%); использованные затем данные литературы позволили рассчитать послетестовую вероятность этого диагноза, которая составила 90%.

Имея представление о ценности каждого фрагмента полученной информации, можно провести точную количественную оценку пре- и по-

слетестовой вероятности каждого из предполагаемых диагнозов. Ниже будет показано, каким образом это осуществляется.

Информативность отдельных данных, получаемых в ходе сбора анамнеза и проведения общего осмотра, не всегда ясна, поэтому при определении их возможного влияния на процесс дифференциальной диагностики приходится опираться на личный опыт и интуицию. В некоторых случаях (в том числе и при диагностике ТЭЛА) интуиция врача часто позволяет прийти к правильному выводу [1].

РАССМОТРЕНИЕ АЛЬТЕРНАТИВНЫХ ВАРИАНТОВ ПРИ ДИФФЕРЕНЦИАЛЬНОЙ ДИАГНОСТИКЕ

Как определить, какие варианты следует рассматривать при дифференциальной диагностике? Если предположить, что вероятность нескольких диагнозов примерно одинакова, и проводить соответствующее обследование одновременно (подход, основанный на рассмотрении всех возможных диагнозов), придется выполнить множество ненужных тестов. Вместо этого опытные врачи в первую очередь исключают или подтверждают наличие наиболее вероятных заболеваний (вероятностный подход), заболеваний с высоким риском развития тяжелых осложнений в отсутствие диагностики и лечения (прогностический подход) или заболеваний, наиболее успешно поддающихся лечению (прагматический подход).

При правильном проведении дифференциальной диагностики следует использовать все три подхода (вероятностный, прогностический и прагматический). Диагноз, наилучшим образом объясняющий наблюдаемые изменения, можно принять в качестве основной гипотезы или рабочей версии. В нашем клиническом примере первичной рабочей версией (до получения результатов тестов) был диагноз инфекции дыхательных путей. В начале обследования в качестве альтернативных вариантов можно рассматривать другие возможные диагнозы (как правило, от 1 до 5) в зависимости от вероятности их наличия, риска развития тяжелых осложнений в отсутствие диагностики и соответствующего лечения, а также от эффективности последнего. У данного больного ТЭЛА была включена в список предполагаемых заболеваний на ранних этапах дифференциальной диагностики в связи с высокими риском развития тяжелых осложнений и эффективностью существующих методов лечения.

Другие причины возникновения определенной клинической проблемы, или дополнительные гипотезы, вероятность которых на момент на-

чала обследования представляется невысокой, могут рассматриваться после того, как рабочая версия и альтернативные варианты будут отвергнуты. В нашем примере такие дополнительные гипотезы, как наличие легочного кровотечения или системного васкулита, не были включены в процесс дифференциальной диагностики. Однако это могло произойти, если бы ни один из альтернативных вариантов не подтвердился.

ПРИЗНАКИ, УКАЗЫВАЮЩИЕ НА НЕОБХОДИМОСТЬ ДОПОЛНИТЕЛЬНОГО ОБСЛЕДОВАНИЯ ИЛИ НЕМЕДЛЕННОГО НАЧАЛА ЛЕЧЕНИЯ

Представим себе больного, у которого на кожных покровах, по ходу одного нерва, появились болезненные сгруппированные пузырьковые высыпания. Опытный врач немедленно диагностирует герпетическую инфекцию (опоясывающий герпес) и начнет подбирать схему лечения. Другими словами, вероятность наличия этого заболевания настолько высока (около 1,0, или 100%), что дополнительное обследование не требуется.

Рассмотрим другой пример. У бейсболиста, прежде не жаловавшегося на здоровье, после случайного удара битой во время игры возникла боль в области ребер с одной стороны грудной клетки. Опытный врач предположит, что причиной посттравматического болевого синдрома стал ушиб (рабочая версия) или перелом ребра (альтернативный вариант); для исключения последнего следует выполнить рентгенографию грудной клетки. Врач также сможет перечислить те диагнозы (например, инфаркт миокарда), которые в силу их низкой вероятности не имеет смысла анализировать. Таким образом, хотя вероятность наличия перелома ребра ниже, чем вероятность ушиба, она все же превышает значение, указывающее на необходимость дополнительного обследования, а вероятность наличия инфаркта миокарда этого значения не достигает.

Рассмотренные примеры иллюстрируют, каким образом можно оценить вероятность наличия того или иного заболевания и соотнести ее с двумя важными пороговыми значениями (см. рисунок). Если вероятность диагноза такова, что при более высоких значениях можно без колебаний начинать лечение, не прибегая к дополнительным тестам, речь идет о верхнем пороговом значении, превышение которого позволяет назначать лечение без дополнительного обследования (обозначим это пороговое значение вероятности как ПЛ). Так, в примере с пузырьковыми высыпаниями вероятность наличия опоясывающего герпеса была явно

От чего зависит выбираемое значение ПЛ? Чем более выражены побочные эффекты терапии, тем выше ПЛ. Так, при наличии ТЭЛА требуется длительное применение антикоагулянтов, существенно повышающее риск возникновения кровотечений; поэтому ошибочная постановка такого диагноза очень опасна для больного. Инвазивность последующего планируемого диагностического теста также сказывается на ПЛ. Если бы в нашем основном клиническом примере после выполнения вентиляционно-перфузионной сцинтиграфии легких наличие ТЭЛА не подтвердилось, пришлось бы выбрать более высокое значение ПЛ. С другой стороны, нежелание применять инвазивные тесты (такие как ангиография легочных артерий), повышающие риск развития тяжелых осложнений, способствует снижению ПЛ. Мы готовы смириться с возможностью гипердиагностики при более низком ПЛ, так как более высокое ПЛ может стать причиной неоправданного применения инвазивных методов исследования у некоторых больных.

Аналогичным образом устанавливают ПДО. Чем опаснее то или иное заболевание в отсутствие лечения, тем ниже ПДО. Поскольку несвоевременно диагностированная ТЭЛА может привести к летальному исходу, мы должны выбрать низкое значение ПДО. Однако и в этом случае следует учитывать возможные осложнения планируемых методов обследования. Если риск развития подобных осложнений невысок, значение ПДО может быть низким; в противном случае оно должно быть установлено на более высоком уровне.

ПРИМЕНЕНИЕ РЕЗУЛЬТАТОВ СИСТЕМАТИЧЕСКИХ ИССЛЕДОВАНИЙ В ПРОЦЕССЕ ДИАГНОСТИКИ

Каким образом врачи отбирают варианты, которые следует включить в процесс дифференциальной диагностики, и оценивают их претестовую вероятность? Они вспоминают сходные клинические случаи из собственной практики; при этом вероятность наличия распространенных заболеваний будет выше, чем редких. Вспомнить случаи из практики достаточно просто, к тому же они отражают характерные особенности данного региона или лечебного учреждения. Однако человеческая память несовершенна, вследствие чего оценка вероятности того или иного заболевания может оказаться ошибочной [6—8].

Для облегчения дифференциальной диагностики можно применять данные, полученные в ходе систематических исследований двух типов. В исследованиях первого типа анализируются клинические проявления

заболевания. Более важны исследования второго типа, в которых изучаются возможные причины появления наблюдаемых симптомов и синдромов (см. главу 7 *Дифференциальная диагностика*). В нашем основном клиническом примере вопрос должен быть сформулирован следующим образом: "В каких случаях остро возникший кашель сочетается с одышкой, какие при этом бывают окончательные диагнозы и какова сравнительная частота установления таких диагнозов?"

Вы уже провели первичную дифференциальную диагностику и определили претестовую вероятность нескольких возможных причин возникновения данного состояния. Как можно использовать дополнительную информацию для установления окончательного диагноза? Анализируя эту информацию, следует сравнить частоту выявления каждого симптома при предполагаемом заболевании и альтернативных вариантах. На этом этапе можно опираться как на собственный опыт, так и на результаты исследований, в которых оценивалась информативность диагностических методов. В нашем примере вы можете использовать данные, полученные в ходе исследования PIOPED (Prospective Investigation of Pulmonary Embolism Diagnosis) [1], посвященного оценке информативности вентиляционно-перфузионной сцинтиграфии легких при диагностике ТЭЛА; приводимые в отчете значения такого показателя, как отношение правдоподобия, позволяют определить, что послетестовая вероятность наличия ТЭЛА равна 90% (см. главу 8 *Диагностические методы*).

В некоторых публикациях приводятся данные, касающиеся как собственно дифференциальной диагностики, так и характеристик диагностических методов. Например, в одном исследовании (259 пожилых больных) сравнивалась информативность анализа крови и костного мозга при диагностике анемии; анемия была диагностирована во всех случаях, а у 94 (36%) больных был выявлен дефицит железа [9]. Хотя основная задача исследования состояла в оценке методов диагностики железодефицитной анемии, была также получена информация о распространенности данного состояния.

В следующих главах будут приведены рекомендации по оценке достоверности двух типов исследований в области диагностики: посвященных анализу клинических проявлений при определенных заболеваниях или изучению информативности диагностических методов. В каждом случае достоверность зависит от ответов на два основных вопроса о структуре исследования: "Правильно ли было организовано включение участников?" и "Применялись ли для постановки окончательного диагноза наиболее информативные методы?". Далее будет показано, каким

образом можно повысить точность диагностики, последовательно применяя результаты исследований каждого из перечисленных типов. Как и при оценке лечения, прогноза и побочных эффектов, наиболее достоверная информация содержится в систематических обзорах всех имеющихся статей, посвященных тому или иному аспекту диагностики (см. главу 10 *Обобщение доказательств*). Чтобы понимать и правильно интерпретировать результаты таких обзоров, необходимо знать принципы оценки отдельных исследований, посвященных диагностике.

Литература

1. *The PIOPED Investigators*. Value of the ventilation/perfusion scan in acute pulmonary embolism. Results of the Prospective Investigation of Pulmonary Embolism Diagnosis (PIOPED). *JAMA* 1990;263:2753–2759.
2. *Sox H.C., Blatt M.A., Higgins M.C., Marton K.I.* Medical Decision Making. Boston: Butterworths: 1988.
3. *Glass R.D.* Diagnosis: A Brief Introduction. Melbourne: Oxford University Press; 1996.
4. *Baroness J.A., Carpenter C.C.J., eds.* Differential Diagnosis. Philadelphia: Lea & Febiger; 1994.
5. *Pauker S.G., Kassirer J.P.* The threshold approach to clinical decision making. *N Engl J Med* 1980;302:1109–1117.
6. *Schmidt H.G., Norman G.R., Boshuizen H.P.* A cognitive perspective on medical expertise: theory and implication. *Acad Med* 1990;65:611–621.
7. *Bordage G.* Elaborated knowledge: a key to successful diagnostic thinking. *Acad Med* 1994;69:883–885.
8. *Regehr G., Norman G.R.* Issues in cognitive psychology: implications for professional education. *Acad Med* 1996;71:988–1001.
9. *Guyatt G.H., Patterson C., AH M., et al.* Diagnosis of iron-deficiency anemia in the elderly. *Am J Med* 1990;88:205–209.

7**ДИФФЕРЕНЦИАЛЬНАЯ
ДИАГНОСТИКА**

У.С. Ричардсон, М. Уилсон, Д. Лиймер,
Г. Гайятт, Д. Кук

В подготовке данной главы принимали также участие члены Рабочей группы по разработке и внедрению принципов доказательной медицины П. Уайер, К.Д. Нэйлор, Д. Крэйг, Л.М. Летельер и В. Мойер

В ЭТОМ РАЗДЕЛЕ

Поиск доказательной информации

Достоверны ли результаты исследования?

Соответствуют ли критерии включения участников задачам исследования? Репрезентативна ли выборка для анализа данной клинической проблемы?

Был ли использован соответствующий диагностический стандарт для постановки окончательного диагноза? Правильно ли был организован диагностический процесс?

Достаточно ли больных, у которых диагноз не был поставлен во время исследования, наблюдались до выяснения причины наблюдаемых изменений?

Каковы результаты исследования?

Какие заболевания рассматривались в процессе дифференциальной диагностики, и какой была вероятность их наличия?

Насколько точно была оценена вероятность наличия того или иного заболевания?

Как можно применить результаты исследования на практике?

Совпадают ли характеристики участников исследования и конкретного больного?

Возможно ли, что с момента публикации полученных доказательств вероятность наличия того или иного заболевания изменилась?

Принятие клинического решения

ВОЗМОЖНАЯ КЛИНИЧЕСКАЯ СИТУАЦИЯ

Отчего 33-летний мужчина жалуется на учащенное сердцебиение?

Представьте себя врачом общей практики. К вам на прием пришел 33-летний мужчина с жалобами на учащенное сердцебиение без перебоев. Приступы возникают несколько раз в сутки с постепенно нарастающей интенсивностью, продолжаются 1—2 мин, не связаны с повседневной активностью и не повлияли на переносимость физических нагрузок. Во время предыдущих визитов вы уже отмечали, что больной склонен к тревожности; вот и сейчас он упоминает о том, что боится развития заболевания сердца. Другие симптомы отсутствуют; больной не принимает никаких лекарственных препаратов; в его личном и семейном анамнезе нет указаний на заболевания сердца. При осмотре частота сердечных сокращений составляет 90 уд/мин; ритм правильный; патологические изменения глаз, щитовидной железы и легких отсутствуют. При аускультации сердца шумы, добавочные тоны, ритм галопа не выслушиваются. Результаты ЭКГ в 12 отведениях в пределах нормы без признаков аритмии или преждевременного возбуждения миокарда.

Вы предполагаете, что в данном случае учащенное сердцебиение скорее всего обусловлено тревожностью, а непосредственной причиной его возникновения является гипервентиляция; возможно, такое сердцебиение относится к признакам приступа паники. Кроме того, несмотря на отсутствие клинических проявлений аритмии или гипертиреоза, вы задаетесь вопросом, достаточно ли часто эти нарушения наблюдаются у подобных больных, и следует ли обратить на них более пристальное внимание. Вероятность наличия феохромоцитомы слишком мала, поэтому этот диагноз вы отвергаете сразу. Таким образом, вы можете перечислить возможные причины учащенного сердцебиения, но хотели бы получить дополнительную информацию о частоте их выявления еще до начала дифференциальной диагностики. Вопрос следует сформулировать следующим образом: "Какова частота выявления разных заболеваний, сопровождающихся приступами учащенного сердцебиения?"

ПОИСК ДОКАЗАТЕЛЬНОЙ ИНФОРМАЦИИ

В вашем кабинете установлен компьютер, соединенный с медицинской библиотекой, в которой имеется база данных MEDLINE (версия CD-ROM). Вы проводите поиск по 3 ключевым терминам: "palpitations" ("учащенное сердцебиение"; 89 ссылок), "differential diagnosis" ("дифференциальный диагноз"; 7039 ссылок) и "cause of causes" ("причина или причины"; 71 848 ссылок). После объединения всех трех терминов в одном запросе остается 17 ссылок. Просмотрев на экране заголовки и краткие рефераты этих публикаций, вы находите статью В.Е. Weber и W.N. Кароог [1], посвященную именно дифференциальной диагностике при учащенном сердцебиении, и открываете полный текст этой статьи.

ДОСТОВЕРНЫ ЛИ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ?

При оценке результатов исследования, в том числе и их достоверности, рекомендуется ответить на целый комплекс вопросов, приведенных в таблице.

Таблица. Рекомендуемый подход к оценке результатов исследования, посвященного дифференциальной диагностике

Достоверны ли результаты исследования?

- Соответствуют ли критерии включения задачам исследования? Репрезентативна ли выборка для анализа данной клинической проблемы?
- Был ли использован соответствующий диагностический стандарт для постановки окончательного диагноза? Правильно ли был организован диагностический процесс?
- Достаточно ли больных, у которых диагноз не был поставлен во время исследования, наблюдались до выяснения причины наблюдаемых изменений?

Каковы результаты исследования?

- Какие заболевания рассматривались в процессе дифференциальной диагностики, и какой была вероятность их наличия?
- Насколько точно была оценена вероятность наличия того или иного заболевания?

Как можно применить результаты исследования на практике?

- Совпадают ли характеристики участников исследования и конкретного больного⁹
 - Возможно ли, что с момента публикации полученных доказательств вероятность наличия того или иного заболевания изменилась⁹
-

СООТВЕТСТВУЮТ ЛИ КРИТЕРИИ ВКЛЮЧЕНИЯ УЧАСТНИКОВ ЗАДАЧАМ ИССЛЕДОВАНИЯ? РЕПРЕЗЕНТАТИВНА ЛИ ВЫБОРКА ДЛЯ АНАЛИЗА ДАННОЙ КЛИНИЧЕСКОЙ ПРОБЛЕМЫ?

Эти вопросы касаются двух взаимосвязанных аспектов: определения клинической проблемы и обеспечения репрезентативности выборки.

Во-первых, какое определение дали исследователи клинической проблеме, подлежащей рассмотрению? Ответ на этот вопрос позволяет выделить популяцию, из которой следовало производить отбор участников. Например, при изучении гематурии в исследование должны быть включены больные с микро- и макрогематурией вне зависимости от наличия симптомов. При изучении же бессимптомной микрогематурии больных с макрогематурией или с другими клиническими проявлениями не следует включать в исследование.

При разных определениях клинической проблемы разной будет и частота выявления заболеваний, которые могут сопровождаться наблюдаемыми изменениями. Так, при включении в исследование больных с макрогематурией и симптомами поражения органов мочевой системы частота выявления острых инфекций мочевых путей будет выше, чем при включении больных только с бессимптомной гематурией. Оценивая достоверность результатов исследования, посвященного дифференциальной диагностике, в первую очередь следует выяснить, насколько точно была сформулирована изучаемая клиническая проблема.

Определив клиническую проблему и соответствующую ей популяцию, исследователи приступают к отбору больных. В идеале все существенные характеристики сформированной выборки и изучаемой популяции должны совпадать, поэтому и распространенность заболеваний или состояний, которые могут служить причиной наблюдаемых изменений, в них должна быть примерно одинаковой. Выборку, удовлетворяющую данным требованиям, называют репрезентативной. Чем репрезентативнее выборка, тем более точна оценка вероятности наличия определенного заболевания или состояния.

Лучшим методом, позволяющим добиться максимальной репрезентативности, следует считать рандомизированное формирование выборки из общей популяции больных с изучаемым набором признаков; однако исследователи редко применяют такой подход. Существуют еще Два метода обеспечения репрезентативности выборки: 1) включение всех больных с соответствующей клинической проблемой, проживающих в определенном регионе, или 2) последовательное включение всех больных с данной проблемой, обращающихся в медицинские учреждения,

на базе которых проводится исследование. Отказ от последовательного включения больных снижает достоверность результатов исследования, поскольку сказывается на распространенности заболеваний, которые могут послужить причиной изучаемых клинических проявлений.

О репрезентативности выборки можно судить и при оценке клинической базы исследования. Больные с одинаковыми жалобами могут обращаться в разные медицинские учреждения. Как правило, в специализированных медицинских центрах доля больных с тяжелыми или более редкими заболеваниями выше, чем в учреждениях первичной медицинской помощи. Так, одно из исследований, посвященное выявлению причин возникновения боли в грудной клетке, проводилось на базе как специализированных медицинских учреждений, так и учреждений первичной медицинской помощи; в первом случае ИБС диагностировали чаще, хотя жалобы и анамнез в обеих группах были сходными [2].

Проводя дальнейшую оценку репрезентативности выборки, следует обратить внимание на методы выявления всех потенциальных участников, на критерии включения и исключения из исследования. Чем полнее выборка по характеристикам, тем больше она отражает общую популяцию больных с изучаемыми признаками и тем выше достоверность полученных результатов. Например, одно из исследований было посвящено изучению колита, вызываемого *Clostridium difficile*, и включало 609 госпитализированных взрослых больных с диареей, у которых имелись результаты анализа кала на наличие цитотоксина [3]. Поскольку больные, которым данный анализ не проводился, в исследование не включались, оценка вероятности наличия *C. difficile* могла оказаться завышенной по сравнению с общей популяцией больных с диареей.

В найденной вами публикации В.Е. Weber и W.N. Кароог [1] дали более широкое определение "сердцебиения", которое включало также перебои в работе сердца. В исследование они включали больных как с впервые возникшими, так и с рецидивирующими приступами. Исследование было проведено на базе отделения неотложной медицинской помощи, терапевтических и хирургических отделений стационара, а также амбулаторной клиники одного университетского медицинского центра, расположенного в среднем по размерам городе на севере США. Из 229 больных, последовательно обратившихся за медицинской помощью с жалобами на сердцебиение, 39 отказались от участия в испытании. Таким образом, в исследовании приняли участие 190 больных (включая 62 больных, обратившихся в отделение неотложной медицинской помощи). Можно предположить, что из исследования не была исключена ни

одна из клинически значимых подгрупп больных, т.е. выборка была вполне репрезентативной.

БЫЛ ЛИ ИСПОЛЬЗОВАН СООТВЕТСТВУЮЩИЙ ДИАГНОСТИЧЕСКИЙ СТАНДАРТ ДЛЯ ПОСТАНОВКИ ОКОНЧАТЕЛЬНОГО ДИАГНОЗА? ПРАВИЛЬНО ЛИ БЫЛ ОРГАНИЗОВАН ДИАГНОСТИЧЕСКИЙ ПРОЦЕСС?

Результаты исследований, посвященных дифференциальной диагностике, достоверны только в том случае, когда окончательный диагноз был поставлен правильно. Для этого исследователи должны разработать и применять четкие критерии, на основании которых у каждого больного можно будет установить окончательный диагноз, т.е. подтвердить или опровергнуть наличие каждого из рассматриваемых заболеваний. Например, опубликованные в медицинской литературе критерии диагностики инфекционного эндокардита включают в себя разные признаки, позволяющие не только выявить это заболевание, но и исключить его наличие [4, 5]. Затем исследователи могут выделить отдельные группы больных с определенными, четко различающимися признаками (за исключением больных, у которых симптомы могут быть обусловлены несколькими этиологическими факторами). Это позволяет очертить круг возможных диагнозов в любых неясных случаях.

Применение диагностических критериев должно подразумевать проведение обследования, достаточно полного для выявления всех возможных причин наблюдаемых изменений. Чем более тщательным будет такое обследование, тем меньше опасность того, что частота выявления того или иного заболевания рассчитана неправильно. Например, в одном ретроспективном исследовании, посвященном изучению частоты развития инсульта при психических расстройствах [6] и включавшем 127 больных, не был проведен тщательный диагностический поиск для выявления всех возможных заболеваний, сопровождавшихся делирием; в 118 случаях причины его возникновения остались неясными. Это ставит под сомнение выводы авторов о вероятности наличия тех или иных психических заболеваний, приводящих к возникновению инсульта.

Четкие, обоснованные диагностические критерии разрабатываются и применяются для того, чтобы предлагаемый метод диагностики был воспроизводимым; окончательно судить о степени воспроизводимости можно на основании формальной оценки совпадения мнений исследователей. Доверие к публикации будет выше, если исследователи сообщали о том, насколько серьезными были разногласия по вопросу об окончательном диагнозе, как это было сделано в исследовании, посвященном изучению причин головокружения [7].

Изучая диагностические критерии, следует помнить, что выявление признаков определенного заболевания не всегда означает установление точной причины недомогания больного. Другими словами, с помощью четких и правильно подобранных диагностических критериев исследователи иногда выявляют у больного сразу несколько заболеваний, наличием которых можно объяснить наблюдаемую симптоматику; в таких случаях остается сомнение в истинной причине ее появления. В наиболее качественных исследованиях при оценке вероятности диагноза должны применяться методы, в той или иной степени позволяющие убедиться в наличии причинно-следственной связи между выявленным заболеванием и изучаемой симптоматикой. Так, в ряде исследований обморочные состояния не связывали с развитием аритмии до тех пор, пока не был проведен анализ данных только о тех больных, у которых оба события регистрировались одновременно [8]. В исследовании, включавшем больных с хроническим кашлем, назначалась терапия, направленная на устранение его предполагаемой причины; положительная ответная реакция на лечение служила лишним подтверждением того, что выявленное заболевание действительно вызвало хронический кашель [9].

Непоследовательное применение способно обесценить самые четкие и полные диагностические критерии. Это вовсе не означает, что у каждого больного обязательно должны применяться все диагностические тесты. Напротив, при решении многих клинических проблем врач собирает подробный анамнез (с учетом конкретных жалоб больного), проводит общий осмотр (обращая особое внимание на отдельные органы и системы, которые в данном случае могут быть вовлечены в патологический процесс) и назначает несколько первичных тестов. Дальнейшее обследование зависит от полученных данных и может быть самым разнообразным. В идеале исследователи также должны проводить всем участникам одинаковое первичное обследование, а последовательность применения других диагностических методов (в зависимости от получаемых данных) должна быть заранее определена. Если у отдельных участников результаты какого-то теста позволят однозначно подтвердить окончательный диагноз, то в этих случаях дальнейшее применение других диагностических методов будет бессмысленным и неэтичным.

Оценить качество диагностического процесса довольно просто, если исследование было проспективным, и в нем использовался заранее определенный алгоритм диагностики. В отсутствие подобной стандартизации обследования судить о его точности труднее. Например, при изучении факторов, способствующих декомпенсации сердечной недостаточности (101 больной) [10], у всех участников исследования собирали

анамнез и проводили общий осмотр, но последующее применение диагностических тестов не было стандартизованным; поэтому судить о том, насколько точно была определена вероятность наличия того или иного заболевания, крайне сложно.

В исследовании В.Е. Weber и W.N. Кароог [1] заранее были разработаны четкие, обоснованные критерии, позволяющие подтвердить наличие каждой из возможных причин учащенного сердцебиения; все эти критерии изложены в приложении к статье вместе с соответствующими ссылками. Обследование участников было проспективным; окончательный диагноз ставился на основании результатов структурированного опроса, проводимого одним из исследователей, и комплексного диагностического обследования (включая сбор анамнеза, общий осмотр и применение разных тестов), порядок которого определял врач во время первого визита. Кроме того, все больные заполняли специальные анкеты, помогающие выявить наличие различных психических расстройств. В большинстве случаев (у 166 из 190 больных) была выполнена ЭКГ, у многих участников применялись и другие методы выявления заболеваний сердца. Аритмию, выявленную при ЭКГ, считали причиной выраженного сердцебиения только в тех случаях, когда эти клинические проявления регистрировались одновременно. Однако о степени согласия исследователей по поводу окончательного диагноза в каждом конкретном случае В.Е. Weber и W.N. Кароог не сообщали.

Таким образом, диагностический процесс был достаточно полным (хотя и не исчерпывающим) для выявления наиболее распространенных причин учащенного сердцебиения. Порядок обследования не был полностью стандартизованным, что могло в некоторой степени сказаться на полученных результатах; однако создается впечатление, что это не оказало существенного влияния на определение вероятности наличия наиболее распространенных причин учащенного сердцебиения, включая психические и сердечно-сосудистые заболевания.

ДОСТАТОЧНО ЛИ БОЛЬНЫХ, У КОТОРЫХ ДИАГНОЗ НЕ БЫЛ ПОСТАВЛЕН ВО ВРЕМЯ ИССЛЕДОВАНИЯ, НАБЛЮДАЛИСЬ ДО ВЫЯСНЕНИЯ ПРИЧИНЫ НАБЛЮДАЕМЫХ ИЗМЕНЕНИЙ?

Даже при последовательном применении четких и тщательно разработанных диагностических критериев иногда не удается выявить причину изучаемых клинических проявлений. Чем больше число таких больных, тем выше риск неверной оценки вероятности наличия того или иного заболевания. Например, в ретроспективном исследовании, посвященном изучению причин головокружения (1194 больных в отоларин-

гологической клинике), диагноз не был установлен в 27% случаев [11]. Это могло привести к тому, что во всей выборке больных вероятность наличия заболеваний, вызывавших головокружение, была определена неверно.

Если во время исследования у части больных диагноз не был поставлен, за такими больными можно наблюдать и в дальнейшем, постоянно выявляя дополнительные данные, необходимые для диагностики, и оценивая прогноз. Чем более длительным и полным будет такое наблюдение, тем выше будет уверенность в благоприятном прогнозе. В течение какого времени следует наблюдать больных, у которых диагноз не был поставлен во время исследования? Однозначного ответа на этот вопрос не существует, но мы полагаем, что при остро возникшей симптоматике, исчезающей без лечения, продолжительность наблюдения должна составлять от 1 до 6 мес, а при хронической симптоматике (рецидивирующей или прогрессирующей) — от 1 года до 5 лет.

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

Вы отмечаете, что В.Е. Weber и W.N. Кароог [1] не выявили причину учащенного сердцебиения у 31 (16,3%) из 190 больных, включенных в исследование. Почти все (96%) участники наблюдались в течение ≥ 1 года. За этот период еще у 1 больного был установлен диагноз (желудочковая экстрасистолия). Ни у одного из больных, диагноз которым так и не был поставлен за время исследования, не развился инсульт, никто из них не умер в период наблюдения.

КАКОВЫ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ?

КАКИЕ ЗАБОЛЕВАНИЯ РАССМАТРИВАЛИСЬ В ПРОЦЕССЕ ДИФФЕРЕНЦИАЛЬНОЙ ДИАГНОСТИКИ, И КАКОМ БЫЛА ВЕРОЯТНОСТЬ ИХ НАЛИЧИЯ?

В отчетах о многих исследованиях, посвященных определению вероятности наличия тех или иных заболеваний, основные результаты представлены в виде таблицы, в которой перечислены поставленные диагнозы, а также указано абсолютное и относительное число больных с этими диагнозами. Иногда у больных выявляют несколько заболева-

ний, которые могут вызывать изучаемые изменения или просто сопутствуют им. В этих случаях авторы, как правило, отдельно указывают основную и сопутствующие диагнозы. Некоторые авторы выделяют больных с множественными причинами наблюдаемых изменений в отдельную группу.

В.Е. Weber и W.N. Кароог [1] представили результаты исследования в виде таблицы, из которой следует, что в общей выборке у 58 (31%) больных причиной учащенного сердцебиения были психические нарушения, у 82 (43%) больных — заболевания сердца, у 5 (2,6%) больных — тиреотоксикоз. Феохромоцитомы не была диагностирована ни у одного из участников. Однако распределение диагнозов различалось в зависимости от особенностей клинической базы. Так, у больных, обращавшихся за медицинской помощью в отделение неотложной медицинской помощи, заболевания сердца выявлялись в 2 раза чаще, чем у больных, наблюдавшихся в амбулаторной клинике.

НАСКОЛЬКО ТОЧНО БЫЛА ОЦЕНЕНА ВЕРОЯТНОСТЬ НАЛИЧИЯ ТОГО ИЛИ ИНОГО ЗАБОЛЕВАНИЯ?

Даже если результаты исследования достоверны, рассчитанная вероятность наличия того или иного заболевания не обязательно совпадает со значением этого показателя в общей популяции. О точности оценки вы можете судить по диапазону доверительного интервала (ДИ), представленного авторами статьи. Если значения ДИ не указаны, их можно рассчитать самостоятельно по следующей формуле:

$$95\% \text{ CI} = P \pm 1,96 \sqrt{P(1-P)} / n,$$

где P — доля больных с каким-либо заболеванием, n — число больных в выборке. Если число больных с рассматриваемым заболеванием не превышает 5, полученные результаты будут неточными; для таких случаев существуют методы приблизительного расчета.

В исследовании В.Е. Weber и W.N. Кароог [1] частота выявления психических расстройств составила 31%. Используя приведенную выше формулу, введем следующие значения: $P = 0,31$, $(1 - P) = 0,69$ и $n = 190$. В результате $\text{ДИ} = 0,31 \pm 0,066$. Таким образом, хотя истинная частота выявления психических заболеваний в качестве причины учащенного сердцебиения скорее всего близка к 31%, она может колебаться от 24,4 до 37,6%.

Ваше мнение о точности определения ДИ будет зависеть от того, каким образом все его значения соотносятся со значениями ПДО и ПЛ

(см. главу 6 *Диагностика*). Если на воображаемой шкале основное значение и все другие значения 95% ДИ находятся по одну сторону от ПДО или ПЛ, полученные результаты достаточно точно отражают вероятность наличия того или иного заболевания и могут быть использованы при планировании дальнейшего обследования или лечения. Напротив, если выбранное значение ПДО или ПЛ попадает в диапазон ДИ, результаты исследования не настолько точны, чтобы сделать однозначный вывод о вероятности наличия определенного заболевания. Достоверные, но не очень точные данные можно использовать, но о возможном влиянии такой неточности следует помнить при планировании обследования и лечения.

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

Определяя вероятность наличия различных причин учащенного сердцебиения, В.Е. Weber и W.N. Кароог не указали значения 95% ДИ [1]. Однако, если вас интересует, каким образом значения 95% ДИ соотносятся с выбранными вами значениями ПДО и ПЛ, можно воспользоваться приведенной выше формулой и рассчитать 95% самостоятельно.

КАК МОЖНО ПРИМЕНИТЬ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ НА ПРАКТИКЕ?

СОВПАДАЮТ ЛИ ХАРАКТЕРИСТИКИ УЧАСТНИКОВ ИССЛЕДОВАНИЯ И КОНКРЕТНОГО БОЛЬНОГО?

Как уже говорилось выше, всегда полезно задаваться вопросом, не отличаются ли условия оказания медицинской помощи и характеристики участников исследования настолько существенно, что вам следует отказаться от применения полученных результатов при лечении конкретного больного [12]. Например, ваш больной мог ранее проживать в местности, где некоторые заболевания, способные стать причиной возникновения изучаемой симптоматики, считаются эндемическими; в этом случае вероятность наличия таких заболеваний будет гораздо выше, чем в исследовании.

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

В исследовании В.Е. Weber и W.N. Кароог [1] были включены 190 больных с учащенным сердцебиением, которые проходили обследование в отделении неотложной медицинской помощи, терапевтических и хирургических отделениях стационара, а также в амбулатории клиники (62 больных) одного университетского медицинского центра, расположенного в среднем по размерам городе на севере США. Вы работаете в отделении неотложной медицинской помощи, поэтому характеристики участников исследования и больных, с которыми вам приходится иметь дело, во многом совпадают. Следовательно, у вашего больного вы можете применять результаты этого исследования для оценки претестовой вероятности наличия заболеваний, вызывающих учащенное сердцебиение.

ВОЗМОЖНО ЛИ, ЧТО С МОМЕНТА ПУБЛИКАЦИИ ПОЛУЧЕННЫХ ДОКАЗАТЕЛЬСТВ ВЕРОЯТНОСТЬ НАЛИЧИЯ ТОГО ИЛИ ИНОГО ЗАБОЛЕВАНИЯ ИЗМЕНИЛАСЬ?

С течением времени данные о распространенности заболеваний устаревают. Многие из них врачи научились успешно лечить, некоторые (например, натуральная оспа) практически полностью исчезли с лица земли [13]. Однако взамен появляются новые заболевания или возникают внезапные эпидемии уже известных болезней. Это может настолько существенно изменить спектр существующих заболеваний и оценку вероятности их наличия, что вполне достоверные результаты ранее проведенных высококачественных исследований станут неприменимыми. Вспомните, как появление ВИЧ-инфекции изменило перечень возможных диагнозов при выявлении причин генерализованной лимфаденопатии, хронической диареи или необъяснимого снижения массы тела.

На частоту выявления и возникновения заболеваний влияют также достижения в области медицины и здравоохранения. Например, в исследованиях по изучению причин лихорадки неясного происхождения применение новых методов диагностики существенно повысило частоту выявления злокачественных заболеваний и снизило долю больных с не установленным диагнозом [14–16]. Использование новых высокоэффективных методов лечения (например, химиотерапии при лейкемии

у детей) повышает не только выживаемость, но и вероятность развития тяжелых осложнений или других заболеваний (например, злокачественных опухолей иной природы через много лет после успешной терапии). В некоторых случаях соотношение возможных диагнозов может измениться в результате широкого применения профилактических мер; так, жесткий контроль за заболеваемостью холерой повысил вероятность наличия других причин возникновения острой диареи.

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

Материалы исследования В.Е. Вебер и W.N. Кароог [1], посвященного изучению причин учащенного сердцебиения, были опубликованы в 1996 г. Само исследование проводилось в течение 8 мес в 1991 г. Вам неизвестно о каких-либо новых обстоятельствах, которые могли бы повлиять на спектр возможных причин данного состояния.

ПРИНЯТИЕ КЛИНИЧЕСКОГО РЕШЕНИЯ

Вернемся к конкретному больному. Рассмотрев возможные причины учащенного сердцебиения, вы предполагаете, что у данного больного оно скорее всего обусловлено острыми приступами тревожности, но не уверены в этом настолько, чтобы исключить другие причины (т.е. претестовая вероятность *основного диагноза* ниже ПЛ). Изучив отчет об исследовании В.Е. Вебер и W.N. Кароог [1], посвященном выявлению причин учащенного сердцебиения, вы решили включить в перечень *альтернативных вариантов* некоторые аритмии (поскольку они довольно широко распространены, представляют опасность для жизни и здоровья, хорошо поддаются терапии) и гипертиреоз (он менее распространен, но достаточно опасен и вполне излечим). Чтобы исключить эти заболевания, вы решаете назначить вашему больному дополнительное обследование (поскольку претестовая вероятность наличия данных заболеваний ниже ПЛ). И наконец, поскольку феохромоцитомы не была выявлена ни у одного из 190 участников исследования В.Е. Вебер и W.N. Кароог [1], и ее клинические признаки отсутствуют у вашего больного, этот диагноз вы отнесли к категории - "*другие гипотезы*" {те претестовая веро-

ятность наличия феохромоцитомы ниже ПДО) и решили временно воздержаться от применения специфических тестов

Литература

- 1 *Weber BE, Kapoor WN* Evaluation and outcomes of patients with palpitations *Am J Med* 1996; 100: 138–148
- 2 *Sox HC, Hickam DH, Marlon KI, et al* Using the patient's history to estimate the probability of coronary artery disease: a comparison of primary care and referral practices *Am J Med* 1990; 89: 7–14
- 3 *Kati DA, Bates DW, Rittenberg E, et al* Predicting Clostridium difficile stool cytotoxin results in hospitalized patients with diarrhea *J Gen Intern Med* 1997; 12: 57–62
- 4 *Reyn CF, Levy BS, Arbeit RD, Fnedland G, Crumacker CS* Infective endocarditis: an analysis based on strict case definitions *Ann Intern Med* 1981; 94: 505–517
- 5 *Durack DT, Lukes AS, Bright DK, and the Duke Endocarditis Service* New criteria for diagnosis of infective endocarditis: utilization of specific echocardiographic findings *Am J Med* 1994; 96: 200–209
- 6 *Benbadis SR, Sila CA, Cristea RL* Mental status changes and stroke *J Gen Intern Med* 1994; 9: 485–487
- 7 *Kroenke K, Lucas CA, Rosenberg ML, et al* Causes of persistent dizziness: a prospective study of 100 patients in ambulatory care *Ann Intern Med* 1992; 117: 898–904
- 8 *Kapoor WN* Evaluation and outcome of patients with syncope *Medicine* 1990; 69: 160–175
- 9 *Pratter MR, Barter T, Aken S, et al* An algorithmic approach to chronic cough *Ann Intern Med* 1993; 119: 977–983
- 10 *Ghali JK, Kadakia S, Cooper R, Ferhaz J* Precipitating factors leading to decompensation of heart failure: traits among urban blacks *Arch Intern Med* 1988; 148: 2013–2016
- 11 *Katsarkas A* Dizziness in aging—a retrospective study of 1194 cases *Otolaryngol Head Neck Surg* 1994; 110: 296–301
- 12 *Glasziou P, Guyatt GH, Dans AL, Dans LF, Straus SE, Sackett DL* Applying the results of trials and systematic reviews to individual patients [editorial] *ACP Journal Club* 1998; 129: A15–A16
- 13 *Barquet N, Domingo P* Smallpox: the triumph over the most terrible of the ministers of death *Ann Intern Med* 1997; 127: 635–642
- 14 *Petersdorf RG, Beeson PB* Fever of unexplained origin: report on 100 cases *Medicine* 1961; 40: 1–30
- 15 *Larson EB, Featherstone HJ, Petersdorf RG* Fever of undetermined origin: diagnosis and follow up of 105 cases 1970–1980 *Medicine* 1982; 61: 269–292
- 16 *Knockaert DC, Vanneste LJ, Vanneste SB, Bobbaen HJ* Fever of unknown origin in the 1980s: an update of the diagnostic spectrum *Arch Intern Med* 1992; 152: 51–55

8

ДИАГНОСТИЧЕСКИЕ МЕТОДЫ

Р. Йешке, Г. Гайятт, Д. Лиймер

В подготовке данной главы принимали также участие члены Рабочей группы по разработке и внедрению принципов доказательной медицины П Уайер, В Мойер, Д Кук, Дж Крэйг, Л М Летельер, Дж Уильяме, КД Нэйлор, УС Ричардсон, М Уилсон и Дж Нишикава

В ЭТОМ РАЗДЕЛЕ

Поиск доказательной информации

Достоверны ли результаты исследования?

Всегда ли точны результаты диагностических тестов"

Применялся ли в группах лечения и контроля слепой метод при сравнении результатов изучаемого диагностического метода и "золотого стандарта""⁷

Повлияли ли результаты изучаемого диагностического теста на решение о необходимости применения "золотого стандарта"?

КАКОВЫ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ?

Какие значения отношения правдоподобия соответствуют возможному диапазону результатов изучаемого диагностического метода?

Как можно применить результаты исследования на практике?

Настолько ли убедительны данные о воспроизводимости результатов изучаемого теста и интерпретация этих данных, чтобы их можно было применить в условиях вашей практики?

Возможно ли применение полученных результатов у конкретного больного?

Окажут ли полученные результаты влияние на вашу тактику ведения больных?

Улучшится ли качество оказания медицинской помощи вследствие применения изучаемого диагностического метода?

Принятие клинического решения

ВОЗМОЖНАЯ КЛИНИЧЕСКАЯ СИТУАЦИЯ

Насколько информативны результаты компьютерной томографии при предполагаемом аппендиците?

В отделение неотложной медицинской помощи обратилась 32-летняя женщина с жалобами на боль в правой подвздошной области. Больная работает в компании, продающей компьютерные Интернет-программы. Она не замужем, ведет половую жизнь (за последний год у нее были три партнера), последняя менструация закончилась 3 нед назад. Со вчерашнего дня ее самочувствие внезапно ухудшилось, пропал аппетит. За последние несколько часов боль в животе усилилась, был озноб, однако температуру тела больная не измеряла. Патологических выделений из влагалища не было. Женщина решила обратиться за медицинской помощью, поскольку испугалась, что такая боль может быть проявлением какого-нибудь тяжелого заболевания.

При общем осмотре вы отмечаете у больной состояние средней тяжести; температура тела составляет 38,2°C. При пальпации в правой подвздошной области выявляются болезненность и напряженность мышц передней брюшной стенки; симптомы раздражения брюшины сомнительные. Смещение шейки матки безболезненно; выделений из нее нет. Число лейкоцитов составляет 11 000 в 1 мкл крови. Вы считаете, что причиной боли могут быть аппендицит, воспалительные заболевания органов малого таза или внематочная беременность. Пока вы размышляете, стоит ли сразу перевести больную в хирургическое отделение или лучше сначала проконсультировать ее у гинеколога, возле вас останавливается рентгенолог, возвращающийся после выполнения экстренной ангиопульмонографии (АПГ). Вы описываете ему ситуацию, и он говорит, что применение компьютерной томографии (КТ) у подобных больных позволяет избежать ненужной лапаротомии в 15% случаев, а госпитализации — в 20% случаев. Он советует "обязательно прочитать весьма интересную статью в журнале *New England Journal of Medicine*", хотя не помнит точно ни названия этой статьи, ни деталей проведенного исследования (в частности, применявшихся диагностических методов или полученных результатов)

Состояние больной стабильное и не внушает опасений; в отделении неотложной медицинской помощи наконец наступило относительное затишье. Ваш коллега готов отпустить вас на 30 мин, и вы решаете посвятить это время поиску и изучению статьи, порекомендованной рентгенологом.

ПОИСК ДОКАЗАТЕЛЬНОЙ ИНФОРМАЦИИ

Вы поднимаетесь в библиотеку, садитесь за компьютер и входите в базу данных PubMed. На ее странице, посвященной клиническим вопросам (www.ncbi.nlm.nih.gov/entrez/query/static/clinical.html), вводите ключевые слова "diagnosis" ("диагноз") и "specificity" ("специфичность"), позволяющие провести поиск исследований, посвященных диагностике. Введя также ключевые слова "СТ" ("КТ") и "appendicitis" ("аппендицит"), вы получаете список из 39 ссылок. После этого вы ограничиваете условия поиска (англоязычные статьи, рефераты которых были опубликованы в течение предыдущих 5 лет); в списке остается 18 ссылок, включая рефераты 2 описательных обзоров, 4 ретроспективных исследований, 2 исследований, посвященных выявлению специфических КТ-признаков аппендицита, и 2 исследований, посвященных обследованию отдельных групп больных. В 2 рефератах отсутствует количественная информация об информативности КТ, 1 реферат опубликован в журнале, который не выписывает больничная библиотека. В оставшихся 5 рефератах сообщается о высокой диагностической точности КТ. Название наиболее поздней статьи указывает на то, что в данном исследовании оценивалась информативность спиральной КТ при проведении дифференциальной диагностики между аппендицитом и острыми гинекологическими заболеваниями; при этом клинические характеристики участников были схожими с характеристиками вашей больной [1]. Статья, опубликованная в *New England Journal of Medicine*, появилась гораздо раньше; в ней преимущественно оценивается эффективность затрат при выполнении КТ, а не информативность этого метода диагностики. Поэтому вы решаете изучить материалы более позднего исследования, ценность которого в вашей ситуации должна быть выше.

В процессе последующего обсуждения достоверности и применимости результатов исследований, посвященных изучению информативности диагностических тестов, мы не раз будем вспоминать два клинических примера с использованием вентиляционно-перфузионной сцинтиграфии легких при предполагаемой тромбоземболии легочной артерии (ТЭЛА; см. главу 6 *Диагностика*) и КТ при предполагаемом аппендиците. При оценке исследований, посвященных интерпретации результатов диагностических тестов, рекомендуется ответить на целый комплекс вопросов, приведенных в табл. 1.

Таблица 1. Рекомендуемый подход к оценке исследования, посвященного интерпретации результатов различных диагностических тестов

Достоверны ли результаты исследования?

- Всегда ли точны результаты диагностических тестов?
- Применялся ли в группах лечения и контроля слепой метод при сравнении результатов изучаемого диагностического теста и "золотого стандарта"?
- Повлияли ли результаты изучаемого диагностического теста на решение о необходимости применения "золотого стандарта"?

Каковы результаты исследования?

- Какие значения отношения правдоподобия соответствуют возможному диапазону результатов изучаемого диагностического теста?

Как можно применить результаты исследования на практике?

- Настолько ли убедительны данные о воспроизводимости результатов изучаемого теста и интерпретация этих данных, чтобы их можно было применить в условиях вашей практики?
 - Возможно ли применение полученных результатов у конкретного больного?
 - Окажут ли полученные результаты влияние на вашу тактику ведения больных?
 - Улучшится ли качество оказания медицинской помощи вследствие применения изучаемого диагностического метода?
-

ДОСТОВЕРНЫ ЛИ ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ?

ВСЕГДА ЛИ ТОЧНЫ РЕЗУЛЬТАТЫ ДИАГНОСТИЧЕСКИХ ТЕСТОВ?

Применение диагностического теста должно помочь различить заболевания со сходной клинической картиной (которые без использования такого теста могут остаться не выявленными). Почти любой диагностический тест должен позволить отличать здоровых людей от больных; однако такая характеристика метода не отражает его клиническую ценность. Об этом можно судить только по результатам исследований, условия проведения которых были максимально приближены к клиническим. Существует и другой подход, позволяющий определить клиническую значимость результатов диагностического теста; для этого следует вернуться к рисунку в главе 6 *Диагностика*. Так, обследование оправдано лишь в тех случаях, когда клиническая картина не позволяет ни отвергнуть рассматриваемый диагноз, ни начать лечение безотлагательно (т.е. вероятность этого диагноза находится между значениями ПДО и ПЛ).

Важность правильного выбора популяции, в которой следует применять определенный диагностический тест, можно продемонстрировать с помощью следующего примера. В одном из исследований [2] определяли уровень карциноэмбрионального антигена; он оказался повышенным у 35 из 36 больных раком толстой кишки в поздней стадии, но был существенно ниже при других заболеваниях. Полученные результаты указывали на то, что такой радиоиммунный анализ может быть полезным при диагностике рака толстой кишки (или даже в качестве скринингового метода). Однако в дальнейших исследованиях, включавших больных раком толстой кишки в более ранних стадиях и больных со злокачественными опухолями иной локализации или другими заболеваниями желудочно-кишечного тракта (т.е. с заболеваниями разной природы, но схожими по клиническим проявлениям), точность данного метода не подтвердилась; поэтому врачи отказались от его использования при диагностике рака и при скрининговых обследованиях. Определение уровня карциноэмбрионального антигена оказалось оправданным только как часть комплексного обследования, проводимого в ходе наблюдения за больными с подтвержденным раком толстой кишки [3].

J.G. Lijmer et al. [4] провели эмпирическую оценку вероятности возникновения систематической ошибки в исследованиях, посвященных диагностике, определяя зависимость между структурой исследования и показателями информативности диагностического метода. Было выявлено существенное завышение информативности изучаемых методов (т.е. возможности подтвердить или отвергнуть наличие определенного заболевания) в тех случаях, когда в контрольную группу включались только лица без этого заболевания (0111=3,0 при 95% ДИ от 2,0 до 4,5).

Более правильно было организовано исследование PIOPED, продемонстрировавшее информативность применения вентиляционно-перфузионной сцинтиграфии легких при диагностике ТЭЛА, поскольку в это исследование включались самые разные больные (с высокой, средней и низкой вероятностью наличия этого заболевания) [5]. В другом исследовании последовательное включение небеременных женщин с предполагаемым острым аппендицитом или острыми гинекологическими заболеваниями, обратившихся в отделение неотложной медицинской помощи крупной многопрофильной больницы, обеспечило формирование репрезентативной выборки для оценки информативности КТ при дифференциальной диагностике данных состояний.

ПРИМЕНЯЛСЯ ЛИ В ГРУППАХ ЛЕЧЕНИЯ И КОНТРОЛЯ СЛЕПОЙ МЕТОД ПРИ СРАВНЕНИИ РЕЗУЛЬТАТОВ ИЗУЧАЕМОГО ДИАГНОСТИЧЕСКОГО МЕТОДА И "ЗОЛОТОГО СТАНДАРТА"?

О точности изучаемого диагностического теста лучше всего судить при сравнении его результатов с результатами "золотого стандарта" — стандартного диагностического метода (например, гистологического исследования биоптатов, хирургического вмешательства или длительного наблюдения), который должен применяться у всех больных наравне с изучаемым тестом [6]. В исследовании ROPED в качестве такого стандарта использовали АПГ, информативность которой настолько высока, насколько это возможно в условиях прижизненного обследования.

Лишь в том случае, когда изучаемый диагностический тест является частью "золотого стандарта", результаты исследования могут оказаться недостоверными. Например, в одном исследовании [7] оценивалась информативность одновременного определения уровня амилазы в сыворотке крови и моче при диагностике панкреатита. В качестве "золотого стандарта" исследователи применяли результаты комплексного обследования, включавшего и указанные тесты. При таком подходе оценка информативности изучаемых методов могла оказаться завышенной. Таким образом, изучаемый диагностический тест и "золотой стандарт" должны быть независимыми методами.

Если при чтении статьи вы сомневаетесь в точности использованного "золотого стандарта" (помня при этом, что идеальных диагностических методов не существует), полученные результаты нельзя считать достоверными. Если вы полагаете, что "золотой стандарт" был выбран правильно, следует задать вопрос, применялся ли слепой метод при сравнении его результатов и результатов изучаемого диагностического теста (при использовании данного метода лица, оценивающие результаты одного теста, не должны знать о результатах другого). Опыт клинической работы указывает на необходимость соблюдения этого правила. Так, увидев узелковое образование на компьютерной томограмме легких, врач обнаружит его и на рентгенограммах (хотя ранее оно не было диагностировано); имея результаты эхокардиографии, можно услышать и те шумы, которых раньше не находили. В уже упоминавшемся исследовании J.G. Lijmeret al. [4] была продемонстрирована пусть небольшая, но реальная возможность возникновения систематической ошибки, связанной с отказом от применения слепого метода при оценке полученных результатов (0111=1,3 При 95% ДИ от 1,0 до 1,9).

Чем выше вероятность того, что знание результатов "золотого стандарта" повлияет на интерпретацию результатов изучаемого диагностического теста (и наоборот), тем большее значение имеет использование слепого метода. В отчете об исследовании PLOPED не было указано, что при оценке результатов диагностических тестов применялся слепой метод. Однако из этого отчета следует, что исследователи приложили максимум усилий для получения воспроизводимых результатов; поэтому можно предположить, что их оценка была независимой. Переписка с одним из авторов подтвердила данное предположение.

В исследовании, посвященном определению информативности КТ при диагностике аппендицита, в качестве "золотого стандарта" использовалось хирургическое вмешательство с последующим гистологическим исследованием операционного материала. В тех случаях, когда от операции было решено воздержаться, за "золотой стандарт" принимали результаты последующего наблюдения продолжительностью не менее 2 мес (данные получены в процессе амбулаторного обследования или опросов по телефону). В отчете об исследовании не сообщалось, в какой степени врачи были осведомлены о результатах КТ. В тех случаях, когда окончательный диагноз ставился после длительного наблюдения, такая осведомленность могла способствовать завышенной оценке информативности КТ.

ПОВЛИЯЛИ ЛИ РЕЗУЛЬТАТЫ ИЗУЧАЕМОГО ДИАГНОСТИЧЕСКОГО МЕТОДА НА РЕШЕНИЕ О НЕОБХОДИМОСТИ ПРИМЕНЕНИЯ "ЗОЛОТОГО СТАНДАРТА"?

Оценка характеристик диагностического теста будет неверной, если на основании его результатов выносилось решение о необходимости подтверждения диагноза с помощью "золотого стандарта". В этих случаях можно говорить о наличии систематической ошибки, связанной с подтверждением диагноза [8, 9] или установлением окончательного диагноза [10, 11]. Подобная систематическая ошибка возникает, например, в ходе обследования больных с предполагаемой ишемической болезнью сердца (ИБС), когда при положительном результате пробы с физической нагрузкой коронарную ангиографию ("золотой стандарт") выполняют чаще, чем при отрицательном результате этой пробы. В своем исследовании J.G. Lijmer et al. [4] отметили высокую вероятность возникновения *систематической ошибки, связанной с порядком обследования больных* (применением различных стандартных диагностических методов при положительных и отрицательных результатах изучаемого теста).

Наличие такой систематической ошибки не исключено в исследовании PLOPED. В тех случаях, когда оценка результатов вентиляционно-

перфузионной сцинтиграфии легких была более благоприятной для больного ("нормальные или почти нормальные", "низкая вероятность наличия ТЭЛА"), АПГ выполнялась реже (69%), чем при менее благоприятных оценках результатов сцинтиграфии (92%). Это не удивительно, поскольку при низкой вероятности наличия ТЭЛА врачи обычно избегают назначать исследование с высоким риском развития тяжелых осложнений.

Если бы авторы исследования **PIOPED** ограничились применением указанных методов, читатели могли сделать вывод о возможном наличии достаточно большой систематической ошибки, обусловленной различной частотой выполнения АПГ при разных результатах сцинтиграфии легких (свидетельствующих о низкой или высокой вероятности наличия ТЭЛА). Однако у 150 больных с "низкой вероятностью наличия ТЭЛА", а также с "нормальными или почти нормальными" результатами сцинтиграфии легких (в 136 случаях АПГ не выполняли, в 14 случаях ее результаты оказались сомнительными) применялся альтернативный "золотой стандарт": диагноз ТЭЛА исключали в тех случаях, когда состояние больных не ухудшалось в отсутствие терапии антикоагулянтами. Таких больных наблюдали в течение 1 года. За весь период наблюдения клинических признаков ТЭЛА не было отмечено ни у одного из этих больных, что позволило сделать заключение об отсутствии клинически значимой ТЭЛА (если под клинически значимой понимается ТЭЛА, при которой необходимо применять антикоагулянты для предотвращения развития осложнений) на момент проведения вентилиционно-перфузионной сцинтиграфии легких.

В исследовании, посвященном определению информативности КТ, "золотой стандарт" применялся у всех больных. При этом не исключено, что на его выбор (операция или наблюдение) повлияли результаты изучаемого диагностического теста. Как упоминалось выше, оценка информативности КТ при постановке окончательного диагноза оказалась существенно завышенной.

КАКОВЫ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ?

КАКИЕ ЗНАЧЕНИЯ ОТНОШЕНИЯ ПРАВДОПОДОБИЯ СООТВЕТСТВУЮТ ВОЗМОЖНОМУ ДИАПАЗОНУ РЕЗУЛЬТАТОВ ИЗУЧАЕМОГО ДИАГНОСТИЧЕСКОГО ТЕСТА?

Диагностический процесс начинается, когда больной с определенными субъективными и объективными клиническими проявлениями

обращается за медицинской помощью. Представьте себе 2 больных (78-летняя женщина через 10 дней после хирургического вмешательства и 28-летний мужчина с выраженной тревожностью) с болью в грудной клетке и одышкой, но без признаков пневмонии, обструкции дыхательных путей или сердечной недостаточности; врач предполагает у них наличие ТЭЛА. Исходная (претестовая) вероятность такого диагноза у данных больных различна: у пожилой женщины она высокая, у молодого мужчины низкая. Даже при получении одинаковых результатов вентиляционно-перфузионной сцинтиграфии легких (например, "промежуточная вероятность наличия ТЭЛА") тактика лечения у них также будет разной. Пожилой больной показано немедленное применение гепарина, молодому мужчине — дополнительное обследование.

Рассмотрение этих клинических примеров позволяет сделать два вывода. Во-первых, результаты сцинтиграфии легких не всегда однозначно свидетельствуют о наличии ТЭЛА. В данных примерах они лишь изменили вероятность предполагаемого диагноза, т.е. послетестовая вероятность оказалась отличной от претестовой. Степень и направленность этого изменения (от претестовой до послетестовой вероятности) зависят от характеристик того или иного диагностического теста, из которых наибольшее значение имеет отношение правдоподобия (ОП).

В табл. 2 представлены результаты исследования PUPED, в котором участвовал 251 больной с ТЭЛА, подтвержденной при АПГ, и 630 больных, у которых этот диагноз был отвергнут после проведения АПГ или в ходе наблюдения. При анализе данных, полученных в ходе этого исследования, были использованы 4 варианта оценки результатов вентиляционно-перфузионной сцинтиграфии легких: "высокая вероятность наличия ТЭЛА", "промежуточная вероятность наличия ТЭЛА", "низкая вероятность наличия ТЭЛА" и "нормальные или почти нормальные". Часто ли "высокая вероятность наличия ТЭЛА" может выявляться у больных с подтвержденным позднее диагнозом? Из табл. 2 следует, что такие результаты сцинтиграфии были получены у 102 (40,6%, или 0,406) из 251 больных. Часто ли "высокая вероятность наличия ТЭЛА" может выявляться у больных с отвергнутым позднее диагнозом? Из той же таблицы следует, что такие результаты сцинтиграфии были получены у 14 (2,2%, или 0,022) из 630 больных с предполагаемым вначале, но затем отвергнутым диагнозом ТЭЛА. Соотношение этих двух показателей называют ОП; для такого результата сцинтиграфии, как "высокая вероятность наличия ТЭЛА", $ОП = 0,406 / 0,022 = 18,3$. Другими словами, вероятность получения подобного результата в 18,3 раза выше при наличии ТЭЛА, чем в ее отсутствие.

Таблица 2. Характеристики вентиляционно-перфузионной сцинтиграфии легких при диагностике ТЭЛА (по данным исследования PLOPED)

Оценка результатов сцинтиграфии	ТЭЛА				ОП
	имеется		отсутствует		
	число больных	доля больных	число больных	доля больных	
"Высокая вероятность наличия ТЭЛА"		102/251 = 0,406	14	14/630 = 0,022	18,3
"Промежуточная вероятность наличия ТЭЛА"	105	105/251 = 0,418	217	217/630 = 0,344	1,20
"Низкая вероятность наличия ТЭЛА"		39/251 = 0,155	273	= 0,433	
"Нормальные или почти нормальные"	5	5/251 = 0,020	126	= 0,200	
Всего		251			

Аналогичным образом можно рассчитать значение ОП для любых результатов диагностического теста. Такой расчет позволяет ответить сразу на 2 вопроса. Во-первых, как часто можно получить определенный результат теста при реальном наличии предполагаемого заболевания? Во-вторых, как часто этот же результат теста может быть получен в отсутствие данного заболевания? Для такого результата вентиляционно-перфузионной сцинтиграфии легких, как "низкая вероятность наличия ТЭЛА", ответ на первый вопрос будет выглядеть следующим образом: $39 / 251 = 0,155$; на второй: $273 / 630 = 0,433$; при этом $ОП = 0,155 / 0,433 = 0,36$. В табл. 3 приводятся значения ОП и для других результатов сцинтиграфии, применявшейся для диагностики ТЭЛА.

Что означают все эти цифры? Величина ОП указывает на соотношение претестовой и послетестовой вероятности наличия определенного заболевания, т.е. на то, в какой степени претестовая вероятность повысится или понизится после получения результатов диагностического теста. Если $ОП=1,0$, послетестовая вероятность равна претестовой. Значение $ОП>1,0$ свидетельствует о повышении вероятности наличия предполагаемого заболевания; чем больше значение ОП, тем выше эта вероятность. И наоборот, при $ОП<1,0$ вероятность наличия этого заболевания снижается; чем меньше значение ОП, тем ниже эта вероятность.

Какое значение ОП можно считать "большим" или "маленьким"? Обычно понимание этого приходит с опытом, однако можно использовать следующее правило приблизительной оценки ОП:

- Значения $ОП > 10$ или $< 0,1$ указывают на существенные, часто решающие различия между пре- и послетестовой вероятностью.
- Значения ОП от 5 до 10 и от 0,1 до 0,2 указывают на умеренно выраженные различия между пре- и послетестовой вероятностью.
- Значения ОП от 2 до 5 и от 0,5 до 0,2 указывают на незначительные (но иногда клинически важные) различия между пре- и послетестовой вероятностью.
- Значения ОП от 1 до 2 и от 0,5 до 1 указывают на небольшое (и редко имеющее клиническое значение) изменение претестовой вероятности.

Каким образом можно рассчитать послетестовую вероятность, зная величину и клиническую значимость ОП? Оперировать значениями правдоподобия напрямую (как оперируют значениями вероятности или процентами) нельзя. Сначала нужно выразить претестовую вероятность в виде шансов, умножить полученный результат на значение ОП, а затем перевести послетестовое значение шансов в послетестовую вероятность. Эти расчеты не очень сложны (см. главу 4 раздела II *Лечение и пониманиерезультатовисследования:измерениекорреляции*), но они однообразны и отнимают время. К счастью, есть более простой способ.

Номограмма, предложенная Т.Д. Fagan [12], позволяет легко определить изменение претестовой вероятности после применения диагностического теста (см. рисунок). В левой колонке номограммы приведены значения претестовой вероятности, в средней колонке — значения ОП, в правой колонке — значения послетестовой вероятности. Чтобы определить послетестовую вероятность, нужно продолжить прямую линию, соединяющую значения претестовой вероятности и ОП.

Вернемся к примеру с пожилой женщиной, у которой предполагается возникновение ТЭЛА после выполнения хирургического вмешательства на органах брюшной полости. Большинство врачей согласятся с тем, что в этом случае претестовая вероятность наличия ТЭЛА достаточно высока и составляет около 70%. Допустим, что результаты вентиляционно-перфузионной сцинтиграфии также свидетельствуют о "высокой вероятности наличия ТЭЛА". Пользуясь номограммой (см. рисунок), послетестовую вероятность можно определить, проведя прямую линию

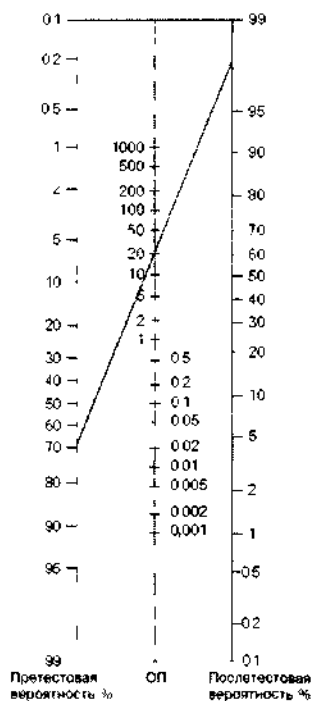


Рисунок. Номограмма, предложенная Т. J Fagan.

между значением претестовой вероятности (70%) и ОП (которое в данной ситуации равно 18,3), а затем продолжив ее до пересечения с третьей колонкой. В данном случае послетестовая вероятность превышает 97%. Если бы результаты вентиляционно-перфузионной сцинтиграфии свидетельствовали о "промежуточной вероятности развития ТЭЛА" (ОП=1,2), послетестовая вероятность была бы гораздо ниже, составляя 74%. При "нормальных или почти нормальных" результатах сцинтиграфии она составляла бы 19%.

Претестовая вероятность оценивается приблизительно. Как уже подчеркивалось, существенную помощь в этом могут оказать публикации, посвященные дифференциальной диагностике (см. главу 6 *Диагностика*) Некоторые неточности в расчетах послетестовая вероятности могут быть обусловлены возможным разбросом значений претестовой вероятности. Например, если у пожилой женщины претестовая вероятность

—

наличия ТЭЛА будет равна 60 или 80%, значения послетестовой вероятности будут иными (табл. 3).

Обратимся ко второму примеру. У 28-летнего мужчины с выраженной тревожностью претестовая вероятность наличия ТЭЛА составляет примерно 20%. Снова воспользуемся номограммой. При результатах скинтиграфии, свидетельствующих о "высокой вероятности наличия

Таблица 3. Значения претестовой вероятности, ОП результатов вентиляционно-перфузионной скинтиграфии легких и послетестовой вероятности у 2 больных с предполагаемой ТЭЛА

Претестовая вероятность, % (диапазон значений*)	Оценка результатов скинтиграфии легких (ОП)	Послетестовая вероятность, % (диапазон значений*)
Женщина в возрасте 78 лет, у которой после хирургического вмешательства на органах брюшной полости внезапно появилась одышка		
70 (от 60 до 80)	"Высокая вероятность наличия ТЭЛА" (18,3)	97 (от 96 до 99)
70 (от 60 до 80)	"Промежуточная вероятность наличия ТЭЛА" (1,2)	74 (от 64 до 83)
70 (от 60 до 80)	"Низкая вероятность наличия ТЭЛА" (0,36)	46 (от 35 до 59)
70 (от 60 до 80)	"Нормальные или почти нормальные" (0,1)	19 (от 1 до 29)
Мужчина в возрасте 28 лет с одышкой и нетипичной болью в грудной клетке		
20 (от 10 до 30)	"Высокая вероятность наличия ТЭЛА" (18,3)	82 (от 67 до 89)
20 (от 10 до 30)	"Промежуточная вероятность наличия ТЭЛА" (1,2)	23 (от 12 до 34)
20 (от 10 до 30)	"Низкая вероятность наличия ТЭЛА" (0,36)	8 (от 4 до 12)
20 (от 10 до 30)	"Нормальные или почти нормальные"	2 (от 1 до 4)

III)

Примечание. * — в скобках указано, в каких пределах может варьировать показатель; так, если наиболее подходящее значение претестовой вероятности составит 70%, значения, равные 60 или 80%, также допустимы.

ТЭЛА", послетестовая вероятность составит 82% (см. рисунок); при "средней вероятности наличия ТЭЛА" — 23%, а при "нормальных или почти нормальных" результатах сцинтиграфии — 2%. В табл. 3 отражены значения претестовой вероятности (20% при возможном диапазоне от 10 до 30%), ОП и послетестовой вероятности, соответствующие каждому из четырех вариантов оценки результатов сцинтиграфии.

В найденном вами исследовании, включавшем женщин с предполагаемым аппендицитом, результаты КТ оказались положительными у всех 32 больных с подтвержденным позднее диагнозом и отрицательными у 66 из 68 больных с отвергнутым позднее диагнозом. Таким образом, для отрицательных результатов теста ОП=0, а для положительных ОП=34. Полученные данные свидетельствуют о высокой диагностической информативности теста. Это означает, что отрицательные результаты КТ позволяют с уверенностью исключить наличие аппендицита, а положительные результаты почти наверняка указывают на наличие этого заболевания.

Получив представление о клинической значимости ОП, вы захотите узнать, где можно найти значения этого показателя для тестов, постоянно применяемых в вашей практике. В журнале *JAMA* публикуются систематические обзоры серии Rational Clinical Examination [13], в которых анализируются диагностические характеристики данных анамнеза и общего осмотра. Достаточно большое количество информации о возможностях современных диагностических тестов можно найти в книге E.R. Black et al. [14].

Чувствительность и специфичность. Если вы дочитали главу до этого места, то уже осознали важность правильной интерпретации результатов диагностических тестов. Полезно также понимать смысл еще двух терминов, широко применяемых для характеристики диагностических тестов: чувствительность и специфичность.

Читатель должен был заметить, что при обсуждении ОП мы не использовали такие понятия, как "нормальные результаты теста" или "наличие патологических изменений". При анализе данных, полученных в ходе исследования PLOPED, мы применяли 4 варианта оценки результатов вентиляционно-перфузионной сцинтиграфии легких и соответствующие им значения ОП. Однако сами исследователи представили полученные данные в другой форме, опираясь на традиционные (но менее полезные) понятия чувствительности и специфичности методов Диагностики.

Под *чувствительностью* подразумевается доля больных с искомым заболеванием, у которых результат диагностического теста оказался

положительным, а под *специфичностью* — доля больных без искомого заболевания, у которых результат диагностического теста оказался отрицательным. Чтобы понять сущность данных терминов, все результаты определенного диагностического теста следует подразделить на 2 категории: "нормальные" и свидетельствующие о "наличии патологических изменений". Другими словами, необходимо построить таблицу с двумя столбцами и двумя строками (формат 2 x 2); примером такой таблицы будет табл. 4. Следует отметить, что любую таблицу формата 4 x 2 (например, табл. 3) можно трансформировать в 3 таблицы формата 2 x 2 (см. табл. 4, а также табл. 5 и 6) в зависимости от того, какие результаты теста мы будем оценивать как "нормальные" или как "наличие патологических изменений" (либо от того, что мы понимаем под отрицательными и положительными результатами теста). Допустим, что результатами, свидетельствующими о "наличии патологических изменений" (или положительными результатами), мы будем считать только те результаты вентилиционно-перфузионной сцинтиграфии легких, которые ранее мы отнесли к категории "высокая вероятность наличия ТЭЛА".

Табл. 5 также представляет собой таблицу формата 2 x 2, в которой сравниваются результаты вентилиционно-перфузионной сцинтиграфии

Таблица 4. Сравнение результатов изучаемого диагностического теста и "золотого стандарта" с помощью таблицы формата 2 x 2

Результаты изучаемого теста	Результаты "золотого стандарта"	
	заболевание имеется	заболевание отсутствует
Заболевание имеется	Истинно положительные (a)	Ложноположительные (b)
Заболевание отсутствует	Ложноотрицательные (c)	Истинно отрицательные (d)

Примечание

$$\text{Чувствительность} = \frac{a}{a + c}$$

$$\text{Специфичность} = \frac{d}{b + d}$$

$$\text{Отношение правдоподобия положительного результата теста (ОП+)} = \frac{\text{чувствительность}}{1 - \text{специфичность}} = \frac{[a/(a + c)]}{[b/(b + d)]}$$

$$\text{Отношение правдоподобия отрицательного результата теста (ОП-)} = \frac{1 - \text{чувствительность}}{\text{специфичность}} = \frac{[c/(a + c)]}{[d/(b + d)]}$$

Таблица 5. Сравнение результатов изучаемого теста (вентиляционно-перфузионная сцинтиграфия легких) и "золотого стандарта" (ангиопульмонография)*

Оценка результатов сцинтиграфии	Результаты ангиопульмонографии	
	ТЭЛА имеется	ТЭЛА отсутствует
"Высокая вероятность наличия ТЭЛА"	102	14
Все другие оценки	149	616
Всего	251	630

Примечание * — положительными (действительно свидетельствующими о "наличии патологических изменений") считались только результаты вентиляционно-перфузионной сцинтиграфии легких, которые ранее были отнесены к категории "высокая вероятность наличия ТЭЛА". Чувствительность изучаемого теста составляет 41%, специфичность — 98%, ОП+ равно 18,3, а ОП всех других результатов — 0,61

легких и ангиопульмонографии, применявшейся в качестве "золотого стандарта".

Чтобы рассчитать чувствительность вентиляционно-перфузионной сцинтиграфии легких на основе данных, представленных в табл. 6, число больных с положительными результатами сцинтиграфии ($n=102$) следует разделить на общее число больных, у которых наличие ТЭЛА было подтверждено с помощью АПГ ($n=251$); чувствительность изучаемого теста составляет примерно 41% ($a / a + c$, т.е. $102 / 251$). Чтобы рассчитать специфичность изучаемого теста, число больных с "нормальными" (отрицательными) результатами сцинтиграфии ($n=616$) следует разделить на число больных, у которых диагноз ТЭЛА не был подтвержден при АПГ ($n=630$); специфичность изучаемого теста составляет 98% ($d/b + d$, т.е. $616 / 630$). На основании этих данных можно рассчитать значения ОП+ и ОП—, которые будут равны 18,3 и 0,61 соответственно.

Если изменить критерии положительного и отрицательного результата вентиляционно-перфузионной сцинтиграфии легких (например, считать отрицательными только те результаты, которые ранее мы отнесли к категории "нормальные или почти нормальные"), характеристики этого теста также изменятся. Как следует из данных, приведенных в табл. 6, его чувствительность повысится до 98% ($246 / 251$, так как из 251 больного с наличием ТЭЛА, подтвержденной при АПГ, у 246 больных результат сцинтиграфии также будет считаться положительным), а специфичность существенно снизится (до 20%). Это обусловлено тем, что результаты сцинтиграфии окажутся отрицательными лишь у 126 из 630 больных с отвергнутым при АПГ диагнозом ТЭЛА. При этом значения

Таблица 6. Сравнение результатов изучаемого диагностического теста (вентиляционно-перфузионная сцинтиграфия легких) и "золотого стандарта" (ангиопульмонография)*

Оценка результатов сцинтиграфии	Результаты ангиопульмонографии	
	ТЭЛА имеется	ТЭЛА отсутствует
"Высокая, промежуточная или низкая вероятность наличия ТЭЛА"	246	504
"Нормальные или почти нормальные"	5	126
Всего	251	630

Примечание * — отрицательными (свидетельствующими об отсутствии патологических изменений) считали только те результаты вентиляционно-перфузионной сцинтиграфии легких, которые мы отнесли к категории "нормальные или почти нормальные" Чувствительность метода составляет 98%, специфичность — 20%. ОП+ и ОП- равны 1,23 и 0,1 соответственно

ОП+ и ОП— составят 1,23 и 0,1 соответственно. Таким образом, использование новой точки разделения не только снижает диагностическую информативность результатов сцинтиграфии, отнесенных нами ранее к категории "высокая вероятность наличия ТЭЛА", но и создает ложное впечатление о повышении вероятности наличия этого заболевания при результатах теста, отнесенных к категориям "промежуточная вероятность наличия ТЭЛА" и "низкая вероятность наличия ТЭЛА". Самостоятельно можно построить еще одну таблицу формата 2x2, поместив точку разделения посередине между рассмотренными выше крайними ситуациями. Если при этом чувствительность и специфичность теста составят 82 и 63% соответственно, а ОП+ и ОП— 2,25 и 0,28 соответственно, значит, вы все сделали правильно.

Когда при оценке информативности того или иного диагностического метода используются понятия чувствительность и специфичность, приходится либо игнорировать часть важной информации, либо рассчитывать значения этих показателей для каждой точки разделения. Мы рекомендуем применять такой показатель, как ОП; подобный подход к оценке результатов теста гораздо проще и эффективнее.

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

Таким образом, мы установили, что результаты исследования PLOPED скорее всего достоверны, и рассчитали значения ОП для различных результатов вентиляционно-перфузионной сцинтиграфии легких. Мы также пришли к выводу, что исследователи могли завязать оценку информативности КТ при диагностике аппендицита, но не настолько, чтобы это заставило нас вообще отказаться от использования полученных данных. Остается ответить на вопрос о возможности применения этого метода в условиях вашей практики.

КАК МОЖНО ПРИМЕНИТЬ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ НА ПРАКТИКЕ?

НАСТОЛЬКО ЛИ УБЕДИТЕЛЬНЫ ДАННЫЕ О ВОСПРОИЗВОДИМОСТИ РЕЗУЛЬТАТОВ ИЗУЧАЕМОГО МЕТОДА И ИНТЕРПРЕТАЦИЯ ЭТИХ ДАННЫХ, ЧТОБЫ ИХ МОЖНО БЫЛО ПРИМЕНИТЬ В УСЛОВИЯХ ВАШЕЙ ПРАКТИКИ?

Ценность любого диагностического теста зависит от возможности получения аналогичного результата в случае применения этого теста у больных со сходными характеристиками. Низкая воспроизводимость результатов теста может быть обусловлена особенностями его выполнения (например, применением разных коммерческих наборов для радиоиммунного определения уровня гормонов в крови) или интерпретации полученных данных (например, разным суждением о степени подъема сегмента *STnа* электрокардиограмме). В идеале в любой статье, посвященной характеристикам диагностического теста, должна оцениваться воспроизводимость его результатов (с внесением поправок на влияние фактора случайности). Это особенно важно в тех случаях, когда при проведении теста или интерпретации его результатов требуется высокая квалификация. Известно, что мнения нескольких врачей зачастую расходятся при анализе одной и той же ЭКГ, данных ультразвукового исследования или КТ, даже если в обсуждении участвуют только опытные специалисты.

При средней воспроизводимости результатов диагностического теста и умеренном расхождении мнений исследователей он может оказаться весьма полезным в вашей практике, если помогает различать больных с наличием или отсутствием предполагаемого заболевания. Если воспроизводимость результатов теста очень высокая, а мнения исследователей практически не различаются, не исключено, что данный диагностический метод прост и не допускает двусмысленного толкования; возможно также, что его результаты интерпретировали специалисты высокого класса. В последнем случае использование теста может оказаться менее эффективным, чем вы рассчитывали, поскольку в условиях вашей практики не все специалисты могут обладать такой квалификацией.

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

В отчете об исследовании, посвященном определению информативности КТ, не сообщается о воспроизводимости результатов метода. Упоминается только, что томограммы сначала анализировали ординаторы, а затем консультанты, которые лишь в одном случае не согласились с мнением ординаторов. Авторы не сообщают также о квалификации рентгенологов, однако участие в исследовании ординаторов позволяет предположить, что для правильной интерпретации результатов КТ особая квалификация не обязательна.

ВОЗМОЖНО ЛИ ПРИМЕНЕНИЕ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ У КОНКРЕТНОГО БОЛЬНОГО?

Характеристики диагностического теста могут зависеть от тяжести основного заболевания, а также от наличия и числа сопутствующих заболеваний. Если у всех участников исследования основное заболевание протекает в тяжелой форме, значения ОП будут существенно отличаться от 1,0 (т.е. чувствительность теста возрастет). При заболевании средней тяжести значения ОП будут в меньшей степени отличаться от 1,0 (чувствительность теста снизится). Если в отсутствие предполагаемого

заболевания сопутствующие болезни имитируют его проявления, результаты теста будут такими же, как и при наличии этого заболевания; при этом значения ОП будут приближаться к 1,0, а сам тест окажется менее пригодным (его специфичность снизится). В других клинических условиях число таких больных может быть меньше; тогда значения ОП будут в большей степени отличаться от 1,0 и тест окажется более полезным для диагностики (чувствительность теста возрастет).

Наиболее ярко возможные различия в информативности тестов, применяемых в разных популяциях больных, были продемонстрированы при использовании ЭКГ с физической нагрузкой для диагностики ИБС. Так, при более тяжелой ИБС, позднее подтвержденной с помощью коронарной ангиографии, вероятность выявления патологических изменений на ЭКГ была выше (т.е. больше были значения ОП+) [15]. Точно так же компрессионное ультразвуковое исследование, применяемое для диагностики проксимального тромбоза глубоких вен, оказалось более информативным у амбулаторных больных с клиническими проявлениями этого заболевания, чем у больных, перенесших обширное хирургическое вмешательство на суставах и не имеющих соответствующей симптоматики [16].

Иногда диагностический метод оказывается неинформативным именно в тех случаях, когда на него возлагались большие надежды. Так, при использовании тест-полосок для экспресс-диагностики инфекций мочевых путей значение ОП— оказалось равным 0,2 у больных с явными клиническими проявлениями, т.е. с высокой вероятностью наличия этих инфекций, но превышало 0,5 в тех случаях, когда она была низкой (что не позволяет использовать этот тест для исключения инфекций мочевых путей) [17]. Если условия вашей медицинской практики соответствуют условиям рассматриваемого исследования, а характеристики конкретного больного удовлетворяют всем критериям включения (и не отвечают ни одному из критериев исключения), можно быть уверенным, что результаты этого исследования в данном случае применимы. Если это не так, необходимо проанализировать ситуацию более тщательно. Как и при оценке применимости лечебных вмешательств, необходимо задать себе вопрос, существуют ли веские причины, мешающие применению результатов исследования у конкретного больного (иная тяжесть основного заболевания или иное сочетание сопутствующих заболеваний). Данную проблему можно решить, если будет найден обзор, в котором обобщаются результаты нескольких исследований [18].

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

В исследовании PLOPED была сформирована репрезентативная выборка больных с предполагаемой ТЭЛА, обратившихся в несколько крупных многопрофильных больниц. Поэтому полученные результаты применимы в большинстве лечебных учреждений Северной Америки. Однако в некоторых ситуациях это может оказаться нежелательным; например, при обследовании больных, находящихся в критическом состоянии. Таких больных не включали в исследование; спектр сопутствующих заболеваний и состояний у них может существенно отличаться.

Участницы исследования, посвященного определению информативности КТ при остром аппендиците, также представляют собой репрезентативную выборку женщин, обращающихся в отделения неотложной медицинской помощи с жалобами на боль в правой подвздошной области. Ваша больная, которой необходимо провести дифференциальную диагностику между острым аппендицитом и воспалительными заболеваниями органов малого таза, вполне отвечает критериям включения в это исследование. Таким образом, вы можете быть уверены, что в данном случае результаты исследования применимы.

ОКАЖУТ ЛИ ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ ВЛИЯНИЕ НА ВАШУ ТАКТИКУ ВЕДЕНИЯ БОЛЬНЫХ?

Принимая клинические решения, передавая и получая такую информацию (в том числе и в процессе обучения), всегда следует учитывать вероятность наличия предполагаемого заболевания. Как уже упоминалось, для этого показателя существует нижняя граница, за которой диагноз можно отвергнуть без дополнительного обследования (ПДО), и верхний порог, за которым можно немедленно начать лечение (ПЛ). Если вероятность наличия заболевания находится между ПДО и ПЛ, дальнейшее обследование надо проводить обязательно (см. главу 6 *Диагностика*).

После выбора значений ПДО и ПЛ тактика лечения будет прямо зависеть от послетестовой вероятности наличия искомого заболевания. Предположим, вы планируете начинать лечение при вероятности наличия ТЭЛА $\geq 80\%$ (понимая, что в 20% случаев лечение будет ненужным) и отвергать этот диагноз при послетестовой вероятности наличия ТЭЛА $< 10\%$. Указанные границы могут меняться, так как значения ПДО и ПЛ

зависят от риска развития осложнений, связанных с лечением (если такой риск высок, необходимость в более точной диагностике возрастает) или самим заболеванием в отсутствие лечения (например, при предполагаемой ТЭЛА от дальнейшего обследования отказываются только при очень низкой послетестовой вероятности этого заболевания). Вернемся к клиническому примеру с 28-летним мужчиной. Если результаты вентиляционно-перфузионной сцинтиграфии легких будут свидетельствовать о "высокой вероятности наличия ТЭЛА", послетестовая вероятность этого диагноза составит 82%; в таком случае надо будет начинать лечение (или по меньшей мере провести дальнейшее обследование). Если результаты сцинтиграфии будут свидетельствовать о "промежуточной вероятности наличия ТЭЛА", послетестовая вероятность этого диагноза составит 23%; в этом случае обследование необходимо продолжить (возможно, выполнить двустороннюю флебографию или ультразвуковое исследование сосудов нижних конечностей либо АПГ). Если результаты сцинтиграфии будут свидетельствовать о "низкой вероятности наличия ТЭЛА" либо будут расценены как "нормальные или почти нормальные", послетестовая вероятность этого диагноза не превысит 10%; в этом случае диагноз ТЭЛА можно исключить без дополнительного обследования. Теперь вернемся к клиническому примеру с 78-летней больной. Если результаты вентиляционно-перфузионной сцинтиграфии легких будут свидетельствовать о "высокой вероятности наличия ТЭЛА", послетестовая вероятность этого диагноза составит 97%; в этом случае следует безотлагательно начинать лечение. Если результаты сцинтиграфии будут свидетельствовать о "промежуточной вероятности наличия ТЭЛА", послетестовая вероятность составит 74%; тогда можно либо начинать лечение, либо продолжить обследование (например, выполнить АПГ). При всех других результатах сцинтиграфии следует проводить дополнительное обследование.

Диагностический тест не очень пригоден, когда у большинства больных его результатам соответствуют значения ОП, приближающиеся к 10. Таким образом, диагностическая ценность любого теста в значительной степени зависит от доли больных с предполагаемым диагнозом, у которых результатам теста соответствует очень высокое или очень низкое значение ОП. В двух рассмотренных выше примерах (с использованием вентиляционно-перфузионной сцинтиграфии легких при предполагаемой ТЭЛА) можно воспользоваться табл. 3 и определить Долю больных с самыми высокими и самыми низкими значениями ОП ("высокой вероятности наличия ТЭЛА" соответствует $ОП > 10$, "нормальным или почти нормальным" результатам сцинтиграфии соответ-

ствует ОП=0,1). Доля таких больных рассчитывается следующим образом: $(102 + 14 + 5 + 126) / 881 = 247 / 881 = 28\%$. Врачи, которые ранее уже были несколько разочарованы тем, что при предполагаемой ТЭЛА результаты скинтиграфии легких часто свидетельствуют о промежуточной или низкой вероятности наличия этого заболевания, должны понимать и то, что доля больных (28%) с другими результатами, т.е. с самыми высокими или самыми низкими значениями ОП, явно недостаточна. Это позволяет говорить об ограниченной диагностической ценности вентилиционно-перфузионной скинтиграфии легких при предполагаемой ТЭЛА.

В заключение следует кратко обсудить вопрос о последовательности применения диагностических тестов. Ранее было продемонстрировано, что каждый этап сбора анамнеза и общего осмотра можно считать диагностическим тестом. Претестовая вероятность наличия определенного заболевания постоянно изменяется после получения новых данных. Важны также результаты лабораторных анализов или визуализирующих методов исследования. Однако, если характеристики двух диагностических тестов сходны, их последовательное применение будет неинформативным или малоинформативным, а выводы, сделанные на основе последовательного анализа значений ОП, — ошибочными. Например, после определения уровня ферритина в сыворотке крови (самый точный метод диагностики железодефицитной анемии) применение дополнительных тестов, таких как определение концентрации железа в сыворотке крови или насыщения трансферрина, не станет источником важных диагностических данных [20]. *Правила клинического прогнозирования* позволяют применять нескольких взаимосвязанных диагностических тестов; совместное рассмотрение результатов этих тестов помогает определить вероятность наличия того или иного заболевания. Так, при анализе клинической ситуации, описанной в начале главы 6 *Диагностика*, оценка состояния дыхательной системы, частоты сердечных сокращений, кровообращения в нижних конечностях, насыщения крови кислородом, результатов ЭКГ, данных анамнеза и общего осмотра позволяет точно квалифицировать вероятность наличия ТЭЛА как высокую, среднюю и низкую [21].

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДОВАННОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

Учитывая крайние значения ОП при выполнении КТ у женщин с болью в животе, можно предположить, что результаты данного диагностического теста скорее всего повлияют на тактику лечения. У всех больных со средней претестовой вероятностью наличия аппендицита положительные результаты КТ будут указывать на необходимость экстренной операции, а отрицательные результаты — на необходимость дальнейшего наблюдения и лечения других возможных заболеваний (в данном случае воспалительных заболеваний органов малого таза).

улучшится ли качество оказания медицинской помощи вследствие применения изучаемого диагностического метода?

О пригодности того или иного теста можно говорить, когда польза от его применения превышает риск развития осложнений [22]. Каким образом можно оценить преимущества и недостатки применения диагностического теста? Для этого необходимо рассматривать его с тех же позиций, что и лечебное вмешательство (см. главу 4 *Исследования, посвященные эффективности лечения*). Оценить соотношение преимуществ и недостатков диагностического теста можно с помощью рандомизации (одни больные распределяются в основную группу, где будет применяться изучаемый метод исследования, другие — в контрольную группу, где этот метод применяться не будет) и последующего наблюдения, позволяющего определить в этих группах частоту развития разных клинических исходов, важных для больного.

В каких случаях выявленная диагностическая точность теста позволяет сразу рекомендовать его для практического использования, а в каких необходимо дополнительное проведение рандомизированных контролируемых испытаний? Целесообразность применения информативного теста не вызывает сомнений, когда отсутствие своевременной диагностики приводит к развитию тяжелых осложнений, применение теста достаточно безопасно, и существуют эффективные методы лечения выявляемого заболевания. Все перечисленное можно отнести к двум Диагностическим тестам, рассмотренным в данном разделе книги. Результаты вентиляционно-перфузионной сцинтиграфии легких, указывающие на "высокую вероятность наличия ТЭЛА", позволяют отказать-

ся от дальнейшего обследования и своевременно начать вполне обоснованную терапию антикоагулянтами; "нормальные или почти нормальные" результаты теста позволяют отказаться от дальнейшего обследования и лечения (в обоих случаях выбранная тактика ведения больного будет положительно влиять на клинические исходы).

Исследователи, определявшие информативность КТ у женщин с болью в животе, просили врачей спланировать лечение до получения результатов теста, после чего сравнивали предложенную тактику с той, которая применялась в действительности. Всего в исследование были включены 100 участниц. После первичного обследования, выполненного до проведения КТ, планировалось оставить 56 участниц для наблюдения в условиях стационара; после проведения КТ 43 из них были выписаны домой, а 13 направлены на экстренную аппендэктомию. Таким образом, очевидные преимущества КТ для больных (которые предпочитают находиться дома, а не в больнице) и для системы здравоохранения в целом (несвоевременное выполнение аппендэктомии повышает риск развития тяжелых осложнений) позволяют отказаться от проведения рандомизированных контролируемых испытаний, в которых сравнивалась бы целесообразность применения КТ и традиционных методов диагностики у женщин с болью в животе.

В некоторых ситуациях диагностический тест может быть информативным и влиять на тактику лечения, но связь между его применением и клиническими исходами может оставаться неясной. Вспомним одну из проблем, касающуюся постановки клинических вопросов (см главу 2 *Поиск доказательной информации*). В этом примере мы сравнивали целесообразность выполнения КТ грудной клетки, на основании результатов которой определяется дальнейшая тактика ведения больного, или первичной медиастиноскопии при немелкоклеточном раке легкого без признаков метастазирования. В такой ситуации недостаточно знать о диагностической точности КТ. Целесообразность применения этих двух диагностических методов необходимо сравнивать в ходе рандомизированного испытания с оценкой клинических исходов у всех больных; и такое испытание было действительно проведено [23]. Другими примерами диагностических тестов с неясным влиянием на клинические исходы могут служить катетеризация правых полостей сердца у больных, находящихся в критическом состоянии, в отсутствие точной информации о состоянии гемодинамики; бронхоальвеолярный лаваж при предполагаемом пульмоногенном сепсисе у больных, находящихся в критическом состоянии; провокационные пробы при бронхиальной астме; преимущественное выполнение магнитно-резонансной томографии

(вместо КТ) при многих заболеваниях. Окончательные выводы о целесообразности применения этих и многих других диагностических методов можно будет сделать только после проведения хорошо спланированных рандомизированных испытаний с достаточной статистической мощностью.

ПРИНЯТИЕ КЛИНИЧЕСКОГО РЕШЕНИЯ

Статья, посвященная оценке информативности КТ, произвела настолько сильное впечатление, что вы решили не приглашать гинеколога на консультацию. Рентгенолог посодействовал в скорейшем выполнении КТ и вскоре сообщил, что полученные результаты подтверждают диагноз аппендицита. Больную срочно отправили в операционную, позднее вы узнали о том, что аппендэктомия была выполнена своевременно, и послеоперационный период протекает без осложнений.

Литература

- 1 Rao P M, Feltmate C M, Rhea J T, Schuhck A H, Novellme R A Helical computed tomography in differentiating appendicitis and acute gynecologic conditions *Obstet Gynecol* 1999,93 417—421
- 2 Thomson D M , Krupcy J, Freedman S O , Gold P The radioimmunoassay of circulating carcinoembryonic antigen of the human digestive system *Proc Natl Acad Sci USA* 1969,64 161 — 167
- 3 Bates S E Clinical applications of serum tumor markers *Ann Intern Med* 1991,115 623—638
- 4 Lijmer J G , Mol B W, HeШезkamp S , et al Empirical evidence of design-related bias in studies of diagnostic tests *JAMA* 1999 282 1061-1066
- 5 The PИOPED investigators Value of ventilation/perfusion scan in acute pulmonary embolism Results of the prospective investigation of pulmonary embolism diagnosis (PИOPED) *JAMA* 1990 263 2753—2759
- 6 Sackett D L , Haynes R B , Guyatt G H Tugwell P Clinical Epidemiology Basic Science for Clinical Medicine 2nd ed Boston Little, Brown and Company, 1991 53—57
- 7 Kemppanen EA , Hedstrom JI, Puolakkamen PA , et al Rapid measurement of urinary trypsinogen-2 as a screening test for acute pancreatitis *N Engl J Med* 1997 336 1788—1793
- 8 Begg C B Greenes RA Assessment of diagnostic tests when disease verification is subject to selection bias *Biometrics* 1983.39 207—215
- 9 Gray R , Begg C B , Greenes RA Construction of receiver operating characteristic curves when disease verification is subject to selection bias *Med Decis Making* 1984 4 151 — 164
- 10 Ransohoff D F, Femstem A R Problems of spectrum and bias in evaluating the efficacy of diagnostic tests *N Engl J Med* 1978,299 926-930
- U Choi B C Sensitivity and specificity of a single diagnostic test in the presence of work up bias *J Chn Epidemiol* 1992 45 581-586
- 12 Fagan TJ Nomogram for Bayes s theorem *N Engl J Med* 1975 293 257
- 13 Sackett D L, Renne D The science and art of the clinical examination *JAMA* 1992 267 2650-2652
- 14 Black ER Bordley D R Tape TG Panzer R J Diagnostic Strategies for Common Medical Problems 2nd ed Philadelphia American College of Physicians 1999

15. *Hlaiky M.A., Pryor D.B., Harrell F.E.* Factors affecting sensitivity and specificity of exercise electrocardiograph[^] *Am J Med* 1984;77-64—71.
16. *Ginsberg J.S., Caco C.C., Brill-Edwards P.A., et al.* Venous thrombosis in patients who have undergone major hip or knee surgery, detection with compression US and impedance plethysmography. *Radiology* 1991;181:651-654.
17. *Lachs M.S., Nachamk I., Edelstem P.H., et al.* Spectrum bias in the evaluation of diagnostic tests: lessons from the rapid dipstick test for urinary tract infection. *Ann Intern Med* 1992;117:135—140.
18. *Irwig L., Tosteson A.N., Gatsonis C. et al.* Guidelines for meta-analyses evaluating diagnostic tests. *Ann Intern Med* 1994;120:667-676.
19. *Sacken D.L., Haynes R.B., Guyatt G.H., Tugwell P.* *Clinical Epidemiology, a Basic Science for Clinical Medicine.* 2nd ed. Boston: Little, Brown and Company, 1991:145—148
20. *Guyatt G.H., Oxman A., AH M., et al.* Laboratory diagnosis of iron-deficiency anemia, an overview. *J Gen Intern Med* 1992;7:145-153.
21. *Wells P.S., Ginsberg J.S., Anderson D.R., et al.* Use of a clinical model for safe management of patients with suspected pulmonary embolism. *Ann Intern Med* 1998;129:997—1005.
22. *Guyatt G.H., Tugwell P.X., Feeny D.H., Haynes R.B., Drummond M.* A framework for clinical evaluation of diagnostic technologies. *CMAJ* 1986;134:587-594.
23. *Canadian Lung Oncology Group.* Investigation for mediastinal disease in patients with apparently operable lung cancer. *Ann Thorac Surg* 1995;60:1382-1389.

9

ПРОГНОЗ

А. Рэндольф, Х. Бачер, У.С. Ричардсон,
Д. Уэллс, П. Тагвелл, Г. Гайятт

В подготовке данной главы принимали также участие члены Рабочей группы по разработке и внедрению принципов доказательной медицины Д. Кук, Д. Крэйг, Дж. Уайятт

В ЭТОМ РАЗДЕЛЕ

Поиск доказательной информации

Достоверны ли результаты исследования?

Была ли репрезентативной выборка больных?

Были ли больные в достаточной степени сходны по всем известным факторам, влияющим на прогноз?

Какова доля больных, наблюдавшихся до конца исследования?

Использовались ли объективные критерии оценки, исключающие возможность возникновения систематической ошибки?

Каковы результаты исследования?

Насколько вероятно развитие того или иного клинического исхода в течение определенного времени?

Насколько точна оценка прогноза?

Как применить результаты исследования на практике?

Соответствуют ли характеристики участников исследования и методы их лечения условиям вашей клинической практики?

Было ли наблюдение достаточно длительным?

Можно ли применять результаты исследования в вашей практике?

Принятие клинического решения

ВОЗМОЖНАЯ КЛИНИЧЕСКАЯ СИТУАЦИЯ

Представьте себе, что вы шведский врач, к которому обратился мужчина 71 года, выздоравливающий после правосторонней нижнедолевой пневмонии. Год назад больной перенес инсульт с локализацией очага в правом полушарии головного мозга; функция левой руки восстановилась лишь частично, но он может передвигаться с помощью костылей. На протяжении предшествующих 15 лет больной страдал артериальной гипертензией, но гипотензивные средства принимал нерегулярно. В настоящее время на фоне приема аспирина (по 300 мг/сут), ингибитора ангиотензинпревращающего фермента (АПФ) и тиазидного диуретика уровень артериального давления (АД) поддерживается в допустимых пределах. Ритм сердца синусовый. При эхокардиографии выявлены признаки гипертрофии и умеренно выраженной дисфункции левого желудочка. При доплеровском исследовании отмечено наличие клинически незначимого стеноза (<50%) обеих сонных артерий.

Из беседы с больным вы узнали, что он знаток французских вин и после раннего выхода на пенсию ежегодно проводит несколько месяцев в собственном небольшом коттедже на юге Франции. Он жалуется на то, что после перенесенного инсульта "все идет не так, как следует". Затем вы поговорили с женой больного, которую очень беспокоит тот факт, что ее муж практически не желает признавать наличие у себя физического недостатка. После перенесенного инсульта содержание двух домов полностью легло на ее плечи. Она считает, что постоянные переезды стали слишком сложны для них, а пневмония является лишним доказательством "истощения" мужа. Ей кажется, информация о риске развития повторного инсульта и возможном летальном исходе поможет убедить ее мужа в необходимости пересмотра стиля жизни семьи в целом. Не обладая достаточными сведениями о прогнозе у выживших после инсульта, вы обещаете жене больного выяснить все интересующие ее данные.

ПОИСК ДОКАЗАТЕЛЬНОЙ ИНФОРМАЦИИ

В вашей больнице нет прямого доступа к электронным базам данных Best Evidence или Cochrane Library, но имеется выход в Интернет. Во время перерыва вы садитесь за компьютер и с помощью поисковой системы PubMed входите в MEDLINE (библиографическая электронная база данных, организованная и поддерживаемая Национальной медицинской библиотекой США). Введя слово "stroke" ("инсульт") и используя словарь синонимов (тезаурус), вы находите соответствующий термин медицинского предметного рубрикатора (MeSH) - "инфаркт мозга" ("cerebral infarction"). Комбинируя результаты этого поиска и поиска по ключевым словам "epidemiology" ("эпидемиология"), "recurrence" ("рецидив") и "prognosis" ("прогноз"), вы находите целый ряд ссылок. Вас заинтересовала одна публикация, включающая материалы исследования Oxfordshire Community Stroke Project и посвященная оценке отдаленного риска развития рецидива после впервые перенесенного инсульта ("Long-term risk of recurrent stroke after a first-ever stroke" [1]), поэтому вы решаете распечатать ее полный текст.

В процессе оказания медицинской помощи больному врач стремится к достижению трех основных целей: установлению правильного диагноза, назначению эффективного и относительно безопасного лечения, а также предоставлению информации о том, что ожидает больного в будущем. Для достижения второй и третьей целей врачам необходимы материалы исследований, посвященных изучению *прогноза* (т.е. возможных клинических исходов заболевания и вероятности их возникновения). Хотя основная задача врача заключается в восстановлении здоровья, иногда он может лишь облегчить страдания больного и предоставить ему или его родственникам наиболее полную и достоверную информацию о неизбежных неблагоприятных исходах.

Чтобы оценить прогноз у конкретного больного, следует изучать исходы у других больных со сходными клиническими проявлениями (например, в группе больных, перенесших инсульт в течение предшествующего года). Для получения более точного прогноза следует проанализировать клинические исходы именно в той подгруппе, к которой относится конкретный больной. Сформировать такие подгруппы можно в зависимости от возраста (у молодых больных прогноз может быть более благоприятным), особенностей заболевания (при геморрагическом или ишемическом инсульте клинические исходы могут различаться) или сопутствующих заболеваний (при наличии артериальной гипертензии, даже на фоне гипотензивной терапии, прогноз может быть хуже). Если ка-

кие-либо из этих факторов действительно позволяют предсказать развитие определенного клинического исхода, их называют *прогностическими факторами*.

Авторы часто различают такие термины, как *прогностические факторы* и *факторы риска*; наличие последних скорее отражает вероятность возникновения болезни. Например, курение относится к важным факторам риска развития рака легкого, но у больных с уже диагностированным раком оно не считается столь же важным прогностическим фактором, как стадия опухолевого процесса. В исследованиях, посвященных изучению прогностических факторов или факторов риска, решаются одинаковые вопросы; во всех случаях оценивается обоснованность отнесения определенных признаков к таким факторам и их клиническая значимость.

Знание прогноза может позволить врачу принять правильное решение о тактике обследования и лечения больного. Если прогноз благоприятен, не следует назначать инвазивные методы обследования, связанные с высоким риском развития осложнений, а также применять дорогостоящие или токсичные лекарственные препараты. При низком риске развития неблагоприятного клинического исхода применение даже самых эффективных методов лечения может оказаться ненужным. Например, терапия, назначаемая для профилактики стрессовых язв и обусловленного ими желудочно-кишечного кровотечения, может оказаться бессмысленной у неинтубированных больных без коагулопатии, поскольку в этих случаях риск возникновения клинически значимых кровотечений крайне низок [2]. У молодых некурящих больных со слабо выраженной гиперхолестеринемией без артериальной гипертонии и наследственной предрасположенности к ишемической болезни сердца риск развития сердечно-сосудистых заболеваний в течение последующих 10—20 лет настолько низок, что они могут не принимать гиполипидемические препараты. С другой стороны, ожидаемый клинический исход у больных может быть неблагоприятным независимо от выбранной врачом тактики лечения. В этих случаях интенсивное лечение только продлит страдания больных и будет сопровождаться большими затратами.

Знание прогноза позволяет не только определять тактику ведения конкретного больного, но и решать более широкий круг вопросов. Иногда делаются попытки сравнить качество медицинской помощи, оказываемой разными врачами или в разных лечебных учреждениях, путем оценки результатов лечения. Однако различия в клинических исходах могут быть обусловлены разной исходной тяжестью заболевания, а не

тактикой лечения, квалификацией врачей или особенностями медицинского учреждения. Имея информацию о прогнозе у конкретного больного, можно сравнить разные популяции, внести поправки на различия в их прогнозе и уточнить, каким образом определенная тактика лечения влияет на клинические исходы (см. главу 2В на компакт-диске).

Данные о прогнозе используются для решения таких вопросов, как информирование больного о дальнейшем течении заболевания, принятие клинических решений, сравнение клинических исходов в разных популяциях для уточнения качества оказания медицинской помощи; способы их решения могут быть разными. В этой главе основное внимание уделено тому, как пользоваться материалами статей, в которых может содержаться достоверная прогностическая информация, полезная при общении с больными (см. таблицу).

Таблица. Рекомендуемый подход к оценке публикации, содержащей информацию о прогнозе

Достоверны ли результаты исследования?

- Была ли репрезентативной выборка больных⁹
- Были ли больные в достаточной степени сходны по всем известным факторам, влияющим на прогноз?
- Какова доля больных, наблюдавшихся до конца исследования⁹
- Использовались ли объективные критерии оценки, исключающие возможность возникновения систематической ошибки⁹

Каковы результаты исследования?

- Насколько вероятно развитие того или иного клинического исхода в течение определенного времени⁹
- Насколько точна оценка прогноза⁹

Как применить результаты исследования на практике?

- Соответствуют ли характеристики участников исследования и методы их лечения условиям вашей клинической практики⁹
 - Было ли наблюдение достаточно длительным⁹
 - Можно ли применять результаты исследования в вашей практике⁹
-

Как и при изучении потенциально вредных воздействий или безопасности лечения (см. главу 5 *Оценка риска*), для изучения прогноза обычно проводятся когортные исследования и исследования случай-контроль. Несомненно, эти вопросы рассматриваются и в ходе РКИ. Результаты, полученные в основной и контрольной группах, в определенном смысле представляют собой прогностическую информацию,

позволяя судить о возможных клинических исходах при лечении и в его отсутствие. С этой точки зрения каждую группу в РКИ можно рассматривать как когортное исследование. Если РКИ отвечает критериям, которые будут изложены ниже, его результаты могут служить источником полезной информации о дальнейшей судьбе больного.

При изучении безопасности лечения выбор соответствующей группы вмешательства и контроля имеет решающее значение. В исследованиях, посвященных изучению прогноза, контрольную группу либо не формируют вообще (например, при схожем прогнозе у всех участников), либо включают в нее участников с наличием иных прогностических факторов. Точно так же, как в исследованиях, посвященных диагностике, изучается информативность тестов, которые позволяют подтвердить или отвергнуть наличие определенного заболевания, в этих исследованиях могут изучаться факторы, позволяющие прогнозировать высокий или низкий риск развития определенного клинического исхода. Однако при анализе результатов исследований, посвященных прогнозу, следует отнестись совсем на другие вопросы.

ДОСТОВЕРНЫ ЛИ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ?

БЫЛА ЛИ РЕПРЕЗЕНТАТИВНОЙ ВЫБОРКА БОЛЬНЫХ?

Наличие любой систематической ошибки приводит к систематическому отклонению получаемых результатов от истинных значений. В исследованиях, посвященных прогнозу, наличие такой ошибки способствует систематической переоценке или недооценке вероятности развития неблагоприятного клинического исхода. Если характеристики участников исследования и других больных с изучаемым заболеванием в общей популяции систематически различаются, в таком исследовании не исключено наличие систематической ошибки, а выборка считается нерепрезентативной.

Как можно распознать нерепрезентативную выборку? Прежде всего необходимо обратить внимание на использование каких-либо дополнительных критериев включения. Если такие "фильтры" применялись, полученные результаты скорее всего будут отличаться от тех, которые можно было бы получить в изучаемой популяции (больные с впервые диагностированным сахарным диабетом, после инфаркта миокарда или инсульта). Одним из таких "фильтров" можно считать последовательное обращение больных сначала в учреждение первичной медицинской

помощи, а затем в специализированный стационар. В специализированных лечебных центрах чаще оказывают помощь больным с редкими или более тяжелыми заболеваниями. Результаты исследований, проведенных в таких учреждениях, могут быть неприменимы в общей популяции больных.

Например, когда ребенка госпитализируют по поводу судорог на фоне лихорадки, его родители хотят знать о риске возникновения таких судорог в дальнейшем. Результаты популяционных исследований свидетельствуют о более низкой частоте развития повторных судорог (от 1,5 до 4,6%), чем результаты исследований, проводимых в специализированных центрах (от 2,6 до 76,9%) [3]. Это связано с тем, что больные, находящиеся на лечении в специализированных центрах, могут страдать другими неврологическими заболеваниями, которые также способствуют возникновению повторных судорог. Чтобы предоставить родителям достоверную информацию, врач должен знать, к какой выборке ребенок ближе по своим характеристикам.

Нечетко сформулированные критерии включения повышают вероятность того, что выборка была нерепрезентативной. Чтобы оценить ее репрезентативность, обязательно обратите внимание на критерии включения и исключения. В статье должны быть четко описаны не только эти критерии, но и объективные критерии, на основании которых был установлен диагноз.

Были ли больные в достаточной степени сходны по всем известным факторам, влияющим на прогноз?

Исследования, посвященные прогнозу, наиболее полезны для клинической практики в том случае, когда характеристики всех участников достаточно схожи, и общий результат применим к каждому участнику. Это возможно только при включении в исследование больных, находящихся в одной и той же, хорошо описанной стадии заболевания, которая не должна быть ранней. Например, при изучении прогноза у детей с повреждением головного мозга в исследование включали только больных, находящихся без сознания в течение предшествующих 90 дней [4].

Когда вы убедитесь, что различия в клинических исходах не могли быть обусловлены разной стадией заболевания, следует проанализировать возможное влияние других важных факторов. Снова обратимся к исследованию, в котором изучался прогноз у детей с повреждением головного мозга. При анализе неврологических исходов без учета причин повреждения мозга (черепно-мозговая травма или другие причины) результат может быть недостоверным из-за различного прогноза в этих

подгруппах. Так, если общая смертность составила 50%, но из выборки можно выделить подгруппы, в одной из которой смертность приближается к 100%, а в другой равна нулю, расчетный показатель будет применим ко всей выборке, но неприменим к любому из участников. Если у больных исходный риск развития неблагоприятного исхода различен, целесообразно выделить отдельно подгруппы высокого и низкого риска.

Например, T. Pincus et al. на протяжении 15 лет наблюдали за когортой больных с ревматоидным артритом [5]. Всех участников исследователи разделили на подгруппы в зависимости от демографических характеристик, особенностей заболевания и функционального состояния. Продолжительность жизни оказалась ниже не только у более пожилых участников, но и при худшем функциональном состоянии (например, при более низких значениях модифицированного показателя, отражающего расстояние, которое больной может пройти за определенное время, и повседневную активность). В другом, указанном выше исследовании у детей с посттравматическим повреждением головного мозга клинические исходы оказались более благоприятными, чем при гипоксическом повреждении мозга. Из 36 больных, перенесших закрытую черепно-мозговую травму, у 23 (64%) восстановление социальной функции было достаточным для того, чтобы выражать свои потребности и желания, а 9 (25%) больных в конечном итоге смогли самостоятельно передвигаться. Из 13 детей с гипоксическим повреждением головного мозга ни у одного не восстановились важные социальные или познавательные функции [4].

Исследователи должны учитывать не только все важные прогностические факторы, но и их взаимодействие. В ходе Фремингемского исследования среди многих показателей изучались и факторы риска развития инсульта [6]. Частота развития инсульта у больных с фибрилляцией предсердий на фоне ревматического порока сердца составила 41 на 1000 человеко-лет, т.е. этот показатель был приблизительно таким же, как при фибрилляции предсердий без ревматического порока. Однако средний возраст больных с ревматическим пороком был гораздо меньше. В этой ситуации, чтобы точнее понять роль ревматического порока сердца в развитии инсульта, исследователи должны были по отдельности проанализировать относительный риск развития инсульта у более молодых и более пожилых участников (при наличии у них порока или его отсутствии). Такой раздельный анализ данных называют *скорректированным анализом*, а его результаты — скорректированными, или стандартизованными. После внесения поправок на возраст, пол и уровень АД оказалось, что у больных с фибрилляцией предсердий наличие

ревматического порока сердца в 6 раз увеличивает риск развития инсульта. Если прогноз может существенно зависеть от множества разных факторов, то для выявления наиболее значимых из них исследователи должны использовать специальные статистические методы (см. главу 2D на компакт-диске). На основе такого анализа могут быть разработаны *правила принятия клинических решений*, позволяющие врачам одновременно учитывать все важные прогностические факторы (см. главу 2С на компакт-диске).

Как можно оценить однородность групп с точки зрения риска развития определенного клинического исхода? На основании собственного клинического опыта и понимания патофизиологии заболевания вы должны решить, были ли учтены все факторы, которые способны повлиять на прогноз. Если на этот вопрос нельзя ответить положительно, достоверность результатов исследования сомнительна.

КАКОВА ДОЛЯ БОЛЬНЫХ, НАБЛЮДАВШИХСЯ ДО КОНЦА ИССЛЕДОВАНИЯ?

Большое число выбывших больных снижает достоверность исследования, посвященного изучению прогноза. С увеличением доли участников, которых не удалось наблюдать до конца исследования, повышается и вероятность возникновения систематической ошибки (у таких больных риск развития изучаемого клинического исхода может быть существенно выше или ниже). Какое число выбывших больных следует считать "большим"? Ответ на этот вопрос зависит от соотношения числа выбывших больных и больных, у которых за время наблюдения развился неблагоприятный клинический исход. Чем больше число участников, судьба которых неизвестна, по отношению к числу больных с неблагоприятным исходом, тем сомнительнее *достоверность* результатов исследования.

Предположим, что при длительном наблюдении неблагоприятный клинический исход (смерть от сердечно-сосудистых заболеваний) отмечен у 30% участников из группы самого высокого риска (например, пожилые больные с сахарным диабетом). Если до конца исследования не наблюдались 10% больных, истинная доля умерших может колебаться от 27 до 40%. Внутри этого диапазона клиническая значимость полученных данных существенно не изменяется, и выбывание участников не оказывает большого влияния на достоверность исследования. Однако в группе более низкого риска (например, практически здоровые мужчины среднего возраста) частота возникновения этого неблагоприятно-

го исхода может составлять 1%. В этом случае (если мы предположим, что все 10% выбывших больных умерли) смертность может составить 11%; клиническая значимость такого результата будет уже совсем иной.

Если доля больных, не наблюдавшихся до конца исследования, была велика, достоверность результатов исследования особенно страдает при наличии существенных различий в клинических исходах у тех участников, найти которых невозможно, и у тех, кого легко найти. Так, в одном исследовании после длительного и непростого поиска удалось проследить судьбу 180 из 186 больных, ранее лечившихся от невроза [7]. Среди 60% больных, найти которых оказалось легко, смертность составила 3%, а среди тех, кого было гораздо труднее найти, — 27%.

Итак, если между выбывшими и наблюдавшимися до конца исследования участниками существуют различия в клинических исходах (а это верно для большинства исследований, посвященных прогнозу), преобладание числа больных, судьба которых неизвестна, над числом больных, у которых развился неблагоприятный клинический исход, ставит под сомнение достоверность результатов.

ИСПОЛЬЗОВАЛИСЬ ЛИ ОБЪЕКТИВНЫЕ КРИТЕРИИ ОЦЕНКИ, ИСКЛЮЧАЮЩИЕ ВОЗМОЖНОСТЬ ВОЗНИКНОВЕНИЯ СИСТЕМАТИЧЕСКОЙ ОШИБКИ?

В ходе исследований изучаются самые разные клинические исходы; они могут быть либо объективными и очевидными (смерть), либо требующими дополнительного обсуждения (инфаркт миокарда), либо очень сложными для оценки (снижение работоспособности, качество жизни). Исследователи должны заранее выбрать изучаемые клинические исходы, дать их четкое определение и по возможности применять объективные критерии оценки этих исходов. Кроме того, следует заранее определить объем и частоту оценки исходов. Чем более субъективна такая оценка, тем важнее применять слепой метод (т.е. лица, оценивающие клинические исходы, не должны знать о наличии тех или иных прогностических факторов).

Еще раз вернемся к исследованию, включавшему детей с повреждением головного мозга; это хороший пример того, какие проблемы возникают при оценке клинического исхода [4]. Установлено, что родственники зачастую слишком оптимистично трактовали возможности своего общения с ребенком. Поэтому авторы исследования потребовали, чтобы наличие положительной динамики подтверждали специалисты, а дату восстановления сознания обязательно стали увязывать с данными, полученными при повторном обследовании. Например, ребенок пребывал в коме на протяжении 1 года после повреждения головного

мозга, затем к нему вернулось сознание; длительность бессознательно-го периода считали равной 1 году только в том случае, если ребенок находился в сознании и во время следующего осмотра в клинике через 16 мес.

Вернемся к нашему основному клиническому примеру и к статье, посвященной прогнозу после впервые перенесенного инсульта. В исследовании Oxfordshire Community Stroke Project проспективное наблюдение за всеми 675 участниками начиналось с момента их поступления в стационар [1]. Таким образом, все больные включались на одной, достаточно ранней стадии заболевания. Поскольку исследование было популяционным, его результаты применимы ко всем случаям впервые выявленного инсульта в Великобритании. Средний возраст больных составил 72 года, 47% из них были мужчины. У 81% больных причиной инсульта послужил инфаркт мозга, у 10% — внутримозговое кровоизлияние, у 5% — субарахноидальное кровоизлияние, что соответствует результатам других исследований, посвященных изучению инсульта.

Можно предположить, что на частоту развития повторного инсульта влияют многие факторы, включая тяжесть первого инсульта, возраст больного, тип инсульта, наличие сахарного диабета или заболеваний сердца, а также уровень АД. Проанализировав все эти факторы, за исключением первого, исследователи не выявили различий между соответствующими подгруппами.

До конца исследования наблюдались 100% больных; такого уровня удалось достичь за счет амбулаторного посещения больных медицинскими сестрами через 1 мес, 6 мес, 1 год и далее ежегодно. Авторы привели четкое определение того, что понималось под термином "инсульт" (например, не учитывались новые поражения, выявляемые только при компьютерной томографии и протекавшие бессимптомно). Однако исследователи даже не пытались скрыть от медицинских сестер, осуществлявших наблюдение за больными, информацию о наличии возможных прогностических факторов.

КАКОВЫ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ? _____.

НАСКОЛЬКО ВЕРОЯТНО РАЗВИТИЕ ТОГО ИЛИ ИНОГО КЛИНИЧЕСКОГО ИСХОДА В ТЕЧЕНИЕ ОПРЕДЕЛЕННОГО ВРЕМЕНИ?

Количественные результаты исследований, посвященных изучению прогностических факторов и факторов риска, представляют в виде числа больных, у которых за время наблюдения развились определенные

клинические исходы. Чтобы продемонстрировать возможности получения дополнительной информации о прогнозе, рассмотрим следующий пример. К врачу обращается мужчина, который хотел бы узнать о прогнозе у его пожилой матери, страдающей деменцией.

Сына интересует, какова вероятность того, что его мать будет жива через 5 лет. Высокодостоверные результаты исследования, посвященного изучению прогноза при деменции [8], позволяют дать четкий и прямой ответ на этот вопрос. Через 5 лет после обращения в клинику по поводу деменции умерли 50% больных. Таким образом, в данном случае шансы выжить в течение 5 лет составляют 50:50.

Сын больной может вспомнить о том, что он лично знает лишь одного человека с болезнью Альцгеймера (его 65-летнего дядю), которому поставили этот диагноз 10 лет назад, и он до сих пор жив. Мужчина удивлен, что риск смерти у его матери столь высок. Это дает врачу возможность обсудить некоторые прогностические факторы смерти при болезни Альцгеймера. Результаты уже упоминавшегося достоверного исследования указывают на то, что раньше остальных умирают более пожилые больные, а также больные с более выраженной деменцией, с поведенческими расстройствами или со снижением слуха.

Следующий вопрос, который интересует сына: может ли прогноз у его матери изменяться со временем (т.е. может ли риск смерти быть низким в первые 2 года, а затем резко увеличиться). Отвечая на этот вопрос, невозможно выразить имеющуюся информацию ни в абсолютных, ни в относительных значениях; для этого необходимо обратиться к *кривой выживаемости*. Такие кривые отражают число клинических исходов, возникших в определенный момент времени (или шанс избежать неблагоприятного клинического исхода, в данном случае смерти, в тот же момент времени; см. главу 2В2 на компакт-диске). При этом оцениваемый клинический исход должен быть дискретным (например, смерть, развитие инсульта или рецидива рака); необходимо также точно знать время его оценки.

На рис. 1 изображены кривые выживаемости; в одном случае эти кривые отражают долю больных, выживших в определенные сроки после перенесенного инфаркта миокарда [9], в другом — долю больных, не просто выживших после протезирования тазобедренного сустава, но и обошедшихся без повторной операции [10]. Следует отметить, что вероятность смерти после перенесенного инфаркта миокарда наиболее высока в самый ранний период (о чем свидетельствует первоначальное резкое снижение кривых, затем они становятся пологими), тогда как после протезирования тазобедренного сустава необходимость в повторных

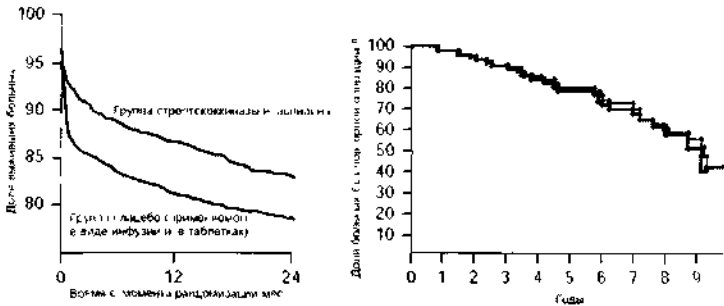


Рис. 1. Примеры кривых выживаемости (приводятся с разрешения The Lancet Publishing Group и The Journal of Bone and Joint Surgery соотв.).

Слева представлены кривые, которые отражают долю больных, выживших в определенные сроки после перенесенного инфаркта миокарда [9], а справа — долю больных, не просто выживших после протезирования тазобедренного сустава, но и обошедшихся без повторной операции [10]

операциях возникает в более поздние сроки (эти кривые начинаются с плато, а затем начинают снижаться).

Характер кривой выживаемости (рис. 2), построенной на основании данных уже упоминавшегося достоверного исследования [8], указывает на то, что при деменции ежегодная вероятность смерти не очень существенно изменяется в течение первых семи лет после обращения к специалисту.

НАСКОЛЬКО ТОЧНА ОЦЕНКА ПРОГНОЗА?

Чем более точной была оценка прогноза в ходе исследования, тем меньше сомнений у нас вызывает этот прогноз и тем полезнее оказывается найденная информация. Обычно риск развития неблагоприятного исхода представляют в виде точечного значения с 95% доверительным интервалом (ДИ). Если исследование достоверно, 95% ДИ определяет разброс значений риска, в пределах которого с очень высокой вероятностью находится истинное значение (см. главу 2B2 на компакт-диске). Например, в исследовании, посвященном изучению прогноза при деменции, через 5 лет после обращения к специалисту выживаемость составляла 49% при 95% ДИ от 39 до 58%. Следует отметить, что при построении большинства кривых выживаемости на начальных этапах наблюдения обычно учитываются результаты большего числа больных, чем на более поздних этапах (это обусловлено тем, что участники включа-

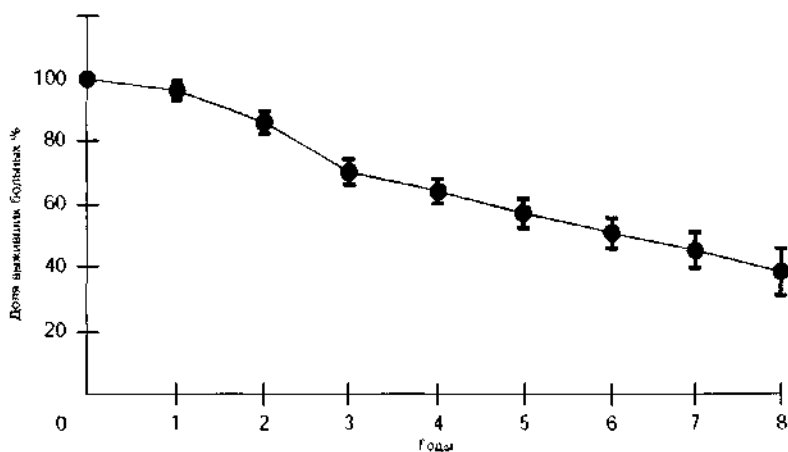


Рис. 2. Кривая общей выживаемости Капмана—Мейера (приводится с разрешения Wiley-Liss, Inc., филиал John Wiley & Sons, Inc.).

На кривой представлены значения стандартной ошибки средней в общей когорте

ются в исследование не одновременно, и многие выбывают до его окончания). Поэтому кривая выживаемости более точна на начальных этапах наблюдения; об этом должны свидетельствовать и более узкие доверительные интервалы в левой части кривой.

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

В найденном вами исследовании Oxfordshire Community Stroke Project [1] было показано, что абсолютный риск смерти в течение первого года после перенесенного инсульта составляет 31% (при 95% ДИ от 27 до 34%), а на протяжении следующих 4 лет — около 4% в год. У выживших больных риск развития повторного инсульта составлял 8,6% (при 95% ДИ от 6,5 до 10,7%) в течение первых 6 мес; 4,6% (при 95% ДИ от 2,6 до 6,6%) в течение следующих 6 мес; 6,7% (при 95% ДИ от 2,7 до 7,3%) в течение 2-го года; 5,0% (при 95% ДИ от 1,0 до 5,6%) в течение 3-го года; 3,3% (значения 95% ДИ не указаны) в течение 4-го года и 1,3% (значения 95% ДИ не указаны) в течение 5-го года после первого инсульта. В 61% случаев повторный инсульт был достаточно

тяжелым и лишал больных возможности выполнять повседневные функции в течение >7 дней, в то время как у 24% больных его симптомы исчезали менее чем за неделю.

Исследователи проанализировали, в какой степени риск развития повторного инсульта зависел от следующих факторов: пол, возраст, курение, наличие сахарного диабета, фибрилляции предсердий, сердечной недостаточности, преходящих нарушений мозгового кровообращения, стенокардии и инфаркта миокарда, перемежающейся хромоты, гипертонии. Показано, что из всех этих факторов только курение повышало риск развития повторного инсульта (ОШ=1,66 при 95% ДИ от 1,10 до 2,51).

КАК ПРИМЕНИТЬ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ НА ПРАКТИКЕ?

СООТВЕТСТВУЮТ ЛИ ХАРАКТЕРИСТИКИ УЧАСТНИКОВ ИССЛЕДОВАНИЯ И МЕТОДЫ ИХ ЛЕЧЕНИЯ УСЛОВИЯМ ВАШЕЙ КЛИНИЧЕСКОЙ ПРАКТИКИ?

В статье характеристики участников должны быть представлены достаточно подробно, чтобы их можно было сравнить с характеристиками ваших больных. Кроме того, важно знать о критериях включения. В исследованиях, посвященных изучению прогноза, редко сообщается о таком важном для клинических исходов факторе, как тактика лечения. Зачастую она существенно различается в разных лечебных учреждениях, да и в целом постоянно изменяется (по мере появления новых методов или периодического роста популярности старых методов). Любое лечебное вмешательство характеризуется определенным соотношением пользы и вреда; в зависимости от того, какие вмешательства мы применяем, клинические исходы в общей популяции больных могут улучшаться или ухудшаться.

БЫЛО ЛИ НАБЛЮДЕНИЕ ДОСТАТОЧНО ДЛИТЕЛЬНЫМ?

Поскольку с момента появления симптомов до развития определенного клинического исхода может проходить много времени, длительность наблюдения должна быть достаточной, чтобы выявить интересующие нас исходы. Например, у некоторых женщин рецидив рака молочной железы может возникнуть через много лет после установления диагноза и успешного лечения [11]. В материалах исследования, посвященного изучению прогноза и соответствующего всем критериям достоверности (см. таблицу), может содержаться достаточно точная и полезная

информация о возможном развитии ранних клинических исходов. Однако эти данные будут практически непригодны, если вашего больного в большей степени интересует долгосрочный прогноз.

МОЖНО ЛИ ПРИМЕНЯТЬ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ В ВАШЕЙ ПРАКТИКЕ?

На основании исследований, посвященных прогнозу, часто принимаются решения о тактике лечения (см. главу 2ВЗ на компакт-диске). Знание о возможном течении заболевания позволяет решить, стоит ли в данном случае вообще начинать лечение. Так, применение варфарина существенно снижает риск возникновения инсульта у больных с фибрилляцией предсердий неревматической природы и показано большинству таких больных [12]. Однако в одном из исследований [13] было показано, что частота развития инсульта при "изолированной" фибрилляции предсердий, которая наблюдается у больных в возрасте ≤ 60 лет без каких-либо признаков других сердечно-легочных заболеваний, за 15 лет составила всего 1,3%. В большинстве таких случаев риск, связанный с проведением антикоагулянтной терапии, может во много раз превышать ее пользу.

Даже если данные о прогнозе не пригодятся вам при выборе соответствующей тактики лечения, их можно использовать при общении с обеспокоенным больным или его родственниками. При бессимптомном течении некоторых заболеваний (например, грыжи пищеводного отверстия диафрагмы или дивертикулеза толстой кишки) общий прогноз настолько благоприятный, что их порой даже не относят к истинным заболеваниям [14]. С другой стороны, данные о крайне плохом прогнозе врач может использовать в качестве отправной точки при подготовке больного и его родственников к мысли о неизбежном неблагоприятном исходе.

ПРИНЯТИЕ КЛИНИЧЕСКОГО РЕШЕНИЯ

Вернемся к основному клиническому примеру. Ознакомившись с критериями достоверности, вы убедились, что при оценке риска развития повторного инсульта авторы когортного исследования свели к минимуму риск возникновения систематической ошибки [1]. По возрасту и типу инсульта ваш больной не отличается от участников исследования, поэтому в данном случае полученные результаты вполне применимы. Минимальная длительность наблюдения в исследовании составляла 2 года, а некоторые больные наблюдались на протяжении 6,5 лет, что позволило оценить частоту возникновения клинических исходов в течение почти 5 лет.

С учетом того, что больной жив в течение года после перенесенного инсульта, на протяжении последующих 4 лет риск смерти у него составит примерно 16% и риск развития повторного инсульта тоже около 16%. Эти результаты характеризуются достаточно узким диапазоном ДИ, поэтому можно быть уверенным в точности их оценки. Поскольку применение аспирина снижает риск развития повторного инсульта примерно на 25%, для предотвращения одного повторного инсульта необходимо на протяжении 4 лет лечить 25 больных, сходных по характеристикам с вашим больным, т.е. ЧБНЛ = $100 / (16 - 12\%)$. Учитывая низкую токсичность аспирина, принимаемого ежедневно в низких дозах, можно с уверенностью рекомендовать вашему больному такое лечение.

Хотя больной не очень доволен состоянием своего здоровья, в целом он смотрит на жизнь оптимистически. Он рад узнать, что на протяжении последующих 4 лет вероятность остаться в живых без дальнейшего ухудшения повседневной активности составляет почти 70%. Он использует эту информацию, чтобы убедить свою жену по-прежнему содержать два дома, по крайней мере в течение ближайшего времени.

Литература

1. *Burn J., Dennis M., Bamford J., Sandercock P., Wade D., Warlow C.* Long-term risk of recurrent stroke after a first-ever stroke: the Oxfordshire Community Stroke Project. *Stroke* 1994;25:333—337.
2. *Cook D.J., Fuller H.D., Guyatt G.H., et al.* Risk factors for gastrointestinal bleeding in critically ill patients. *N Engl J Med* 1994;330:377—381.
3. *Ellenberg J.J., Nelson K.B.* Sample selection and the natural history of disease: studies of febrile seizures. *JAMA* 1980;243:1337-1340.
4. *Kriel R.L., Krach L.E., Jones-Saete C.* Outcome of children with prolonged unconsciousness and vegetative states. *Pediatr Neurol* 1993;9:362—368.
5. *Pincus T., Brooks R.H., Callahan L.F.* Prediction of long-term mortality in patients with rheumatoid arthritis according to simple questionnaire and joint count measures. *Ann Intern Med* 1994; 120:26—34.
6. *Dawber T.R., Kannel W.B., Lyell L.P.* An approach to longitudinal studies in a community: The Framingham study. *Ann NY Acad Sci* 1963; 107:539.
7. *Sims A. C.* Importance of high tracing-rate in long-term medical follow up studies. *Lancet* 1973;2:433.
8. *Walsh J.S., Welch G., Larson E.B.* Survival of outpatients with Alzheimer-type dementia. *Ann Intern Med* 1990;113:429-434.
9. *ISIS-2 (Second International Study of Infarct Survival) Collaborative Group.* Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17.187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet* 1988;2:349-360.
10. *Dorey F., Amstutz H.* The validity of survivorship analysis in total joint arthroplasty. *J Bone J Surg AM* 1989;71A(4):544—548.
11. *Early Breast Cancer Trialists' Collaborative Group.* Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy: 133 randomised trials involving 31000 recurrences and 24000 deaths among 75000 women. *Lancet* 1992;339:1-15.
12. *Segal J.B., McNamara R.L., Miller M.R., et al.* Prevention of thromboembolism in atrial fibrillation: a meta-analysis of trials of anticoagulants and antiplatelet drugs. *J Gen Intern Med* 2000; 15:56—67.
13. *Kopecky S.L.* The natural history of lone atrial fibrillation: a population-based study over three decades. *N Engl J Med* 1987;317:669-674.
14. *Meador C.K.* The art and science of iron disease. *N Engl J Med* 1965;272:92.

10**ОБОБЩЕНИЕ
ДОКАЗАТЕЛЬСТВ**

Э. Оксман, Г. Гайятт, Д. Кук, В. Монтори

В подготовке данной главы принимали также участие члены Рабочей группы по разработке и внедрению принципов доказательной медицины Р. Хатала, Э. МакКиббон, Т. Гринхол, Дж. Крэйг, Р. Йешке

В ЭТОМ РАЗДЕЛЕ

Поиск доказательной информации

Достоверны ли результаты найденного обзора?

Рассматривается ли в обзоре интересующий вас конкретный клинический вопрос?

Проведен ли детальный и исчерпывающий поиск исследований по данной тематике?

Характеризовались ли оригинальные исследования, включенные в обзор, высоким методологическим качеством?

Воспроизводимы ли результаты оценки оригинальных исследований?

Каковы результаты обзора?

Были ли сходными результаты разных первичных исследований?

Каковы общие результаты обзора?

Насколько точны результаты обзора?

Как применить результаты обзора на практике?

Как правильно интерпретировать результаты обзора, чтобы наилучшим образом использовать их на практике?

Оценивались ли в обзоре все клинически значимые исходы?

Сопоставимы ли преимущества изучаемого вмешательства с материальными затратами и потенциальным риском?

Принятие клинического решения

ВОЗМОЖНАЯ СИТУАЦИЯ

Следует ли применять тромболитические препараты при остром ишемическом инсульте?

Представьте себя одним из невропатологов, работающих в учебном медицинском центре. До сих пор в этом центре при лечении больных с острым ишемическим инсультом тромболитическая терапия не применялась. Однако некоторые из ваших коллег убеждены, что проведение тромболизиса у таких больных снизит отдаленную смертность и частоту развития осложнений; они считают, что применение тканевого активатора плазминогена (ТАП) целесообразно в тех случаях, когда с момента возникновения первых симптомов инсульта до поступления больного в стационар прошло не более нескольких часов. Другие ваши коллеги гораздо более сдержанно относятся к назначению тромболитических препаратов в такой ситуации. Вы еще не составили собственного мнения по данному вопросу.

Группа невропатологов решила подойти к решению вопроса формально. Вас включили в состав комитета, члены которого должны собрать достоверную информацию и подготовить предварительный отчет. Эту работу было решено начать с поиска систематического обзора.

ПОИСК ДОКАЗАТЕЛЬНОЙ ИНФОРМАЦИИ

Вы начинаете поиск с просмотра материалов в базе данных Cochrane Library (1-й выпуск за 2000 г.). После введения терминов "stroke" ("инсульт") и "tissue plasminogen activator" ("тканевой активатор плазминогена") в одном из трех основных разделов этой базы (Cochrane Database of Systematic Reviews) находите систематический обзор по интересующей вас тематике и отмечаете, что в последний раз он обновлялся в июле 1999 г. [1].

При поиске ответа на большинство конкретных клинических вопросов можно найти несколько систематических обзоров. Как и при проведении оригинального исследования, при обобщении результатов нескольких исследований очень важно использовать строгие методы, позволяющие предотвратить возникновение систематической и случайной ошибок. В обычных обзорах литературы, публикуемых в медицинских

журналах или руководствах, приводится, как правило, множество сведений о каком-либо заболевании или состоянии. Эти обзоры могут включать в себя обсуждение разных аспектов этиологии, диагностики, прогнозирования или лечения; в них может освещаться целый ряд клинических и теоретических вопросов.

Например, обзорная статья или глава книги, посвященные бронхиальной астме, могут содержать разделы об этиологии, диагностике, прогнозе, различных методах лечения и профилактики этого заболевания. Как правило, авторами таких обзоров практически не применяется систематический подход к формулированию обсуждаемого вопроса, поиску и обобщению имеющихся данных. Однако студенты-медики и практикующие врачи, которых интересует базовая информация, находят эти обзоры полезными, поскольку они позволяют составить общее представление об определенной проблеме (см. главу 2 *Поиск доказательной информации*).

К сожалению, даже опытные составители подобных обзоров часто приходят к противоречивым выводам; их рекомендации порой являются устаревшими или не соответствуют результатам испытаний, проведенных на высоком методологическом уровне [2]. Отчасти это обусловлено применением несистематических методов отбора и обобщения имеющихся данных. Так, в одном из исследований было показано, что самостоятельная, несистематическая оценка достоверности данных обратно пропорциональна методологическому качеству обзора [3].

В этом разделе книги основное внимание будет уделено обзорам, посвященным решению частных клинических вопросов. Врачи, которые ищут ответ на вопрос, относящийся к тактике ведения больного, сочтут эти обзоры чрезвычайно полезными (см. главу 2 *Поиск доказательной информации*).

Иногда авторы считают синонимами термины " обзор", "систематический обзор" и "мета-анализ". Под *обзором* мы понимаем любой обзор, посвященный решению определенного клинического вопроса; под *систематическим обзором* — любой обзор, посвященный решению определенного клинического вопроса и проводимый с использованием методов, которые снижают вероятность возникновения систематической ошибки; под *мета-анализом* — обзор, в котором используются количественные методы обобщения результатов. В процессе подготовки систематического обзора исследователям приходится принимать множество решений, в том числе при формулировании цели обзора, выявлении, отборе и критической оценке оригинальных исследований (в дальнейшем мы будем называть их *первичными исследованиями*), отборе и обра-

ботке (количественной или качественной) необходимых данных, формулировании выводов. Необходимо четко спланировать проведение систематического обзора во избежание ошибок и точно описать применяемые методы, чтобы читатель мог оценить достоверность результатов этого обзора.

За последнее десятилетие появилось много статей, в которых описываются методы проведения систематических обзоров, в том числе отчеты об исследованиях, предоставляющих эмпирическую базу для оценки качества методов обобщения данных [4—6]. В этой главе выделены ключевые моменты оценки результатов систематического обзора, важные для врача, которому необходимо принять клиническое решение относительно ведения конкретного больного.

Используя рекомендуемый подход к оценке данных, вы должны хорошо представлять себе процесс проведения систематического обзора. На рисунке (см. стр. 232) показано, что процесс начинается с формулирования клинического вопроса, а это напрямую связано с определением критериев включения первичных исследований в обзор, т.е. с описанием популяции, вмешательств или потенциально неблагоприятных воздействий, а также изучаемых клинических исходов (см. главу 2 *Поиск доказательной информации*). Должны быть определены также минимальные требования к методологическому качеству первичных исследований. Например, в систематические обзоры, посвященные оценке эффективности лечения, обычно включаются только рандомизированные контролируемые испытания (РКИ).

После точного определения критериев включения и исключения первичных исследований авторы обзора должны провести всесторонний поиск, в результате чего будет найдено большое число названий и рефератов публикаций по данной тематике. Применение критериев отбора к этим названиям и рефератам существенно сократит список публикаций; тогда уже можно изучать полный текст этих публикаций, снова применяя критерии отбора. После этого авторы оценивают методологическое качество публикаций о каждом из оставшихся исследований и отбирают необходимые данные. В конце концов эти данные обобщают, применяя при необходимости методы количественного анализа или мета-анализ. Анализ данных включает в себя выявление различий между первичными исследованиями и попытку объяснить причину различий в их результатах (т.е. изучение возможной гетерогенности исследований), получение общих результатов, оценку их точности и достоверности. Ниже приводится рекомендуемый подход к оценке достоверности ре-

Формулирование клинического вопроса

Определите критерии включения и исключения первичных исследований с учетом:
популяции,
вмешательства или потенциально неблагоприятного воздействия,
клинического исхода,
методологического качества исследований

Сформулируйте предварительную гипотезу, объясняющую гетерогенность результатов первичных исследований

Поиск доказательной информации

Определите источники информации базы данных, контакты с экспертами, представителями финансирующих организаций и фармацевтических компаний, ручной поиск, личные подборки данных, реестры, библиографические списки в найденных статьях

Определите критерии, ограничивающие поиск- временные рамки, возможность включения неопубликованных данных, язык публикации

Отберите названия и рефераты публикаций по интересующей тематике

Применение критериев включения и исключения

Примените критерии включения и исключения к названиям и рефератам найденных публикаций

Найдите полный текст публикаций, название и реферат которых отвечают критериям включения

Примените критерии включения и исключения к полному тексту публикаций

Проведите окончательный отбор публикаций

Определите, насколько согласованным было мнение авторов, принимавших участие в отборе публикаций

Отбор данных для систематического обзора

Проведите отбор данных об участниках, вмешательствах в основных и контрольных группах, структуре исследований

Отберите интересующие вас результаты первичных исследований

Проведите отбор данных, характеризующих методологическое качество первичных исследований

Определите, насколько согласованным было мнение авторов, принимавших участие в оценке достоверности результатов первичных исследований

Анализ отобранных данных

Определите метод обобщения результатов первичных исследований

Обобщите результаты (если это возможно)

Решите вопрос о возможности дополнительного получения отсутствующих данных

Оцените гетерогенность первичных исследований, проведите анализ чувствительности и анализ в подгруппах

Оцените возможность наличия систематической ошибки, связанной с предпочтительной публикацией положительных результатов исследований

зультатов любого обзора и использованию его результатов в собственной клинической практике (табл. 1).

Таблица 1. Рекомендуемый подход к оценке качества обзорных статей и их применению в повседневной практике

Достоверны ли результаты обзора?

Рассматривается ли в обзоре интересующий вас конкретный клинический вопрос?

Проведен ли детальный и исчерпывающий поиск исследований по данной тематике?

Характеризовались ли оригинальные исследования, включенные в обзор, высоким методологическим качеством?

Воспроизводимы ли результаты оценки оригинальных исследований?

Каковы результаты обзора?

Были ли сходными результаты разных первичных исследований?

Каковы общие результаты обзора?

Насколько точны результаты обзора?

Как применить результаты обзора на практике?

Как правильно интерпретировать результаты обзора, чтобы наилучшим образом использовать их на практике?

Оценивались ли в обзоре все клинически значимые исходы?

Сопоставимы ли преимущества изучаемого вмешательства с материальными затратами и потенциальным риском?

ДОСТОВЕРНЫ ЛИ РЕЗУЛЬТАТЫ ОБЗОРА?

РАССМАТРИВАЕТСЯ ЛИ В ОБЗОРЕ ИНТЕРЕСУЮЩИЙ ВАС КОНКРЕТНЫЙ КЛИНИЧЕСКИЙ ВОПРОС?

Представьте себе систематический обзор, авторы которого обобщили информацию обо всех возможных методах консервативного лечения при всех существующих видах рака, чтобы дать общую оценку влияния такого лечения на уровень смертности. Представьте себе еще один обзор, авторы которого обобщили данные о влиянии терапии любыми антиагрегантами (включая аспирин, сульфинпиразон и дипиридамо́л) в любых дозах на смертность и частоту развития тяжелых тромбоемболических осложнений (включая инфаркт миокарда, инсульт, острую недостаточность кровообращения в нижних конечностях) при атероскле-

розе с клиническими проявлениями (вне зависимости от того, какие артерии преимущественно поражены — сердца, головного мозга или нижних конечностей). И наконец, представьте себе третий обзор, в котором оценивается эффективность приема аспирина в разных дозах для профилактики ишемического инсульта после преходящего нарушения мозгового кровообращения в бассейне сонных артерий. Вряд ли врачи сочтут первый из этих обзоров полезным, они решат, что его тематика слишком широка. По той же причине большинству врачей не понравится и второй обзор, хотя заслуживающая доверия группа опытных исследователей сочла эту тему правомерной и опубликовала результаты соответствующего мета-анализа в одном из ведущих медицинских журналов [7—9]. Тема третьего обзора удовлетворит практически всех врачей, хотя некоторые из них могут выразить сомнения по поводу обобщения результатов исследований, в которых аспирин применялся в разных дозах.

Что делает тему систематического обзора слишком широкой или слишком узкой? Ранее уже отмечалось, что четкое определение популяции, вмешательства или потенциально вредного воздействия, а также определенных клинических исходов необходимо при формулировании клинической задачи (см. главу 2 *Поиск доказательной информации*). Оценивая значимость клинического вопроса, решению которого посвящен обзор, врач должен обратить внимание на определение изучаемого заболевания (от этого зависит, будет ли эффект данного вмешательства одинаковым у разных больных). Одновременно следует обращать внимание и на другие составляющие основного клинического вопроса. Например, такова ли биологическая сущность заболевания или состояния, что при использовании разных вмешательств и оценке разных клинических исходов предполагается более или менее сходный эффект лечения? Аналогичные вопросы уместны и при анализе результатов исследований других типов. Например, будут ли сопоставимыми отношения правдоподобия результатов диагностического теста у разных больных при использовании разных модификаций изучаемого теста и "золотого стандарта" (см. главу 8 *Диагностические методы*) [10]?

Практикующие врачи откажутся рассматривать первый систематический обзор, поскольку знают, что при определенных разновидностях рака некоторые методы лечения эффективны, а применение других может нанести вред. При обобщении результатов исследований, в ходе которых применялись разные виды лечения при разных видах рака, рассчитанный размер эффекта будет бессмысленным по сути и неприменимым ни к одному из рассматриваемых вмешательств.

Врачи, отвергающие второй обзор, могут аргументировать свое решение тем, что лечебный эффект разных антиагрегантов может в значительной степени различаться из-за особенностей их фармакологических свойств. Они могут также указать на существенные различия в патофизиологии атеросклероза сосудов сердца, головного мозга и нижних конечностей. Кроме того, поскольку врачам приходится принимать клинические решения относительно ведения конкретного больного, они обычно склонны искать обзоры о вмешательстве, которое применялось у больных, в наибольшей степени соответствующих по своим характеристикам этому конкретному больному.

Врачи, считающие написание второго обзора целесообразным, могут возразить, что патофизиологические механизмы атеросклероза разной локализации и механизмы действия разных антиагрегантов в принципе похожи; поэтому и размер лечебного эффекта должен быть примерно одинаковым. Кроме того, более точные предположения о том, каким будет размер эффекта при лечении конкретного больного, бывает проще сделать на основании результатов обзора с более широкой темой, чем с более узкой. Это обусловлено тремя причинами.

Во-первых, когда в обзоре рассматриваются ограниченная популяция больных (например, с наиболее тяжелым течением атеросклероза), определенное вмешательство (например, прием аспирина в какой-то одной дозе при преходящем нарушении мозгового кровообращения) или ограниченный круг исследований (например, только те из них, материалы которых опубликованы на английском языке), повышается вероятность получения ошибочных результатов [11, 12]. Во-вторых, при рассмотрении данных об определенной подгруппе больных высока вероятность возникновения систематической ошибки, если на критерии, используемые при формировании подгруппы, влияют и другие факторы, определяющие эффект лечения. Так, критерием отбора исследований могут стать только определенные характеристики больных, даже если методологическое качество отбираемых исследований будет сравнительно низким; это может привести к ошибочной оценке лечебного эффекта. В-третьих, включение в обзор максимального числа исследований по определенной тематике позволяет объяснить возможные различия в их результатах (с учетом различий в характеристиках участников, применяемых вмешательствах или способах оценки клинических исходов). Таким образом, на основании более широких обзоров легче оценить эффект конкретного вмешательства при конкретном заболевании и судить о достоверности результатов анализа в подгруппах, чем на основании обзоров

по узкой тематике, в которых результаты такого анализа нередко сомнительны (см. главу 2Е на компакт-диске).

Вернемся к третьему примеру. Большинство врачей согласятся с тем, что фармакологическое действие аспирина скорее всего будет одинаковым при преходящем нарушении мозгового кровообращения с преимущественной ишемией участков правого или левого полушария головного мозга; у больных старше 75 лет и более молодых участников; у мужчин и у женщин; при использовании препарата в любых дозах; при наблюдении в течение 1 года или 5 лет; после инсульта, выявленного лечащим врачом или комиссией экспертов. При едином механизме действия препарата размер эффекта лечения тоже скорее должен быть одинаковым. Однако даже в рамках этого более узко поставленного клинического вопроса могут оставаться другие различия в характеристиках больных, в применяемых вмешательствах, в критериях оценки клинических исходов и методах проведения первичных исследований. Поэтому остается актуальной и необходимость выявления возможных причин несоответствия результатов первичных исследований (см. главу 2Е на компакт-диске). Получается, что ответить на вопрос о возможности обобщения результатов разных исследований нельзя до тех пор, пока не будут рассмотрены сами результаты. Если в нескольких исследованиях эффект вмешательства будет одинаковым, результаты таких исследований можно обобщать; в противном случае возникнет вопрос о достоверности любых выводов, которые будут сделаны после обобщения.

Таким образом, врач должен решить, в каких случаях эффект вмешательства будет одинаковым (несмотря на различия в характеристиках больных, особенностях потенциально вредных воздействий или применяемых вмешательств, а также в критериях оценки клинических исходов). Это возможно только в тех случаях, когда в обзоре четко сформулированы критерии включения первичных исследований с перечислением соответствующих характеристик больных, вмешательств и клинических исходов. Кроме того, должны быть указаны критерии, позволяющие оценить соответствие темы включенных исследований основному клиническому вопросу обзора. В целом эти критерии имеют много общего с критериями, применяемыми для оценки достоверности первичных исследований (табл. 2). Наличие таких четких критериев не только облегчает для читателя оценку значимости основного клинического вопроса, но и снижают вероятность того, что в обзор включались преимущественно те исследования, результаты которых свидетельствуют в пользу заранее сформулированной авторами концепции.

Таблица 2. Рекомендуемый подход к отбору публикаций, результаты которых могут быть наиболее достоверными [3]

Направленность исследования	Вопросы, на которые следует ответить
Лечение	Применялась ли рандомизация больных? Все ли больные наблюдались до конца исследования?
Диагностика	Была ли выборка репрезентативной по отношению ко всей популяции таких больных? Был ли диагноз подтвержден с помощью "золотого стандарта", применение которого не зависело от наличия определенной симптоматики?
Оценка риска	Продемонстрировали ли исследователи схожесть по всем известным факторам, способным повлиять на клинический исход; были ли учтены имеющиеся различия при анализе данных? Была ли достаточной доля больных, наблюдавшихся до конца исследования?
Прогноз	Была ли выборка репрезентативной, характеристики участников — четко определенными, а стадия заболевания при включении в исследование — одинаковой? Была ли достаточной доля больных, наблюдавшихся до конца исследования?

Наличие систематической ошибки, связанной с отбором публикаций, характерно как для систематических обзоров, так и для отчетов об исследованиях (в которых раздел, посвященный обсуждению полученных результатов, часто содержит сравнение с результатами других исследований). P.C. Gotsche, например, проанализировал библиографические списки в отчетах об испытаниях новых нестероидных противовоспалительных средств при ревматоидном артрите [13]. Из 77 отчетов, в которых авторы ссылались на результаты других испытаний определенного препарата, примерно в 44 (60%) случаях преимущественно упоминались испытания с положительными результатами (т.е. подтверждающие эффективность и безопасность нового препарата). В 22 отчетах о контролируемых испытаниях гипохолестеринемических препаратов U. Ravnskov выявил наличие той же систематической ошибки, обусловленной преимущественным цитированием исследований с положительными результатами [14]. При анализе 26 отчетов об РКИ, опубликованных в ведущих медицинских журналах, M. Clarke и I. Chalmers нашли только две статьи, в которых обсуждение результатов проводилось в контексте

недавно обновленного систематического обзора [15]. Читатель должен проявлять осторожность при интерпретации результатов исследования, не включенного в систематический обзор.

ПРОВЕДЕН ЛИ ДЕТАЛЬНЫЙ И ИСЧЕРПЫВАЮЩИЙ ПОИСК ИССЛЕДОВАНИИ ПО ДАННОЙ ТЕМАТИКЕ?

Авторы систематического обзора должны провести детальный и исчерпывающий поиск исследований, соответствующих критериям включения. Для этого следует использовать библиографические базы данных, такие как MEDLINE и EMBASE, Кокрановский реестр контролируемых испытаний (Cochrane Controlled Trials Register; содержит данные более чем о 250 000 РКИ) и базы данных исследований, проводимых на момент поиска [16]. Следует также изучить библиографические списки в найденных статьях и установить личные контакты со специалистами в интересующей области медицины. Важным может оказаться ознакомление с недавно опубликованными материалами научных конгрессов, а также поиск в редко используемых базах данных (содержащих, например, тезисы докторских диссертаций или материалы исследований, которые финансируются фармацевтическими компаниями). При перечислении этих источников информации становится очевидной недостаточность поиска только в базе данных MEDLINE. Если авторы обзора не указали способы поиска данных, читателю сложно будет оценить вероятность того, что некоторые исследования, содержащие важную информацию, не были найдены.

Существуют две причины, по которым желателен личный контакт авторов обзора со специалистами в определенных областях медицины. Во-первых, это позволяет выявить опубликованные материалы исследований, которые могли быть пропущены при первичном поиске (включая сданные в печать статьи и материалы, которые еще не прошли индексацию). Во-вторых, такой контакт помогает найти неопубликованные материалы исследований и включить их в обзор для предотвращения *систематической ошибки, связанной с преимущественной публикацией положительных результатов*.

Такая систематическая ошибка возникает, когда решение о публикации статьи зависит от результатов исследования или от их статистической значимости. В ряде случаев отчеты об исследованиях, в которых эффективность вмешательства не подтвердилась, не публикуются. Если при проведении систематического обзора такие исследования не будут учитываться, оценка эффекта вмешательства может оказаться завышенной [17—21] (см. главу 2Е на компакт-диске).

При включении неопубликованных материалов в обзор авторам следует получить полный текст отчета, а затем провести такую же оценку достоверности, как и при рассмотрении опубликованных материалов. Авторы обзора могут также использовать статистические методы для определения вероятности наличия систематической ошибки, связанной с предпочтительной публикацией положительных результатов исследований и систематических ошибок других типов, хотя чувствительность этих методов ограничена [22]. При проведении систематических обзоров, включающих ограниченное число исследований с небольшими выборками, наиболее высока вероятность возникновения систематической ошибки, связанной с предпочтительной публикацией положительных результатов исследований; применять результаты таких обзоров следует с осторожностью. Результаты, которые выглядят слишком хорошими, могут на самом деле оказаться ошибочными.

Составители обзора могут пойти дальше обычных контактов с авторами оригинальных исследований и пригласить этих исследователей к сотрудничеству; в этом случае появляется возможность получить данные об отдельных участниках. Это повышает статистическую мощность обзора и достоверность его выводов.

ХАРАКТЕРИЗОВАЛИСЬ ЛИ ОРИГИНАЛЬНЫЕ ИССЛЕДОВАНИЯ, ВКЛЮЧЕННЫЕ
В ОБЗОР, ВЫСОКИМ МЕТОДОЛОГИЧЕСКИМ КАЧЕСТВОМ?

Даже если в обзор включались только РКИ, следует располагать информацией о их методологическом качестве. К сожалению, достоверность результатов исследования не может быть гарантирована одним только фактом внешнего рецензирования его материалов (см. главу 4 *Исследования, посвященные эффективности лечения*) [23]. В предыдущих главах уже было показано, что при оценке отчетов об оригинальных исследованиях врач должен в первую очередь выяснить достоверность представленных результатов; по этой же причине при рассмотрении систематического обзора необходимо сначала убедиться в достоверности включенных в него первичных исследований.

Различия в методах исследований могут стать причиной существенных различий в их результатах [24—26]. Например, в исследованиях с использованием менее строгих методов отмечена тенденция к завышению оценки эффективности лечения или профилактики [27]. Но даже при схожести результатов первичных исследований важно убедиться в их достоверности. Сходные результаты выглядят менее убедительно, если они получены в ходе исследований низкого, а не высокого методологического качества.

Особенно подозрительны сходные результаты обсервационных исследований. Врачи могут систематически отбирать больных с хорошим прогнозом для проведения изучаемой терапии; такая практика весьма распространена и не зависит от времени или места проведения исследования. Например, во всех обсервационных исследованиях, включенных в систематический обзор [28], сообщалось о том, что на фоне заместительной гормональной терапии относительный риск развития тяжелых сердечно-сосудистых осложнений снижается в среднем на 50%. Однако уже при проведении первого крупного РКИ было выявлено отсутствие подобного эффекта заместительной гормональной терапии [29].

Точных способов оценки методологического качества исследований не существует, однако читатель должен помнить, что при рассмотрении любого систематического обзора основное внимание следует уделить оценке достоверности результатов первичных исследований; при этом надо с осторожностью относиться к использованию специальных шкал [30, 31]. Некоторые исследователи используют длинные списки параметров, по которым оценивается методологическое качество, другие анализируют 3 или 4 ключевых аспекта. При решении вопроса о том, можно ли доверять результатам обзора, вам надо проверить, совпадают ли критерии, применявшиеся авторами обзора, с критериями, представленными в других главах этой книги (см. главы 4 *Исследования, посвященные эффективности лечения*, 6 *Диагностика*, 5 *Оценка риска* и 9 *Прогноз*). Авторы обзора должны применять эти критерии как при отборе исследований, так и при оценке достоверности результатов включенных исследований (см. рисунок и табл. 2).

ВОСПРОИЗВОДИМЫ ЛИ РЕЗУЛЬТАТЫ ОЦЕНКИ ОРИГИНАЛЬНЫХ ИССЛЕДОВАНИЙ?

Как уже упоминалось, при составлении систематического обзора необходимо определить, какие исследования в него включать, насколько достоверны их результаты и какие данные следует отбирать. Принятие таких решений требует экспертной оценки, при этом также не исключено возникновение случайных и систематических ошибок. Участие нескольких экспертов в принятии каждого решения предотвращает появление ошибок; если степень согласованности мнений экспертов была высокой, врач может с большей уверенностью доверять результатам обзора (см. главу 2С на компакт-диске).

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ ОБЗОРА

Вернемся к нашему основному примеру. В найденный вами Кокрановский систематический обзор включались испытания, в которых у участников был диагностирован ишемический инсульт (наличие геморрагического инсульта исключали с помощью компьютерной томографии) [1]. Все участники рандомизированно распределялись в группы тромболитической терапии или ее отсутствия; во всех испытаниях анализ данных, проводимый в зависимости от назначенного лечения, либо точно проводился, либо мог быть проведен (см. главу 2B1 на компакт-диске). Такой анализ подразумевает, что данные обо всех участниках рассматриваются в соответствии с первоначальным отнесением к той или иной группе, независимо от реально полученного лечения. Вы предполагаете, что результаты лечения могут существенно различаться при начале лечения в ранние или более поздние сроки, при более выраженных или незначительных неврологических нарушениях, на фоне применения разных тромболитических препаратов, при использовании различных способов оценки функционального статуса и при разной длительности наблюдения. Однако вы не знаете, в какой степени указанные факторы влияют на клинические исходы, и предполагаете, что обобщение данных обо всех больных, вмешательствах и клинических исходах может оказаться информативным.

Составители обзора проводили поиск в Кокрановском реестре контролируемых испытаний и базе данных EMBASE. Дополнительно они осуществляли ручной поиск в ряде журналов на японском языке, лично обращались к представителям 321 фармацевтической компании и к ведущим исследователям из Европы, США, Японии и Китая; посетили несколько международных симпозиумов по лечению инсульта; изучили библиографические ссылки из найденных статей. Похоже, им удалось выявить все исследования по данной тематике.

В 7 из 17 включенных в обзор исследований для сокрытия порядка отнесения больных к группе лечения или контроля использовался метод централизованной рандомизации. В 13 исследованиях слепой метод применялся в отношении больных и медицинского персонала; для этого тромболитические препараты и плацебо заранее расфасовывались в одинаковые внешне флаконы, которые затем были запечатаны. Однако из-за возможного развития геморрагических осложнений, связанных с использованием тромболитических препаратов, сокрытие информации от больных и врачей могло оказаться неэффективным, что повышало важность применения слепого метода в отношении лиц,

оценивающих клинические исходы. Тем не менее это было сделано лишь в 4 исследованиях. В обзоре не сообщалось о числе участников, которые не наблюдались до конца каждого из исследований.

Один из авторов обзора принимал решение о соответствии найденных исследований критериям включения; другой отбирал необходимые данные, затем уточнял их с помощью руководителей первичных исследований и исправлял все выявленные ошибки. В 10 исследованиях авторы обзора смогли оценить функциональный статус отдельных больных (в баллах по шкале Рэнкина). Если этот показатель составляет менее 2 баллов из 5 возможных, больной не нуждается в постороннем уходе [32], поэтому оценку, составляющую от 3 до 5 баллов, исследователи считали признаком неблагоприятного клинического исхода. В двух других исследованиях не удалось получить данных об отдельных больных, там в качестве такого признака была выбрана оценка, равная 2 баллам и более.

В целом методологическое качество найденного вами систематического обзора (и включенных в него исследований) можно считать достаточно высоким.

КАКОВЫ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ?

БЫЛИ ЛИ СХОДНЫМИ РЕЗУЛЬТАТЫ РАЗНЫХ ПЕРВИЧНЫХ ИССЛЕДОВАНИЙ?

В большинстве систематических обзоров указываются важные различия между включенными первичными исследованиями в характеристиках больных, вмешательствах, критериях оценки клинических исходов и методах. Поэтому, как правило, нельзя ответить определенно на предварительный вопрос о том, будет ли эффективность лечения одинаковой у всех больных, при любых применявшихся вмешательствах и при рассмотрении любых клинических исходов.

К счастью, существует решение этой проблемы. По завершении обзора исследователи должны представить результаты в таком виде, чтобы можно было проверить обоснованность исходного допущения о схожести результатов отдельных исследований.

Прежде чем сделать вывод о схожести результатов, достаточной для расчета общих показателей эффективности лечения, которые будут применимы у участников всех исследований, при всех применявшихся вмешательствах и при любых клинических исходах, необходимо оценить два важных момента (см главу 2Е на компакт-диске) Во-первых, насколько

ко сходны значения размера эффекта при его *точной оценке* в ходе отдельных исследований? Чем больше различаются эти значения, тем больше сомнений должно вызывать у врача решение авторов обзора обобщить результаты.

Во-вторых, насколько выявленные различия между результатами отдельных исследований превосходят те, которые можно было бы ожидать при влиянии фактора случайности? Читатель может провести предварительную оценку, исследуя степень, в которой доверительные интервалы перекрывают друг друга. Чем она выше, тем большее доверие вызывает решение авторов обзора обобщить результаты этих исследований. Если доверительные интервалы вообще не перекрываются, то это свидетельствует о наличии существенных различий в результатах, требующих дополнительного разъяснения (см. главу 2Е на компакт-диске).

Практикующие врачи могут также обратиться к формальному статистическому анализу (тесты на гетерогенность), с помощью которого оценивается степень различия между разными выборками, группами или популяциями. Если при проведении теста на гетерогенность значение p низкое (например, $<0,05$), наблюдаемые различия в размере эффекта скорее всего нельзя объяснить лишь влиянием фактора случайности (см. главу 2В2 на компакт-диске). К сожалению, при более высоких значениях p (0,1 или даже 0,3) нельзя однозначно исключить наличие важной гетерогенности, поскольку тест обладает недостаточной статистической мощностью в тех случаях, когда число включенных в обзор исследований и размер выборки в каждом из них были небольшими. Поэтому при наличии существенных различий между отдельными исследованиями в величине лечебного эффекта следует крайне осторожно относиться к общему показателю, рассчитанному авторами обзора (даже если результаты теста на гетерогенность оказались статистически незначимыми). И наоборот, если различия между результатами отдельных исследований клинически незначимы, выявленная гетерогенность (даже статистически значимая) не имеет большого значения (см. главу 2Е на компакт-диске).

Авторы обзоров всегда должны пытаться объяснить различия в результатах включенных в обзор исследований. Возможными объяснениями могут быть различия в характеристиках больных (так, тромболитическая терапия при остром инфаркте миокарда может оказаться гораздо более эффективной у больных, поступивших в стационар вскоре после возникновения болей в грудной клетке, по сравнению с больными, госпитализированными позднее), вмешательствах (например, ТАП может

быть эффективнее стрептокиназы), методах оценки клинических исходов (например, размер эффекта может существенно различаться при оценке клинического исхода через месяц или через год после инфаркта миокарда) или методах исследования (например, размер эффекта может быть меньше в тех случаях, когда применяется слепой метод или до конца исследования наблюдается большее число участников). Но даже тщательный поиск причин неоднородности результатов, будучи желательным и даже необходимым, может дезинформировать читателя (см. главу 2Е на компакт-диске). Что же должен делать врач, если эти причины не вполне ясны? Мы рассмотрим этот вопрос при обсуждении применимости результатов систематического анализа на практике.

КАКОВЫ ОБЩИЕ РЕЗУЛЬТАТЫ ОБЗОРА?

В ходе клинических исследований собираются данные об отдельных больных. Из-за ограниченных возможностей человеческого мозга при обработке большого объема информации исследователи пользуются статистическими методами для ее обобщения и анализа. Авторы систематических обзоров собирают данные об отдельных исследованиях и также должны обобщить информацию, причем все чаще применяют для этого количественные методы.

Простое сопоставление числа исследований с отрицательными и положительными результатами недостаточно для их обобщения. При таком подходе крупные и небольшие исследования будут иметь одинаковый вес; и (пусть это выглядит неправдоподобно) результаты какого-либо определенного исследования один автор обзора может счесть положительными, а другой — отрицательными [33]. Например, небольшой и статистически незначимый эффект вмешательства можно расценить как положительный (с точки зрения его клинической значимости) и как отрицательный (по причине его статистической незначимости). Отмечается тенденция к игнорированию небольших, но важных с клинической точки зрения эффектов вмешательств, если результаты исследования (статистически незначимые, но важные с клинической точки зрения) рассматриваются как отрицательные [34]. Более того, читатель не может судить о размере эффекта при простом сопоставлении числа исследований с положительными и отрицательными результатами, даже если эти результаты классифицированы правильно с использованием дополнительных критериев.

Обычно при проведении мета-анализа отдельным исследованиям присваивают определенный вес (более крупные исследования характеризуются при этом большим весом). Таким образом, обобщенный ре-

зультат отражает средневзвешенное значение результатов индивидуальных исследований (см. главу 2Е на компакт-диске). Иногда больший или меньший вес исследования определяется его качеством, тогда исследования самого низкого методологического качества могут иметь нулевой вес и исключаться либо на этапе первичного анализа данных, либо при проведении вторичного анализа, в ходе которого оценивается степень зависимости между разными исходными допущениями и соответствующими результатами (анализ чувствительности).

Подход к рассмотрению общих результатов систематического обзора должен быть таким же, как и при оценке результатов первичных исследований. В систематическом обзоре, посвященном методам лечения, важно обратить внимание на такие показатели, как относительный риск и снижение относительного риска или отношение шансов (см. главу 2В2 на компакт-диске). В систематических обзорах, посвященных вопросам диагностики, следует обращать внимание на итоговую оценку отношения правдоподобия (см. главу & *Диагностические методы*).

Иногда критерии оценки клинических исходов, применявшиеся в разных исследованиях, похожи, но не идентичны. Например, для оценки функционального статуса могли быть использованы разные шкалы. Если характеристики больных и вмешательств в достаточной мере схожи, расчет среднего эффекта вмешательства (его влияния на функциональный статус) может быть оправданным. В частности, это можно сделать, если выразить результаты каждого из исследований в виде такого показателя, как размер эффекта [35]. *Размер эффекта* рассчитывается как разница между величиной изучаемого клинического исхода в основной и контрольной группе, деленная на стандартное отклонение. Расчет размера эффекта для каждого исследования позволяет выразить это различие в единицах стандартного отклонения, после чего авторы обзора могут подсчитать средневзвешенный размер эффекта даже в тех случаях, когда в разных исследованиях использовались разные критерии оценки клинического исхода.

Однако по размеру эффекта не всегда легко судить о его клинической значимости. Например, вы узнаете, что средневзвешенное значение размера эффекта составляет половину стандартного отклонения. Можно ли считать такой эффект клинически важным? Для ответа на подобный вопрос необходимо выразить результаты в другой форме, более пригодной для оценки клинической значимости (например, снова перевести размер эффекта в первичные показатели клинического исхода) [36]. Так, врачи могут хорошо представлять себе клиническую значимость различий в результатах теста с ходьбой, выявляемых при хро-

нических заболеваниях легких. Исследователи или авторы обзора могут рассчитать размер эффекта на основании целого ряда показателей, позволяющих судить о функциональном статусе (например, объединив результаты не только теста с ходьбой, но и лестничной пробы); однако для удобства врачей размер эффекта можно выразить только в виде результатов теста с ходьбой [37].

Хотя количественное обобщение результатов весьма желательно, оно не всегда возможно. В таких случаях авторы обзора должны представить таблицы или диаграммы, в которых сопоставляются результаты первичных исследований.

НАСКОЛЬКО ТОЧНЫ РЕЗУЛЬТАТЫ ОБЗОРА?

При анализе результатов нескольких первичных исследований можно рассчитать не только средний размер эффекта, но и его доверительный интервал, т.е. диапазон значений, которые с заранее определенной вероятностью (обычно равной 95%) включают в себя истинный размер эффекта (см. главу 2В2 на компакт-диске).

ИСПОЛЬЗОВАНИЕ РЕКОМЕНДУЕМОГО ПОДХОДА ПРИ ОЦЕНКЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

Снова вернемся к нашему основному примеру. В 4 включенных в обзор РКИ применялась стрептокиназа, в 3 РКИ — урокиназа, в 2 РКИ — проурокиназа, в 8 РКИ — ТАП. В 6 РКИ смертность в течение первых 7—10 дней после инсульта составила 16,6% среди больных, получавших тромболитические препараты, и 9,8% в контрольной группе (ОШ=1,85 при 95% ДИ от 1,48 до 2,32). При проведении теста на гетерогенность значение p приближалось к уровню статистической значимости; при включении в анализ только РКИ, в которых применялся ТАП, его значение было ниже и не достигало уровня статистической значимости (ОШ=1,24 при 95% ДИ от 0,85 до 1,81). При анализе данных, полученных в ходе 11 РКИ, авторы обзора показали, что тромболитическая терапия повышает частоту развития смертельных внутричерепных кровоизлияний с 1,0 до 5,4% (ОШ=4,15 при 95% ДИ от 2,96 до 5,84; результаты всех РКИ были достаточно сходными).

При итоговой оценке клинических исходов (в 6 РКИ — через 1 мес после инсульта, в 9 РКИ — через 3 мес, в 2 РКИ — через 6 мес) было выявлено повышение смертности с 15,9 до 19% (ОШ=1,31 при 95% ДИ от 1,13 до 1,52). Результаты разных РКИ оказались весьма неоднородными ($p < 0,01$).

Проведение тромболитической терапии снижало комбинированный показатель смертности и частоты наступления инвалидности (55,2% у больных, получавших тромболитические препараты, и 59,7% в контрольной группе; ОШ=0,83 при 95% ДИ от 0,73 до 0,94; результаты разных РКИ были однородными).

Авторы обзора попытались выяснить возможные причины гетерогенности данных о смертности. Статистически значимыми оказались только существенные различия в точечной оценке смертности (для урокиназы ОШ=0,71, для стрептокиназы ОШ=1,43, для ТАП ОШ=1,16); другие различия между применяемыми препаратами не достигли уровня статистической значимости. При сочетанном применении стрептокиназы и аспирина смертность была выше, чем при монотерапии стрептокиназой. Авторам не удалось выявить зависимость между частотой развития осложнений и смертностью в контрольной группе; при анализе результатов в основной группе было отмечено, что для более точной оценки взаимосвязи между исходной тяжестью инсульта и эффектами тромболитической терапии (как благоприятными, так и неблагоприятными) необходимы данные об отдельных больных. В нескольких РКИ больных распределяли по группам либо в течение первых 3 ч после появления симптомов инсульта, либо позднее; при этом различий в смертности не было выявлено.

КАК ПРИМЕНИТЬ РЕЗУЛЬТАТЫ НА ПРАКТИКЕ?

КАК ПРАВИЛЬНО ИНТЕРПРЕТИРОВАТЬ РЕЗУЛЬТАТЫ, ЧТОБЫ НАИЛУЧШИМ ОБРАЗОМ ИСПОЛЬЗОВАТЬ ИХ НА ПРАКТИКЕ?

Даже если в нескольких исследованиях истинный эффект вмешательства был одинаковым, результаты лечения будут различаться из-за влияния фактора случайности (см. главу 1 раздела II *Лечение и риск: причины ^достоверности результатов — систематические и случайные ошибки*). В ходе проведения систематического обзора эти кажущиеся различия могут еще увеличиться. Предположим, что в исследованиях, включавших более пожилых больных, из-за влияния случайных факторов был выявлен менее выраженный эффект лечения. Авторы обзора могут сделать ошибочный вывод о том, что лечение менее эффективно у пожилых больных. Чем больше анализов в подгруппах проводят авторы обзора, тем выше вероятность ложных заключений.

Врач может применить несколько критериев, позволяющих различить достоверные и недостоверные результаты анализа в подгруппах (см.

главу 2Е на компакт-диске). О достоверности выявленных различий можно говорить в тех случаях, когда выводы основаны на результатах сравнений в рамках отдельного исследования, а не между исследованиями; размер эффекта существенно различается между подгруппами; различия в размере эффекта характеризуются высоким уровнем статистической значимости (чем ниже значение p , тем выше достоверность выявленных различий); основная гипотеза о возможности различий между подгруппами была сформулирована еще до начала исследования, а при анализе данных проверялись всего несколько гипотез; различия между подгруппами выявляются во всех исследованиях; выявленные различия имеют косвенное (например, биологическое) обоснование. Если эти условия не соблюдены, результаты анализа в подгруппах скорее всего не заслуживают доверия. В таком случае логичнее считать, что общая оценка эффекта лечения, основанная на использовании различных вмешательств у разных больных, применима к конкретному больному и конкретному планируемому вмешательству в большей степени, чем данные об эффекте лечения в подгруппе.

Что делать, если и после анализа в подгруппах не удастся в полной мере объяснить гетерогенность результатов исследований? В таких случаях применяются разные подходы, в том числе и полный отказ от обобщения результатов отдельных исследований. Однако мы полагаем, что в отсутствие дополнительных исследований, результаты которых могли бы объяснить имеющиеся несоответствия, врачи должны ориентироваться на суммарные показатели, рассчитанные с использованием результатов исследований наиболее высокого методологического качества; это позволит лучше всего оценить влияние лечебного вмешательства или потенциально вредного воздействия на клинические исходы [38—40].

ОЦЕНИВАЛИСЬ ЛИ В ОБЗОРЕ ВСЕ КЛИНИЧЕСКИ ЗНАЧИМЫЕ ИСХОДЫ?

Чтение обзорных статей по узкой тематике полезно, поскольку в них чаще содержится достоверная информация; но это вовсе не означает, что следует игнорировать клинические исходы, которые не оценивались в этих обзорах. Так, к потенциальным преимуществам заместительной гормональной терапии относятся снижение риска возникновения переломов и развития ишемической болезни сердца, а к потенциальным недостаткам — повышение риска развития рака молочной железы или эндометрия. В обзорах по узкой тематике скорее всего будут отдельно представлены достоверные данные о влиянии заместительной гормональной терапии на какой-то один из этих четырех клинических исхо-

дов, однако при принятии клинического решения необходимо одновременно учитывать все указанные исходы.

В систематических обзорах часто не сообщается о побочных эффектах применявшихся лечебных вмешательств. Это отчасти обусловлено тем, что в отдельных исследованиях побочные эффекты либо оценивались по-разному, либо не оценивались вообще, что затрудняет количественное обобщение данных или их простое суммирование. Стоимость лечебных вмешательств (как дополнительный критерий их оценки) также редко анализируется в систематических обзорах.

СОПОСТАВИМЫ ЛИ ПРЕИМУЩЕСТВА ИЗУЧАЕМОГО ВМЕШАТЕЛЬСТВА
С МАТЕРИАЛЬНЫМИ ЗАТРАТАМИ И ПОТЕНЦИАЛЬНЫМ РИСКОМ?

В процессе совместного принятия клинического решения врач и больной осознанно или подсознательно сопоставляют преимущества любого планируемого вмешательства с его стоимостью и потенциальным риском (см. главу 11 *От доказательств к практике*). Несомненно, для полноправного участия в принятии решения о применении лечебного или профилактического вмешательства больной должен знать о причинах своего заболевания и о прогнозе; однако предоставление ему этой информации может принести не только пользу, но и вред. Так, знание о неблагоприятном влиянии на здоровье загрязненной атмосферы современного города может привести к снижению пребывания на воздухе, что потенциально полезно; но это может также создать дополнительные неудобства для городских жителей или даже способствовать развитию тревожности. Сообщив женщине с бессимптомным течением только что выявленной раковой опухоли о прогнозе, мы поможем ей наилучшим образом спланировать предстоящую жизнь, но такая информация может способствовать развитию тревожности, возникновению чувства отчужденности от здоровых людей и увеличению длительности периода, в течение которого она будет считать себя больной.

Обзорная статья с достоверными результатами представляет собой наилучшую основу для количественной оценки ожидаемых клинических исходов, но эти исходы следует рассматривать с учетом жизненных ценностей больного и его взглядов на возможные последствия принимаемого клинического решения. В конечном счете, сопоставляя преимущества и недостатки вмешательства, врач неизбежно принимает во внимание жизненные ценности больного (см. главу 11 *От доказательств к практике*); в процессе принятия индивидуального решения эти ценности^{ст} и Должен определять сам больной (см. главу 2F на компакт-диске).

Вернемся к нашему основному примеру. Члены комитета, в который вы были включены, решили, что на основании результатов систематического обзора можно уверенно сделать два вывода. Во-первых, проведение тромболитической терапии повышает риск развития внутричерепного кровоизлияния примерно в 3–6 раз (при точечной оценке — примерно в 4 раза). Если выразить этот эффект в абсолютных цифрах, то при назначении тромболитической терапии 23 больным у одного из них возникнет внутричерепное кровоизлияние. Во-вторых, применение тромболитических препаратов снижает комбинированный показатель смертности и частоты наступления инвалидности через 3 мес после инсульта примерно на 5–30% (при точечной оценке — на 17%, что соответствует ОШ=0,83). Если выразить этот эффект в абсолютных цифрах, то для предотвращения смерти или наступления тяжелой инвалидности у 1 больного тромболитическую терапию в течение 3 мес должны получать 22 больных. Вполне вероятен и третий вывод: сочетанное применение аспирина и тромболитических препаратов повышает риск возникновения внутричерепного кровоизлияния.

Члены комитета пришли к заключению, что многие вопросы еще остаются невыясненными. Действительно ли риск смерти в течение 3 мес после инсульта ниже на фоне приема ТАП, чем при оценке обобщенных данных обо всех вмешательствах? Каково влияние тромболитической терапии на комбинированный показатель смертности, частоты развития кровоизлияния или наступления инвалидности, оцениваемый с учетом тяжести инсульта и особенностей первичной симптоматики. Учитывая важность невыясненных вопросов, а также сохраняющиеся сомнения в целесообразности и своевременности внедрения этого метода, комитет постановил, что тромболитическую терапию следует назначать только тем больным, которые согласны на повышение риска ранней смерти в обмен на снижение риска развития дальнейших осложнений.

Литература

- 1 *Wardlaw J M, del Zoppo G, Yamaguchi T* Thrombolysis for acute ischaemic stroke Cochrane Database Syst Rev 2000,2 CD000213
- 2 *Antman EM, Lau J, Kupelmeck B, Mosteller F, Chalmers TC* A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts treatments for myocardial infarction JAMA 1992,268 240-248

- 3 Oxman AD, Guyatt G H The science of reviewing research *Ann N Y Acad Sci* 1993,703 125—133, discussion 133—134
- 4 Clarke M, Olsen K L , Oxman A D , eds The Cochrane Review Methodology Database In The Cochrane Library Oxford Update Software, 2000, issue 1
- 5 Clarke M, Oxman A D , eds Cochrane Reviewers' Handbook 4.0 [updated Jul\ 1999] In The Cochrane Library Oxford Update Software, 2000, issue 1
- 6 Egger M, Davey Smith G, Altman D G , eds Systematic Reviews in Health Care Meta-Analysis in Context 2nd ed London BMJ Books, 2000
- 7 Antiplatelet Trialists' Collaboration Collaborative overview of randomised trials of antiplatelet therapy, I prevention of death, myocardial infarction, and stroke by prolonged antiplatelet therapy in various categories of patients *BMJ* 1994,308 81-106
- 8 Antiplatelet Trialists' Collaboration Collaborative overview of randomised trials of antiplatelet therapy, II maintenance of vascular graft or arterial patency by antiplatelet therapy *BMJ* 1994,308 159—168
- 9 Antiplatelet Trialists' Collaboration Collaborative overview of randomised trials of antiplatelet therapy, III reduction in venous thrombosis and pulmonary embolism by antiplatelet prophylaxis among surgical and medical patients *BMJ* 1994,308 235-246
- 10 Irwig L , Tobieson A N, Gatsonis C, et al Guidelines for meta-analyses evaluating diagnostic tests *Ann Intern Med* 1994,120 667-676
- 11 Counselor C B, Clarke M J, Slattery J, Sandercock P A The miracle of DICE therapy for acute stroke fact or fictional product of subgroup analysis' *BMJ* 1994,309 1677-1681
- 12 Clarke M J, Halsey J D I C E 3 the need for cautious interpretation of metaanalyses Paper presented at First Symposium on Systematic Reviews Beyond the Basics, January 1998, Oxford
- 13 Gøtzsche P C Reference bias in reports of drug trials *Br Med J (Chn Res Ed)* 1987,295 654-656
- 14 Ravnskov U Cholesterol lowering trials in coronary heart disease frequency of citation and outcome *BMJ* 1992,305 15-19
- 15 Clarke M, Chalmers I Discussion sections in reports of controlled trials published in general medical journals islands in search of continents" *JAMA* 1998,280 280-282
- 16 The metaRegister of Controlled Trials (mRCT) Current Controlled Trials Available at www.controlled-trials.com/ Accessed January 31 2001
- 17 Dickersm K The existence of publication bias and risk factors for its occurrence *JAMA* 1990,263 1385—1389
- 18 Dickersm K, Mm i , Memert C L Factors influencing publication of research results *JAMA* 1992,267 374-378
- 19 Dickersm K How important is publication bias? A synthesis of available data *AIDS Educ Prev* 1997,9(suppl 1) 15-21
- 20 Stern J M , Simes R J Publication bias evidence of delayed publication in a cohort study of clinical research projects *BMJ* 1997 315 640-645
- 21 Ioannidis J P Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials *JAMA* 1998,279 281—286
- 22 Egger M, Davey Smith G, Schneider M, Minder C Bias in meta-analysis detected by a simple, graphical test *BMJ* 1997 315 629-634
- 23 Williamson J W, Goldschmidt P G , Colton T The quality of medical literature analysis of validation assessments In Bailar J C, Mosteller F, eds *Medical Uses of Statistics* 2nd ed Waltham NEJM Books, 1992 370-391
- 24 Horwitz R I Complexity and contradiction in clinical trial research *Am J Med* 1987,82 498—510
- 25 Detsky A S , Naylor C D , O'Rourke K, McGeer A L, Abbe K A Incorporating variations in the quality of individual randomized trials into meta-analysis *J Chn Epidemiol* 1992 45 255—265
- 26 Moher D , Pham B , Jones A , et al Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998,352 609—613

27. *Kun: R., Oxman A.D.* The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *BMJ* 1998;317:1185—1190.
28. *Stampfer M.J., Colditz G.A.* Estrogen replacement therapy and coronary heart disease: a quantitative assessment of the epidemiologic evidence. *Prev Med* 1991;20:47—63.
29. *Hulley S., Grady D., Bush T., et al.* Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. Heart and Estrogen/progestin Replacement Study (HERS) Research Group. *JAMA* 1998;280:605-613.
30. *Moher D., Jadad A.R., Nichol G., Penman M., Tugwell P., Walsh S.* Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials* 1995; 16:62—73.
31. *Juni P., Witschi A., Bloch R., Egger M.* The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999;282:1054-1060.
32. *de Haan R., Limburg M., Bossuyt P., van der Meulen J., Aaronson N.* The clinical meaning of Rankin 'handicap' grades after stroke. *Stroke* 1995;26:2027—2030.
33. *Glass G.V., McGaw B., Smith M.L.* Meta-analysis in Social Research. Beverly Hills: Sage Publications; 1981:18-20.
34. *Cooper H.M., Rosenhal R.* Statistical versus traditional procedures for summarizing research findings. *Psychol Bull* 1980;87:442-449.
35. *Rosenthal R.* Meta-analytic Procedures for Social Research. 2nd ed. Newbury Park: Sage Publications; 1991.
36. *Smith K., Cook D., Guyatt G.H., Madhavan J., Oxman A.D.* Respiratory muscle training in chronic airflow limitation: a meta-analysis. *Am Rev Respir Dis* 1992; 145:533—539.
37. *Lacasse Y., Wong E., Guyatt G.H., King D., Cook D.J., Goldstein R.S.* Meta-analysis of respiratory rehabilitation in chronic obstructive pulmonary disease. *Lancet* 1996;348:1115—1119.
38. *Peto R.* Why do we need systematic overviews of randomized trials? *Stat Med* 1987;6:233—244.
39. *Oxman A.D., Guyatt G.H.* A consumer's guide to subgroup analyses. *Ann Intern Med* 1992; 116:78—84.
40. *Yusuf S., Wittes J., Probstfeld J., Tyroler H.A.* Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *JAMA* 1991;266:93—98.

Г. Гайятт, Р. Хэйурд, УС. Ричардсон,
Л. Грин, М. Уилсон, Дж. Синклер, Д. Кук,
П. Глажиу, А. Детски, Э. Басе

В подготовке данной главы принимал также участие
П.Дж. Деверо

В ЭТОМ РАЗДЕЛЕ

Поиск доказательной информации

Клинические рекомендации, посвященные тактике лечения

Процесс разработки клинических рекомендаций

Систематические обзоры

Анализ принятия решений

Практические рекомендации

Оценка рекомендаций, посвященных тактике лечения

Учтены ли в рекомендациях все категории больных с определенным заболеванием, все вмешательства и клинические исходы?

Существует ли систематический обзор, в котором по каждому рассматриваемому клиническому вопросу обобщается информация о причинно-следственной связи между вмешательствами и клиническими исходами?

Был ли определен перечень предпочтений и жизненных ценностей, учитываемых при рассмотрении возможных клинических исходов?

Сообщается ли о степени достоверности предлагаемых рекомендаций?

Нужны ли вообще клинические рекомендации?

Принятие клинического решения

ВОЗМОЖНАЯ КЛИНИЧЕСКАЯ СИТУАЦИЯ

Применение варфарина при фибрилляции предсердий: целесообразно ли назначать этот препарат конкретной больной?

Представьте себя врачом общей практики, которому нужно оценить целесообразность назначения варфарина 76-летней больной с застойной сердечной недостаточностью и хронической формой фибрилляции предсердий (ФП). Вы только недавно начали наблюдать эту больную. Она страдает ФП в течение 10 лет; единственным антиагрегантом, который она принимала за это время, был аспирин. После 40 лет у больной была выявлена артериальная гипертония I степени, по поводу которой она принимает гидрохлортиазид и беназеприп. Судя по записям врача, у которого она наблюдалась ранее, в последние годы ее систолическое артериальное давление колебалось от 130 до 140 мм рт. ст., диастолическое — от 80 до 90 мм рт. ст. В настоящее время артериальное давление составляет 136/84 мм рт. ст. при частоте сердечных сокращений 76 уд/мин, что свидетельствует о достаточной эффективности гипотензивной терапии. У больной нет поражения клапанов сердца, сахарного диабета и других сопутствующих заболеваний; кроме того, она не курит.

Наличие ФП в течение длительного времени указывает на нецелесообразность применения кардиоверсии или антиаритмических препаратов. Больная живет одна. Хотя у нее еще ни разу не отмечались падения, сопровождавшиеся тяжелыми последствиями, вы предполагаете, что повышенный риск развития внутричерепного кровоизлияния на фоне приема варфарина может нивелировать преимущества этого препарата, применяемого для профилактики инсульта. Однако в разговоре с больной вы выясняете, что ее больше заботит предупреждение инсульта, чем риск развития тяжелого кровотечения. Больная "не любит лечиться", но готова терпеть неудобства, связанные с регулярными обследованиями для оценки эффективности антикоагулянтной терапии.

Стоит ли назначать антикоагулянты при ФП, не связанной с поражением клапанов сердца, и если да, то в какие сроки? Этот вопрос часто возникает в вашей повседневной практике, однако ваше мнение и мнение ваших коллег по данной проблеме зачастую расходятся. Все вы согласны с тем, что применение варфарина предупреждает развитие инсульта, однако некоторые Врачи не без оснований полага-

ют, что антикоагулянтная терапия слишком часто сопровождается развитием геморрагических осложнений. У нескольких ваших больных с ФП прием аспирина не предупредил развитие ишемического инсульта, а у двух больных, получавших варфарин, возникло тяжелое желудочно-кишечное кровотечение. Еще большее разногласие внесло недавнее заявление одного из ваших коллег, который всегда имеет особое мнение, о том, что при ФП неклапанного генеза лучше применять клопидогрел.

Вы пока не вносите никаких изменений в схему лечения, но к следующему визиту больной собираетесь проанализировать проблему и на следующей неделе обсудить ее на врачебной конференции.

ПОИСК ДОКАЗАТЕЛЬНОЙ ИНФОРМАЦИИ

Вам не очень хочется просматривать огромный объем оригинальных публикаций, в которых оценивались частота развития инсульта и частота возникновения кровотечений на фоне проведения антикоагулянтной терапии, однако вы надеетесь найти научно обоснованные рекомендации, которые, возможно, помогут вам и вашим коллегам. Вы решили, что для этого подходят два типа источников информации: практические рекомендации и результаты анализа принятия решений (анализ принятия решений).

В Интернете вы заходите на свой любимый поисковый сервер Google.com, вводите термин "practical guidelines" ("практические рекомендации"), и в списке результатов поиска замечаете ссылку на сайт "National Guidelines Clearinghouse" по адресу www.guidelines.gov. Название сайта ("Национальный центр анализа и синтеза информации") выглядит многообещающим, здесь же представлены данные о том, что этот сервер поддерживается Агентством США по проведению и оценке исследований в области здравоохранения (US Agency for Healthcare Research and Quality; AHRQ), ранее известным как Агентство по разработке стратегий и исследованиям в области здравоохранения (Agency for Health Care Policy and Research; AHCPR), которое, насколько вы помните, выработало целый ряд практических рекомендаций с использованием методов доказательной медицины [1].

Вы заходите на сайт National Guidelines Clearinghouse и видите заголовок "Guideline Syntheses" ("Синтез данных, представленных в практических рекомендациях") и подпись под ним, которая гласит, что дан-

ный раздел содержит: "...синтез данных, представленных в отобранных практических рекомендациях по сходной тематике. В процессе синтеза оценивались такие важные моменты, как тематика базовых публикаций; рассматриваемые вмешательства и подходы; основные рекомендации с соответствующими схемами градации их достоверности; вопросы, по которым мнения составителей совпадали или расходились".

Такое описание почти зеркально отражает ваши представления о научно обоснованных практических рекомендациях. К сожалению, среди подготовленных таким образом синтезов не числятся обобщенные клинические рекомендации по лечению ФГТ. Вы возвращаетесь на главную страницу, вводите термин "atrial fibrillation" ("фибрилляция предсердий") и находите 22 ссылки на практические рекомендации. Первый же заголовок выглядит многообещающе: "Fifth ACCP Consensus Conference on Antithrombotic Therapy" ("Материалы пятой согласительной конференции по применению антитромботической терапии", проведенной в 1998 г. Американской коллегией пульмонологов). На сайте National Guidelines Clearinghouse помещен реферат статьи, которая была опубликована в рецензируемом журнале [2]. Вы нажимаете кнопку "Complete Summary" ("Полный текст реферата") и распечатываете появившийся текст, а также запрашиваете по электронной почте копию опубликованной статьи из больничной библиотеки. Вы ожидаете эту статью с некоторой настороженностью, поскольку для многих рекомендаций (даже представляемых профессиональными обществами) зачастую характерен невысокий методологический уровень [3, 4].

Прежде чем покинуть сайт Google.com, вы вводите фразу "atrial fibrillation decision analysis" ("фибрилляция предсердий, анализ принятия решений"), и обнаруживаете следующую ссылку: www.thelancet.com/newlancet/sub/issues/vol355no9208/body.article956.html. Эта статья, содержит результаты недавно проведенного анализа принятия решений, она была опубликована в журнале Lancet и, по всей видимости, соответствует теме поиска [5].

КЛИНИЧЕСКИЕ РЕКОМЕНДАЦИИ, ПОСВЯЩЕННЫЕ ТАКТИКЕ ЛЕЧЕНИЯ

Ежедневно врачи принимают десятки решений о методах и схемах ^{лс}чения. Некоторые из таких решений не имеют большого значения, в ^{то} время как другие могут изменить течение заболевания. Любое из них подразумевает сравнение преимуществ и недостатков, положительных

и отрицательных эффектов лечения, определенный совет больному или назначение какого-то вмешательства в его интересах. Принимая клиническое решение, врач оценивает имеющуюся информацию, обобщает ее на интуитивном уровне, а также взвешивает вероятные преимущества и недостатки лечения, исходя из предпочтений больного. В процессе принятия решения можно с успехом использовать структурированные рефераты отчетов об исследованиях вмешательств и клинических исходов; систематические обзоры, посвященные взаимосвязям между вмешательствами и клиническими исходами; и рекомендации, в которых указаны наилучшие вмешательства в той или иной ситуации. В данном разделе книги описан процесс разработки клинических рекомендаций, предлагаются систематические подходы к этому процессу и представлена классификация, в которой различаются самые надежные (т.е. достоверные) и менее надежные (т.е. способные ввести в заблуждение) клинические рекомендации.

Традиционно в дискуссионный раздел отчетов об оригинальных (первичных) исследованиях лечебных вмешательств включают рекомендации по их применению в клинической практике. В систематических обзорах и мета-анализах также обычно обсуждаются возможные последствия практического использования их результатов. При этом авторы исследований названных типов, как правило, не рассматривают всех возможных терапевтических альтернатив, а сравнивают вместо этого лишь два или три метода. Кроме того, могут остаться не выявленными подгруппы больных, в которых эффект вмешательства существенно различается. И наконец, рекомендации в систематических обзорах не учитывают предпочтений больных или интересов общества.

Поэтому в разных оригинальных статьях и систематических обзорах по сходной тематике можно встретить противоречивые рекомендации, даже если они основаны на одной и той же информации. Например, при одинаковых результатах нескольких мета-анализов данных о селективной деконтаминации кишечника при помощи антибиотиков, осуществляемой для профилактики пневмонии у больных, которые находятся в критическом состоянии, их авторы предложили разные рекомендации. В одном случае вмешательство было признано полезным, в другом равноценным отсутствию профилактики, а в третьем нецелесообразным [6—9]. Подобные разногласия связаны с тем, что авторы первичных исследований и мета-анализов редко используют при разработке своих клинических рекомендаций общие стандарты или правила.

В тех случаях, когда соотношение преимуществ и недостатков вмешательства очевидно и мало изменяется в рамках исследуемой популя-

дии, при разработке клинических рекомендаций вполне можно руководствоваться интуицией. Однако такие ситуации редки. В большинстве случаев составители клинических рекомендаций, полагающиеся на интуицию, не могут избежать влияния систематических или случайных ошибок и вводят в заблуждение врачей и больных.

Поэтому прежде чем следовать рекомендации, необходимо тщательно оценить ее методологическое качество. Цель данного раздела заключается в том, чтобы снабдить врачей инструментами, необходимыми для такой оценки.

Клинические рекомендации, способные повлиять на распределение ресурсов здравоохранения, могут быть полезными для администраторов, которые определяют его политику, однако эта книга рассчитана в первую очередь на практикующих врачей. Сначала будет описан процесс разработки клинических рекомендаций и представлены два способа, которые используют при этом авторы клинических исследований, эксперты и органы управления здравоохранением — анализ принятия решений и составление практических рекомендаций. Затем мы предложим критерии, позволяющие оценить методологическое качество клинических рекомендаций, и представим их классификацию, которая может оказаться полезной для врачей.

ПРОЦЕСС РАЗРАБОТКИ КЛИНИЧЕСКИХ РЕКОМЕНДАЦИЙ

На рис. 1 представлены стадии и методы разработки клинических рекомендаций. Перед принятием клинического решения следует определить саму проблему. Для этого нужно составить перечень возможных вмешательств и клинических исходов. Часто лечение направлено на отсрочку или предупреждение неблагоприятных клинических исходов, таких как инсульт, смерть или инфаркт миокарда. Обычно их называют *оцениваемыми клиническими исходами*. Лечение также может приводить к развитию неблагоприятных клинических исходов, а именно побочных эффектов, токсических реакций или значительного изменения привычного Распорядка жизни больного (неудобства). Кроме того, применение новых вмешательств часто может существенно повысить или понизить стоимость лечения. В идеале при определении проблемы учитывают все возможные вмешательства, благоприятные и неблагоприятные клинические исходы. В нашем примере и других случаях ФП неклапанного Г_сНеза можно не проводить профилактику либо назначить аспирин или ан_тикоагулянты (варфарин). На фоне проводимой терапии или в ее от_{ст}вие может развиваться эмболический инсульт в легкой или тяжелой

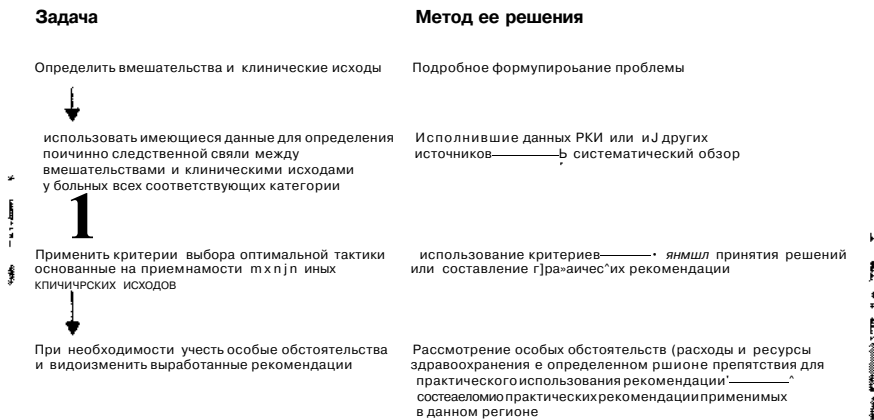


Рис. 1. Схема **процесса разработки клинических рекомендаций, посвященных тактике лечения.**

форме, внутримозговое кровоизлияние, желудочно-кишечное кровотечение, незначительное кровотечение иной локализации; с лечением и мониторингом его результатов могут быть связаны определенные неудобства и расходование средств больного, ресурсов здравоохранения или бюджета государства.

Определив вмешательства и клинические исходы, необходимо оценить причинно-следственную связь между ними. С какими преимуществами и недостатками может быть связано применение разных вмешательств [10, 11]? Насколько вероятно различие возможных преимуществ и недостатков лечения у больных разных категорий [11, 12]? Когда ответы на эти вопросы получены, следует оценить *приемлемость* разных клинических исходов на основании информации о *жизненных ценностях* и *предпочтениях* сторон, участвующих в лечении или профилактике. Мы считаем очень близкими по смыслу эти понятия, имеющие непосредственное отношение к сопоставлению благоприятных и неблагоприятных последствий применения альтернативных подходов.

В последнее время были разработаны и успешно применяются научные принципы выявления, отбора и обобщения данных, а также оценки приемлемости клинических исходов. Мы кратко рассмотрим систематический подход к выявлению, отбору и обобщению данных, более подробно описанный в главе 10 *Обобщение доказательств*, а затем предло-

жим два пути, которые ведут от доказательств к выработке клинических рекомендаций — анализ принятия решений и практические рекомендации.

СИСТЕМАТИЧЕСКИЕ ОБЗОРЫ

Отказ от использования систематического подхода к поиску и отбору данных может привести к возникновению систематической ошибки. Так, эффективность лечения бывает занижена или чаще завышена, а побочные эффекты — преувеличены или преуменьшены. Но даже применение систематического подхода при поиске и отборе информации не предотвращает смещения результатов обзора, если аналогичные методы не использовались при ее обобщении. Авторы подобных обзоров могут рекомендовать потенциально опасные вмешательства либо игнорировать эффективные. Например, эксперты советовали вводить лидокаин всем больным с острым инфарктом миокарда, хотя доступные данные свидетельствовали о бесполезности и даже вреде этого вмешательства, и в то же время не рекомендовали применять тромболитики, несмотря на наличие доказательств их эффективности [13].

Составители систематических обзоров могут избежать ошибок, четко определив критерии включения и исключения, осуществив исчерпывающий поиск данных, и обобщив их в соответствии со строгими правилами, в том числе с оценкой различий по размеру эффекта у разных разных категорий (см. главу 10 *Обобщение доказательств*). Систематический обзор, в котором обобщают данные различных исследований, чтобы количественно оценить эффект лечения, называется *мета-анализом*. Результаты систематических обзоров надежны, если методологическое качество первичных исследований высокое, а выборки большие. И наоборот, выводы авторов систематических обзоров малодостоверны, если методологическое качество включенных исследований низкое, а число их участников невелико. Поскольку субъективная оценка применяется на разных этапах проведения систематических обзоров (включая определение критериев включения и исключения, их применение по отношению к выявленным исследованиям, оценку методологического качества первичных исследований и выбор методов статистической обработки данных), они также не застрахованы от возникновения систематической ошибки. Тем не менее риск связанной с ней ошибочной оценки причинно-следственной связи между применением тех или иных вмешательств и клиническими исходами в систематических обзорах ниже в связи со строгими правилами выявления и обобщения данных

АНАЛИЗ ПРИНЯТИЯ РЕШЕНИЙ

В результате правильно проведенного *анализа принятия решений* обобщают данные о благоприятных и неблагоприятных эффектах лечения с учетом их приемлемости. При этом используются ясные количественные методы, позволяющие оценить те или иные решения в ситуации, когда соотношение их преимуществ и недостатков неочевидно. Анализ принятия решений позволяет врачам сравнить ожидаемые последствия применения различных вмешательств. В соответствующих статьях освещаются все стадии принятия решений, что делает возможным их обсуждение и видоизменение [14—16].

Под *анализом принятия клинических решений* мы подразумеваем исследование, в котором используются формальные математические подходы для оценки решений, принимаемых в клинической практике, например о проведении скрининговых обследований для выявления того или иного заболевания или о выборе того или иного метода лечения. Хотя анализ принятия решений можно провести для получения ответа на вопрос о ведении конкретного больного ("Нужно ли рекомендовать прием варфарина этой 76-летней женщине с ФП?"), чаще к нему прибегают при разработке общей тактики [17] ("Нужно ли рекомендовать прием варфарина всем больным с ФП?"). В качестве примеров частного и общего подхода можно соответственно привести исследование, отчет о котором мы собираемся рассмотреть, и рекомендации пожилой женщине с аортальным стенозом отказаться от хирургической коррекции этого порока сердца или согласиться на предложенное вмешательство [18].

Анализ принятия решений также может применяться для решения глобальных вопросов оказания медицинской помощи с учетом интересов общества в целом или последствий для национальной системы здравоохранения. Можно привести примеры анализа принятия решений о проведении скрининговых обследований для выявления рака предстательной железы и скринингового определения уровня липидов в крови с последующим проведением гипохолестеринемической терапии [20]. Методы анализа принятия решений, используемые организаторами здравоохранения и практикующими врачами, имеют много общего [21], однако есть и различия, обсуждение которых не входит в задачи этой книги.

В большинстве случаев при клиническом анализе принятия решений строят так называемый алгоритм принятия решений, который отображают в статье на одной или нескольких диаграммах. Мы рассмотрим их наглядно, чтобы помочь читателю понять принцип анализа принятия решений. На рис. 2 схематически представлен алгоритм принятия

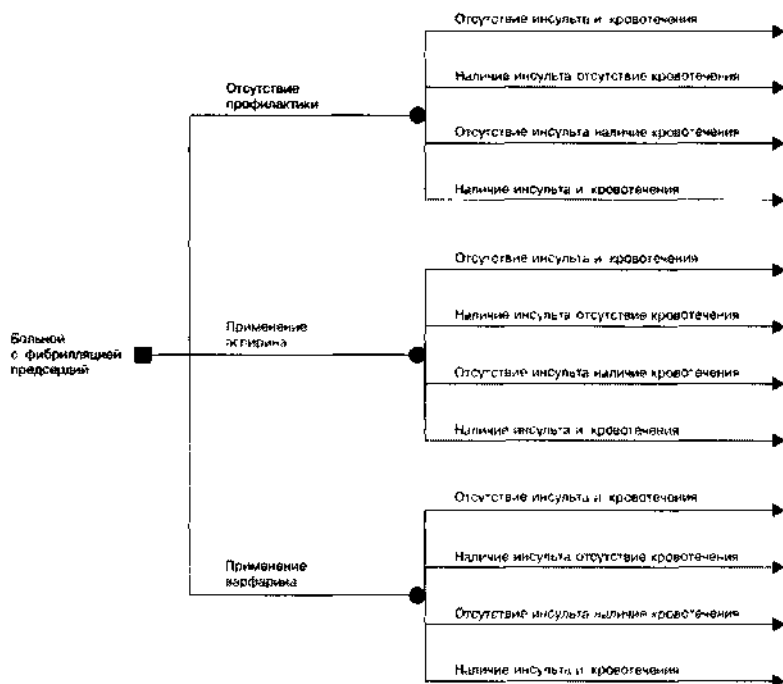


Рис. 2. Упрощенный алгоритм принятия клинических решений при фибрилляции предсердий.

решений, относящихся к вышеупомянутому примеру с ФП. У больных с ФП, которым не показана антиаритмическая терапия для достижения и поддержания синусового ритма, можно выбрать одну из трех схем лечения: отсутствие профилактики, применение аспирина или применение варфарина. Независимо от того, какой выбор был сделан, у некоторых больных могут развиваться эмболические осложнения (в частности, инсульт). Профилактические вмешательства снижают этот риск, но повышают вероятность возникновения кровотечений. Такая упрощенная Модель не учитывает некоторые важные последствия лечения (включая неудобства, связанные с регулярным обследованием при терапии варфарином, а также с возникновением незначительных кровотечений).

Как видно из рис. 2, алгоритм принятия решений графически разделен на 3 части, т.е. слева представлена клиническая проблема, в

центре — сравниваемые методы ее решения, справа — клинические исходы. Исходный пункт принятия клинических решений отмечен маленьким черным квадратом (его еще называют *узлом решений*). Выходящие из него линии соответствуют вмешательствам, которые можно применить в данной ситуации. Совокупность вероятных клинических исходов отмечена маленькими черными кругами (или *узлами вероятностей*), а отдельные последствия принятого решения — треугольниками (как на рис. 2) или прямоугольниками (такое обозначение также можно встретить в литературе). Когда помимо клинических исходов при анализе принятия решений рассматривают стоимость различных методов лечения или профилактики, его называют *анализом эффективности затрат*, отображающим соотношение между клинической эффективностью вмешательств и связанным с ними расходом ресурсов здравоохранения [22, 23] (см. главу 2F на компакт-диске).

После разработки алгоритма принятия решений следует сделать количественные предположения о *вероятности* наступления того или иного клинического исхода, которая может колебаться от 0 (событие невозможно) до 1,0 (событие обязательно произойдет). Вероятность должна быть назначена для каждой линии, выходящей из узла вероятностей, а сумма всех вероятностей для каждого узла должна составить 1,0. Возвращаясь к рис. 2, рассмотрим одну из выбранных схем — отсутствие профилактики. Это решение соответствует верхней линии, исходящей из узла решений. Она заканчивается узлом вероятностей четырех возможных комбинаций двух клинических исходов (инсульта и кровотечения). На рис. 3 представлены значения вероятности для верхней линии алгоритма принятия решений (отсутствие профилактики), подсчитанные после предположения о том, что вероятность развития кровотечения составляет 1%, а вероятность инсульта — 10%, и что эти два события независимы друг от друга. Тогда вероятность развития и инсульта, и кровотечения составит 0,1% (0,001), вероятность развития кровотечения без инсульта — 0,9% (0,009), вероятность развития инсульта без кровотечения — 9,9% (0,099), а вероятность отсутствия как инсульта, так и кровотечения — 89,1% (0,891).

При анализе принятия решений необходимо оценить вероятности и для двух других возможных схем лечения. Можно предположить, что для средней линии (прием аспирина) вероятность развития кровотечения будет выше, а вероятность развития инсульта — ниже. Для нижней линии (прием варфарина) вероятность кровотечения окажется максимальной, а вероятность развития инсульта — минимальной.

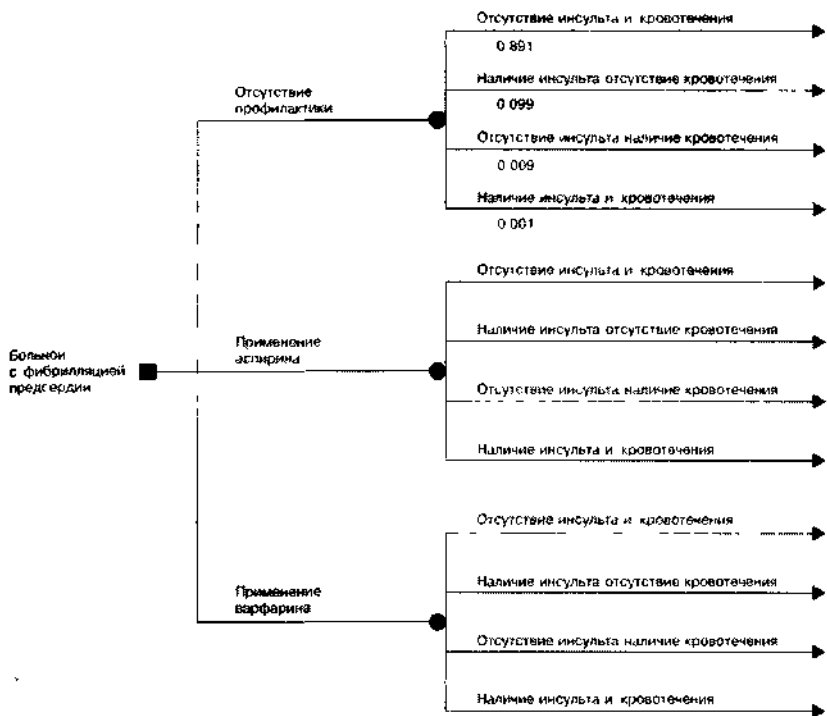


Рис. 3. Алгоритм принятия решений с расчетом вероятности возникновения отдельных клинических исходов в отсутствие профилактики.

Знание вероятности возникновения этих осложнений не облегчает выбор определенной тактики, поскольку более выраженное снижение риска кровотечения сопровождается более выраженным повышением риска развития инсульта и наоборот. Правильность решения будет определяться относительной ценностью предупреждения инсульта или снижения риска кровотечения для конкретного больного. При анализе принятия решений обычно указывают приемлемость каждого клинического исхода, причем этот показатель изменяется от 0 (смерть) до 1,0 (полное здоровье). На рис. 4 отмечены гипотетические показатели приемлемости четырех клинических исходов, рассчитанные для верхней ветви дерева решений: 1,0 для отсутствия инсульта и кровотечения; 0,8 для отсутствия инсульта и наличия кровотечения; 0,5 для наличия ин-

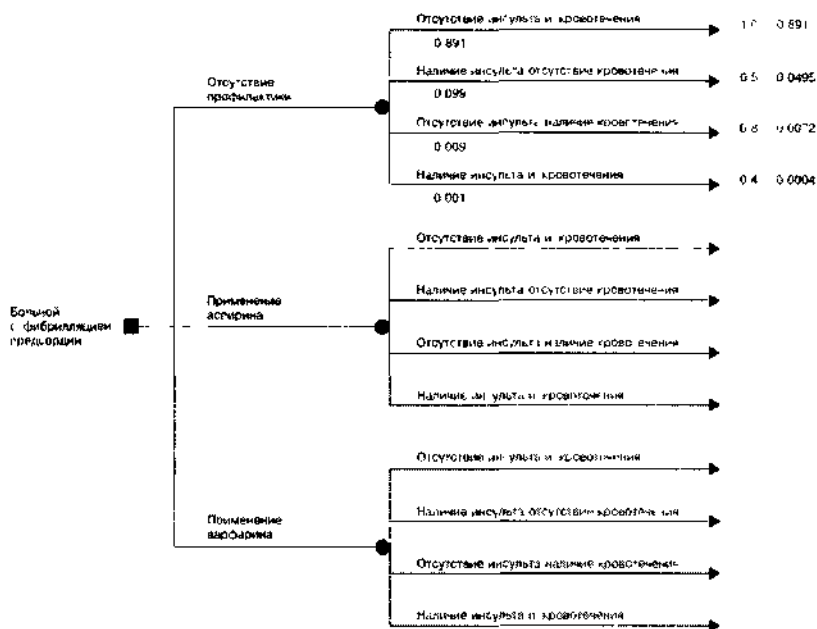


Рис. 4. Алгоритм принятия решений с расчетом вероятности и приемлемости отдельных клинических исходов в отсутствие профилактики.

сульта и отсутствия кровотечения и 0,4 для наличия и инсульта, и кровотечения.

На последнем этапе анализа принятия решений рассчитывают интегральный показатель приемлемости и вероятности для каждого вмешательства. Для представленных значений вероятности и приемлемости при отказе от профилактики расчет этого показателя будет выглядеть следующим образом: $(0,891 \times 1,0) + (0,009 \times 0,8) + (0,099 \times 0,5) + (0,001 \times 0,4) = 0,948$. В зависимости от вероятностей, выбранных для средней и нижней линий, можно будет оценить приемлемость применения аспирина или варфарина по сравнению с отсутствием профилактики. Если интегральный показатель приемлемости и вероятности для любой из этих линий будет превышать 0,948, соответствующее вмешательство окажется более приемлемым по сравнению с отсутствием профилактики, если нет — менее приемлемым.

Алгоритм, представленный на рис. 2—4, в значительной степени усложнен. Так, в нем не учитываются сроки наступления клинического ис-

хода и возможность последовательного развития у больного нескольких клинических исходов. При анализе принятия решений часто используют компьютерные программы, с помощью которых можно смоделировать клинические исходы в гипотетической когорте больных при нескольких циклах наблюдения определенной продолжительности (например, равной году). Модель может также учитывать возможность изменения состояния здоровья с течением времени. Например, у кого-то из больных в течение первого цикла наблюдения может развиваться инсульт в легкой форме, затем на протяжении нескольких циклов у него будут отмечаться слабо выраженные нарушения функционального статуса, после чего возникнет желудочно-кишечное кровотечение и, наконец, тяжелый инсульт. Такие модели переходных состояний, или модели Маркова, обычно отражают ситуацию во всей ее сложности, близко к действительности и поэтому повышают точность анализа принятия решений.

ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ

Практические рекомендации ("систематически разрабатываемые положения, направленные на облегчение принятия врачом и больным решения о тактике лечения в конкретной ситуации" [24]) представляют собой еще один способ обобщения имеющихся данных и оценки предпочтений с последующей выработкой клинических рекомендаций о выборе тактики лечения [1, 25—30]. При составлении практических рекомендаций количественному анализу придают меньшее значение, чем при анализе принятия решений. Учитываются совпадающие мнения группы экспертов, в идеале включающей специалистов, практикующих врачей и больных, которые тщательно оценивают данные о вмешательствах и принимают решение об их применении. Составители практических рекомендаций наделяются определенными временными полномочиями и могут принимать участие в подобных проектах на уровне страны, области, города, больницы или амбулаторной практики. Такие рекомендации, основанные на одинаковых данных, могут различаться в зависимости от того, составляются ли они в Филиппинах или США, "Редназначены ли для применения в городской или сельской местности" > в крупной университетской клинике или небольшой больнице, в поликлинике, расположенной в уважаемых или бедных кварталах. Та^К в Р^а ч^и практикующие в сельских районах слаборазвитых стран, могут сразу отказаться от применения варфарина при ФП, поскольку не имеют возможности проводить регулярное тщательное исследование свертывающей системы крови

Таблица 1. Рекомендуемый подход к оценке достоверности клинических рекомендации, посвященных тактике лечения

Учтены ли в клинических рекомендациях все категории больных с определенным заболеванием, вмешательства и клинические исходы?

Существует ли систематический обзор, в котором по каждому рассматриваемому клиническому вопросу обобщается информация о причинно-следственной связи между вмешательствами и клиническими исходами⁹

Был ли определен перечень предпочтений и жизненных ценностей, учитываемых при рассмотрении возможных клинических исходов⁹

Сообщается ли о степени достоверности клинических рекомендаций?

И анализ принятия решений, и практические рекомендации могут обладать высоким или низким методологическим качеством, что сказывается на достоверности конечных клинических рекомендаций. В табл. 1 мы предлагаем четыре принципа ее оценки.

ОЦЕНКА КЛИНИЧЕСКИХ РЕКОМЕНДАЦИЙ

УЧТЕНЬ! ЛИ В КЛИНИЧЕСКИХ РЕКОМЕНДАЦИЯХ ВСЕ КАТЕГОРИИ БОЛЬНЫХ С ОПРЕДЕЛЕННЫМ ЗАБОЛЕВАНИЕМ, ВСЕ ВМЕШАТЕЛЬСТВА И КЛИНИЧЕСКИЕ ИСХОДЫ?

Клинические рекомендации предназначены для использования в процессе принятия решений о применении определенных методов лечения или профилактики у больных определенных категорий с учетом возможных последствий. Вне зависимости от того, посвящены ли клинические рекомендации вопросам диагностики, профилактики, лечения или реабилитации, в них должны быть учтены все возможные категории больных, методы традиционной и нетрадиционной (альтернативной) медицины. Например, в анализ принятия решений о том или ином лечении при предполагаемом герпетическом энцефалите авторы включили три вмешательства, доступные для врачей на момент составления рекомендаций: биопсию мозга, применение видарабина и отсутствие лечения [31]. Эта модель отражала положение дел в то время, когда были опубликованы рекомендации, но впоследствии появился ацикловир, который теперь широко используется при данном заболевании. Посколь-

ку исходная модель не включала оценку лечения ацикловиром, она уже не может считаться надежной основой для принятия решения.

В качестве другого примера можно привести практические рекомендации Американской коллегии врачей, разработанные по результатам систематического обзора [32], в которых были предложены подходы к медикаментозной профилактике инсульта [33]. Хотя авторы упоминали каротидную эндартерэктомию как альтернативный метод профилактики, эта операция не была включена в рекомендации. Последние были бы гораздо достоверней, если бы применение лекарственных препаратов для предотвращения преходящих нарушений мозгового кровообращения рассматривалось совместно с высокоэффективным хирургическим вмешательством [34].

Клинические рекомендации часто различаются по тому, на какие категории больных они рассчитаны. В частности, у больных с наименьшим риском развития изучаемого клинического исхода, на предотвращение которого направлено вмешательство, вероятность эффективности этого вмешательства будет ниже, чем у больных с максимальным исходным риском (см. главу 2В3 на компакт-диске). Например, Американская коллегия врачей разработала отдельные рекомендации по использованию заместительной гормональной терапии в постменопаузе у женщин разных категорий: в общей популяции; после экстирпации матки; при высоком риске развития сердечно-сосудистых осложнений; при высоком риске развития рака молочной железы [35].

При подготовке клинических рекомендаций должны учитываться не только все соответствующие категории больных и вмешательства, но также все важные последствия применения последних. Данные о влиянии лечения или профилактики на частоту развития осложнений, смертность или качество жизни имеют большое значение для больных, а необходимость эффективного использования ресурсов определяет повышенный интерес к затратам на медицинскую помощь. Рассмотрение последних при разработке клинических рекомендаций может повлиять на их содержание, т.е. на выводы их составителей, независимо от того, учитывались ли в первую очередь интересы больных, страхователей, всей системы здравоохранения или конечной целью было решение общих вопросов, таких как влияние разных подходов к ведению больных на продолжительность нетрудоспособности (см. главу 2F на компакт-диске). Поэтому анализ принятия решений, который включает оценку экономической эффективности вмешательств, называется *экономическим анализом*.

Составление клинических рекомендаций по проведению скрининговых обследований требует особого внимания к определению всех воз-

можных клинических исходов. Попытки выявить болезнь в отсутствие симптомов могут привести к ряду нежелательных последствий, с которыми не сталкиваются врачи, обследующие и лечащие больных с теми или иными клиническими проявлениями. Люди, у которых при скрининговых обследованиях было диагностировано неизлечимое заболевание, в течение более длительного времени ощущают свое нездоровье, что связано с дополнительными переживаниями. Это особенно проблематично, если заболевание, на выявление которого направлено скрининговое обследование, еще долго останется бессимптомным. Представьте себе мужчину, которому в результате обследования был поставлен диагноз рака предстательной железы, но умереть суждено от заболевания сердца до того, как появятся первые симптомы рака. Кроме того, лицам, у которых при более тщательном обследовании наличие заболевания не подтвердилось, скрининговое обследование может нанести психологическую травму, а тот, кто при скрининговом обследовании был признан здоровым, но впоследствии заболел, будет чувствовать себя обманутым (см. главу 2F на компакт-диске).

В рекомендациях по использованию заместительной гормональной терапии, разработанных Американской коллегией врачей, были использованы данные о риске развития на протяжении жизни рака эндометрия, рака молочной железы, перелома шейки бедренной кости, ишемической болезни сердца, инсульта и о медиане продолжительности жизни; это позволило оценить преимущества и недостатки такого лечебного вмешательства в разных подгруппах. Были подтверждены отрицательное влияние заместительной гормональной терапии на концентрации липопротеинов в крови, половую и мочевыделительную функцию, повышение риска развития маточного кровотечения и необходимость инвазивных исследований, однако эту информацию не включили в математическую модель, использованную при обработке данных. Влияние заместительной гормональной терапии на качество жизни, которое могло в наибольшей степени определить выбор самих больных, подробно не рассматривалось.

В ходе анализа принятия решений о применении антикоагулянтов при дилатационной кардиомиопатии [4] авторы использовали модель, в которой учитывались все возможные клинические исходы, имеющие значение для больных (инсульт, тромбоэмболия других сосудов, кровотечение и т.п.). Клинические исходы оценивались по шкале ожидаемой продолжительности жизни, скорректированной с учетом ее качества, что позволяет интегрировать данные о продолжительности и качестве жизни. Такой количественный подход кажется целесообразным в **ситуации**

ях, подобных описанной в примере с ФП, поскольку можно предположить, что прием варфарина повлияет как на продолжительность жизни, так и на ее качество.

СУЩЕСТВУЕТ ЛИ СИСТЕМАТИЧЕСКИЙ ОБЗОР, В КОТОРОМ ПО КАЖДОМУ РАССМАТРИВАЕМОМУ КЛИНИЧЕСКОМУ ВОПРОСУ ОБОБЩАЕТСЯ ИНФОРМАЦИЯ О ПРИЧИННО-СЛЕДСТВЕННОЙ СВЯЗИ МЕЖДУ ВМЕШАТЕЛЬСТВАМИ И КЛИНИЧЕСКИМИ ИСХОДАМИ?

После того как были определены возможные вмешательства и клинические исходы, следует оценить вероятность наступления последних. В результате может возникнуть ряд специфических вопросов. В случае с заместительной гормональной терапией первый из них будет звучать приблизительно так: "Каково влияние альтернативных вмешательств на частоту развития переломов шейки бедренной кости, рака молочной железы, рака эндометрия, инфаркта миокарда или внезапной смерти?". В клинических рекомендациях должна быть собрана воедино и обобщена с использованием соответствующих методов, сводящих к минимуму риск возникновения систематической ошибки, вся актуальная информация. При этом необходимо найти или провести систематический обзор с рассмотрением данных по каждому из возникших вопросов. В главе 10 *Обобщение доказательств* предложены рекомендации, которые позволят определить вероятность того, что сбор и обобщение данных свободны от ошибок.

В идеале в клинических рекомендациях должны быть описаны критерии включения доказательств, методы их отбора и обобщения, наглядно представлена ключевая информация, а также нужно сообщать результаты РКИ, в которых соотносились вмешательства и клинические исходы. Однако такие РКИ могут оказаться недоступными. В подобной ситуации авторы обзора отказались бы от его проведения, сославшись на отсутствие высококачественных исследований с обобщаемыми результатами. Так не могут поступить составители клинических рекомендаций. Надежная доказательная информация по важным вопросам, изучение которых затруднено по этическим, техническим или экономическим соображениям, может отсутствовать. Поскольку для составления **клинических** рекомендаций необходимо использовать самые достоверные **данные** (которые часто не выдерживают никакой критики) из имеющихся, необходимо учитывать результаты исследований с различной структурой (опубликованных и неопубликованных), мнение экспертов и потребителей медицинских услуг. Это означает, что обоснованность ин-

формации, на которой строятся клинические рекомендации, может быть различной. Таким образом, даже клинические рекомендации, появившиеся в результате тщательного отбора и вдумчивого обобщения данных, могут оказаться ненадежными, если сами данные недостоверны. К этому вопросу мы вернемся позднее в этой главе (см. табл. 1).

БЫЛ ЛИ ОПРЕДЕЛЕН ПЕРЕЧЕНЬ ПРЕДПОЧТЕНИЙ И ЖИЗНЕННЫХ ЦЕННОСТЕЙ, УЧИТЫВАЕМЫХ ПРИ РАССМОТРЕНИИ ВОЗМОЖНЫХ КЛИНИЧЕСКИХ ИСХОДОВ?

Выявление связи между вмешательствами и клиническими исходами по большей части зависит от действительных причинно-следственных отношений и методов исследования. Приемлемость тех или иных клинических исходов, напротив, является функцией предпочтений и жизненных ценностей сторон, участвующих в лечении или профилактике. Можно по-разному оценить приемлемость повышения риска развития рака молочной железы при заместительной гормональной терапии с целью снижения риска переломов. Поэтому важно, чтобы авторы рекомендаций сообщали об основных источниках данных о приемлемости клинических исходов и методах достижения согласия.

Врач должен искать сведения о том, кто прямо или косвенно определял приемлемость клинических исходов. В разработке клинических рекомендаций нередко участвуют группы экспертов и согласительные комиссии. Необходимо знать, что понимается под термином "эксперт", поскольку при формировании таких групп и комиссий преимущественно из врачей профильной специальности высока вероятность возникновения систематических ошибок, связанных с различиями в уровне интеллекта, территориальными и финансовыми различиями. Группы и комиссии, в которые в продуманном соотношении включают экспертов по методологии медицинских исследований, врачей общей практики, специалистов и представителей общественности, скорее всего будут обсуждать разные точки зрения, чего не произойдет, если коллектив авторов будет представлен специалистами в одной области медицины.

Даже если группа составителей сформирована правильно по составу, организация обсуждения разногласий может повлиять на качество клинических рекомендаций. Поэтому необходимо искать описание специальных методов, используемых для достижения согласия. Отказ от применения таких методов может привести к доминированию отдельных членов группы или комиссии, особенно председателя. Продуманные подходы к достижению согласия при расхождении мнений или прямое заявление о числе проголосовавших "за", "против" и воздержавшихся повышают надежность клинических рекомендаций.

Знание того, в какой степени учтены предпочтения больных, особенно важно. Во многих рекомендациях о предпочтениях больных не упоминается, но предполагается, что разработчики учли основные интересы больных. Хотя об этих интересах сообщается редко, читатель должен знать, какие принципы (независимость больного от стационарной или амбулаторной медицинской помощи; принцип "не навреди"; справедливое распределение ресурсов здравоохранения) были приоритетными при выборе тактики лечения или профилактики. В рекомендациях наиболее высокого методологического качества сообщается, направлены ли они на наиболее полное удовлетворение потребностей больных, финансирующих организаций или общества в целом. В идеале должны быть перечислены основные ценностные принципы, на которых базируются данные рекомендации.

Например, в рекомендациях по медикаментозной профилактике инсульта, разработанных Американской коллегией врачей, при преходящих нарушениях мозгового кровообращения аспирин предлагается в качестве препарата первого ряда, а тиклопидин — в качестве препарата второго ряда (в случаях непереносимости аспирина) [24]. В оптимальном варианте при преходящем нарушении мозгового кровообращения применение тиклопидина по сравнению с аспирином снижает относительный риска развития инсульта на 15%, что соответствует предупреждению одного дополнительного инсульта у каждых 70 больных, принимающих тиклопидин (при исходном риске инсульта, составляющем 10%). Рекомендация принимать аспирин вместо тиклопидина при преходящих нарушениях мозгового кровообращения основана, по меньшей мере отчасти, на большей стоимости тиклопидина, а также на необходимости в периодическом определении числа и соотношения различных форм лейкоцитов у больных, принимающих этот препарат. Если бы авторы назвали жизненные ценности, которые они рассматривали в процессе принятия решения, такое суждение о предпочтительности могло бы быть оспорено, а рекомендация стала бы от этого более достоверной.

В отличие от практических рекомендаций для анализа принятия решений не характерны серьезные проблемы, связанные с недостаточным освещением или сокрытием данных о предпочтениях и жизненных ценностях, которые определяли выбор того или иного вмешательства. Из рис. 4 видно, что причина этого заключается в обязательности наглядного количественного представления предпочтений при анализе принятия решений. Показатели приемлемости отражают предпочтения лица, рассматривающего различные клинические исходы. Существует

несколько методов их непосредственной оценки [5, 7, 24, 25] (см. главу 2В2 на компакт-диске), но пока неясно, какой из них наилучший.

Какой бы метод оценки предпочтений ни использовался, авторы рекомендаций должны указать источник соответствующих данных. При анализе принятия решений для выбора тактики лечения конкретного больного самый, и, возможно, единственно правильный способ оценки предпочтений — узнать мнение самого больного. Если анализ принятия решений проводится для разработки общих рекомендаций, информацию о приемлемости клинических исходов можно получить из трех источников. Во-первых, можно опросить большую группу больных с изучаемым заболеванием, к которым могут быть применимы результаты анализа принятия решений. Во-вторых, можно изучить отчеты об опубликованных исследованиях качества жизни у подобных больных, как это было сделано в недавно проведенном анализе принятия решений при постоянной форме ФП [26]. В-третьих, можно исходить из оценок, сделанных такими же крупными группами представителей общей популяции. Лица, оценивающие приемлемость клинических исходов при проведении различных вмешательств, должны иметь представление о соответствующих клинических исходах и методах лечения или профилактики. Чем больше члены соответствующей группы знают об изучаемом заболевании или состоянии, тем правдоподобнее будет результат оценки приемлемости клинических исходов.

СООБЩАЕТСЯ ЛИ О СТЕПЕНИ ДОСТОВЕРНОСТИ ПРЕДЛАГАЕМЫХ РЕКОМЕНДАЦИЙ?

Достоверность (обоснованность) клинических рекомендаций определяется множеством факторов, в том числе методологическим качеством первичных исследований, включенных в систематический или описательный обзор; размером и постоянством лечебного или профилактического эффекта вмешательства в разных исследованиях; выраженностью побочных эффектов; неудобствами для больного, связанными с практическим применением и стоимостью предложенных методик; расходом физических и материальных ресурсов здравоохранения, а также с относительной приемлемостью возможных клинических исходов. На противоположных полюсах находятся клинические рекомендации, основанные на данных систематических обзоров крупных РКИ, з которых показано значительное положительное влияние вмешательств на важные для больных клинические исходы при минимальных побочных эффектах, затратах и неудобствах для больных (наиболее обоснованные клинические рекомендации), и те рекомендации, при подготов-

ке которых использовались результаты наблюдательных исследований, свидетельствующих о небольшом размере эффекта вмешательств при выраженных побочных реакциях и высокой стоимости (клинические рекомендации низкого методологического качества).

Есть два способа, при помощи которых разработчики клинических рекомендаций могут указать их достоверность (обоснованность). Во-первых (что наиболее приемлемо для практических рекомендаций), можно оценить этот показатель по специальной шкале. Во-вторых (что лучше всего подходит для анализа принятия решений), можно определить степень, в которой изменение предположений о влиянии вмешательства на клинические исходы изменяет характер рекомендаций (анализ чувствительности). Мы обсудим оба этих подхода.

СТЕПЕНЬ ДОСТОВЕРНОСТИ КЛИНИЧЕСКИХ РЕКОМЕНДАЦИИ

Специальная канадская группа разработчиков рекомендаций по периодическим профилактическим обследованиям (The Canadian Task Force on the Periodic Health Examination) предложила первую формальную классификацию клинических рекомендаций, основанную на научной обоснованности использовавшихся при их составлении первичных исследований [36—38]. Мы видоизменили эту классификацию, принимая во внимание, что практические рекомендации наиболее высокого методологического качества должны опираться на выводы систематических обзоров, в которых обобщаются самые доказательные данные доступных оригинальных исследований (табл. 2).

Степень достоверности клинических рекомендаций мы будем обозначать буквами А, В, С+ и С (см. табл. 2). Они отражают методологическое качество обзоров, авторы которых могли обобщать информацию, полученную в РКИ с однородными результатами (степень А), или же Данные наблюдательных исследований с неоднородными результатами. Поскольку выводы об эффекте вмешательства становятся менее обоснованными при наличии необъяснимых и значительных различий по эффективности лечения или профилактики между различными исследованиями, неоднородность результатов первичных исследований снижает надежность рекомендаций даже в тех случаях, когда первичными исследованиями являются РКИ (степень В). Обоснованность клинических рекомендаций, опирающихся на информацию, которая была получена в наблюдательных исследованиях, еще ниже (степень С).

Встречаются ситуации, когда РКИ, тематика которых строго соответствовала бы клиническому вопросу, недоступны. В таких ситуациях авторы практических рекомендаций могут успешно применить данные

об эффективности вмешательства в одной подгруппе больных к другой подгруппе. Например, в РКИ показано значительное снижение частоты развития инсультов с ФП без поражения митрального клапана. Имеющаяся информация о патогенезе ФП позволяет предположить, что эти результаты применимы к больным с ФП клапанной этиологии. Во-вторых, в обсервационных исследованиях может быть выявлен большой и постоянный размер эффекта лечения или профилактики, например при изучении терапии инсулином при диабетическом кетоацидозе. И в том, и в другом случае достоверность составленных таким образом клинических рекомендаций можно отнести к категории С+.

Если данные о причинно-следственной связи между вмешательствами и клиническими исходами были получены в систематических обзорах первичных исследований, врачи могут применить для оценки клинических рекомендаций критерии достоверности результатов систематического обзора и схему, представленную в табл. 2.

В табл. 2 число категорий первой и второй групп отражает баланс между преимуществами и недостатками лечения. Если преимущества изучаемого вмешательства явно перевешивают его недостатки (или недостатки альтернативного вмешательства перевешивают его преимущества), а выбор почти всех больных одинаков, клиническим рекомендациям можно присвоить категорию 1. Если соотношение преимуществ и недостатков не столь очевидно, а потому разные больные могут принять разные решения, клинические рекомендации соответствуют категории 2.

Сомнения в балансе между преимуществами и недостатками лечения или профилактики могут быть обусловлены рядом причин, включая существенные различия в предпочтениях самих больных и широкие доверительные интервалы при оценке размера благоприятного или неблагоприятного эффекта (см. главу 2F на компакт-диске).

Если клинические рекомендации разработаны на основе обсервационных исследований, или оценка эффективности лечения приближительна, можно сделать вывод о том, что достоверность клинических рекомендаций сравнительно невелика. Проверить его правильность можно, проанализировав результаты практического применения рекомендаций на клинические исходы. Например, S.R. Weingarten et al. [39] оценили последствия практического применения рекомендаций, согласно которым всех больных с низким риском развития сердечно-сосудистых осложнений, помещенных в отделения интенсивной кардиологии, следует раньше выписывать из стационара. В течение 1 года в течение четных месяцев врачам раздавали памятки с соответствующими выдержками из упомянутых рекомендаций. В те месяцы, когда врачи следовали его по-

Таблица 2. Подход к оценке степени достоверности клинических рекомендаций, разработанных на основе систематических обзоров

Степень достоверности клинических рекомендаций	Соотношение преимуществ и недостатков	Структура оригинальных исследований и их методологическое качество	Комментарий
IA	Очевидное	РКИ без существенных недостатков	Достоверные рекомендации; безоговорочно применимы у большинства больных в большинстве клинических ситуаций
IB	Очевидное	РКИ с существенными недостатками (неоднородность результатов, методологические ограничения*)	Достоверные рекомендации; вероятнее всего, применимы у большинства больных
IC+	Очевидное	РКИ, тематика которых строго соответствует клиническому вопросу, отсутствуют, но можно быть уверенным в применимости результатов РКИ по сходной тематике либо обсервационных исследований	Достоверные рекомендации; применимы у большинства больных в большинстве клинических ситуаций
IC	Очевидное	Обсервационные исследования	Достоверность клинических рекомендаций промежуточная, она может измениться при получении более обоснованных данных
2A	Неочевидное	РКИ без существенных недостатков	Достоверность клинических рекомендаций промежуточная, выбор наилучшей тактики решения может изменяться в зависимости от клинической ситуации, предпочтений больного или ценностной ориентации общества

Продолжение табл. 2 на с.258

Таблица 2 (продолжение)

Степень достоверности клинических рекомендаций	Соотношение преимуществ и недостатков	Структура оригинальных исследований и их методологическое качество	Комментарий
2В	Неочевидное	РКИ с существенными недостатками (неоднородность результатов, методологические ограничения*)	Достоверность клинических рекомендаций низкая; у некоторых больных в некоторых ситуациях более эффективными могут оказаться другие подходы к лечению
2С	Неочевидное	Обсервационные исследования	Недостоверные рекомендации; другие подходы могут быть не менее эффективны

Примечание. * -- включая отказ от использования слепого метода при использовании субъективных критериев оценки клинических исходов, что может вызвать систематическую ошибку, связанную с измерением; и выбывание большого числа участников.

Поскольку клинические рекомендации, которым были присвоены категории В или С, основаны на недостоверных данных, чаще всего их можно отнести к группе категорий 2.

Клинические рекомендации относят к группе категорий 1 или 2 в зависимости от размера эффекта лечения (профилактики) и точности его оценки; риска развития клинического исхода, который стремится предотвратить; конкретных преимуществ, которыми пользуются больные основной группы по сравнению с контрольной, и выраженности побочных эффектов; разнообразия предпочтений больных; региональных различий по доступности ресурсов здравоохранения и подходам к оказанию медицинской помощи; стоимости лечения (профилактики). Оценка этих факторов неизбежно несет в себе элемент субъективности.

ложениям, продолжительность пребывания больных в отделении интенсивной кардиологии была меньше примерно на 1 день, а средняя стоимость лечения — более чем на 1000 долларов США. В обеих группах были одинаковыми смертность и общее состояние через месяц после выписки. Если подобное исследование проведено на высоком методологическом уровне, оно повышает достоверность данных, положенных в основу клинических рекомендаций, и надежность последних.

Рекомендации по заместительной гормональной терапии, которые были упомянуты ранее, демонстрируют ограничения любых рекомендаций, опирающихся на результаты первичных исследований низкого методологического качества [35]. Хотя авторы не указали степень надежности своих рекомендаций, известно, что они основаны главным образом на результатах обсервационных исследований, а потому могут быть отнесены к категории 2С в соответствии со схемой, представленной в табл. 2. В частности, эти рекомендации базируются в основном на мета-анализе результатов обсервационных исследований, посвященных оценке влияния заместительной гормональной терапии на течение ишемической болезни сердца; в этом мета-анализе сообщается, что снижение относительного риска развития сердечно-сосудистых осложнений составляет 0,35. Впоследствии в ходе первого же крупного РКИ, включавшего женщин с подтвержденной ишемической болезнью сердца, снижение частоты развития таких осложнений на фоне заместительной гормональной терапии не подтвердилось [40]. Ясно, что врач должен с большой осторожностью применять на практике клинические рекомендации, относящиеся к категории С.

АНАЛИЗ ЧУВСТВИТЕЛЬНОСТИ

При анализе принятия решений специалисты могут использовать систематические методы для оценки неопределенности данных, т.е. определять степень, в которой изменение предположений об исходном риске развития тех или иных клинических исходов, преимуществах вмешательства и предпочтениях сторон, участвующих в лечении (профилактике), изменяет характер рекомендаций. Такой *анализ чувствительности* предпринимают для того, чтобы ответить на вопрос: влияет ли неопределенность данных о вероятности развития клинических исходов или их приемлемости на принятие конкретного решения? Если при этом изменяют значения одного показателя, анализ чувствительности называется однофакторным, если двух или более — многофакторным. Например, исследователи, проводящие анализ принятия решений о назначении антибиотиков для профилакти-

ки заражения *Mycobacterium avium-intracellulare* у ВИЧ-инфицированных, обнаружили, что эффективность затрат снижается, если предполагается большая продолжительность жизни больных либо меньшая эффективность препарата [41]. Если предполагается одновременно и большая продолжительность жизни больных, и меньшая эффективность препарата (двухфакторный анализ чувствительности), гипотетическая эффективность затрат на подобную профилактику станет существенно ниже. Обращаясь к результатам анализа чувствительности, необходимо искать сводный перечень переменных, учитываемых при анализе, а также сведения о рассматриваемых значениях каждой переменной и список переменных, влияющих на выбор тактики лечения или профилактики (если таковые имеются).

В целом с помощью анализа чувствительности должны быть оценены все предположения о вероятности развития клинических исходов. Диапазон рассматриваемых значений зависит от источника данных. Если вероятность была рассчитана в крупных РКИ высокого методологического качества при небольших доверительных интервалах, может быть выбран узкий диапазон рассматриваемых значений. Он будет тем шире, чем ниже методологическое качество исследования и точность расчетов.

Необходимо выверять при помощи одноименного анализа чувствительность используемой модели к изменениям предположений о приемлемости клинических исходов, причем и в этом случае диапазон рассматриваемых значений зависит от источника данных. Если сходные предпочтения были определены в больших группах больных или представителей общей популяции, информированных о клинических исходах, диапазон рассматриваемых значений может быть узким. Если же оценка проводилась в небольшой группе больных или мнения экспертов значительно расходятся, следует выбрать более широкий диапазон рассматриваемых значений. Когда при изменении предположений о вероятности и приемлемости клинических исходов основные результаты анализа принятия решений остаются прежними, рекомендация надежна. В противном случае ее нельзя считать обоснованной.

В табл. 3 клинические рекомендации, посвященные лечению, классифицированы по их методологическому качеству. В этой классификации основное внимание уделяется трем ключевым компонентам: рассмотрению всех соответствующих вмешательств и клинических исходов, систематическому обобщению данных и объективной и/или количественной оценке предпочтений больных или интересов общества.

Таблица. 3. Классификация клинических рекомендаций в зависимости от их методологического качества

Методологическое качество	Систематическое обобщение данных	Рассмотрение всех вмешательств и клинических исходов	Рассмотрение приемлемости клинических исходов	Форма публикации, в которую были включены рекомендации
Высокое	Применялось	Применялось	Описано	Практические рекомендации или анализ принятия решений*
Промежуточное	Применялось	Применялось или не применялось	Не описано	Систематический обзор*
Низкое	Не применялось	Применялось или не применялось	Не описано	Описательный обзор; оригинальное исследование

Примечание. * — методологическое качество клинических рекомендаций может быть высоким или низким вне зависимости от формы публикации, в которую они были включены. Например, если авторы практических рекомендации или анализа принятия решений не применяют систематических методов при поиске или обобщении информации, не рассматривают во всей полноте предпочтения больных или интересы общества, методологическое качество предлагаемых рекомендаций будет низким. Наоборот, если при подготовке систематического обзора учитываются все соответствующие вмешательства и хотя бы качественно оцениваются предпочтения, методологическое качество рекомендаций будет высоким.

НУЖНЫ ЛИ ВООБЩЕ КЛИНИЧЕСКИЕ РЕКОМЕНДАЦИИ?

Подходы, которые мы описали, указывают на то, что принятие клинического решения зависит не только от доказательств эффективности вмешательств, но и от предпочтений больных или интересов общества. Приемлемость методов лечения или профилактики, скорее всего, будет существенно различаться в разных ситуациях. Так, мониторинг терапии антикоагулянтами может оказаться неприемлемым в сельской местности, где для его осуществления потребуются преодолевать большие расстояния, или в условиях дефицита ресурсов здравоохранения, когда расходы на мониторинг заставят сократить закупки некоторых необходимых лекарственных препаратов, например антибиотиков.

Не менее важны и различия в предпочтениях отдельных больных. Даже при одинаковых условиях оказания медицинской помощи относительная приемлемость мониторинга терапии антикоагулянтами, инсульта и желудочно-кишечного кровотечения будет широко различаться. Логично спросить: если принятие решений в такой степени зависит от приемлемости клинических исходов, зачем вообще нужны клинические рекомендации?

Здесь можно было бы отметить следующее. Вместо того чтобы составлять клинические рекомендации с сомнительной обоснованностью, исследователи должны осуществлять систематический поиск, отбор и обобщение данных, чтобы передать в распоряжение врачей информацию об эффективности методов лечения (профилактики) и клинических исходах, имеющих значение для больных. В свою очередь врачи будут принимать клинические решения, учитывая предпочтения конкретного больного и особенности своей повседневной практики.

С этим нельзя не согласиться. Однако успешным применение такого подхода на практике будет только при наличии стандартных методов обобщения и представления информации для врачей в удобной для них форме. Кроме того, при этом врачи должны располагать временем и владеть приемами, необходимыми для оценки приемлемости клинических исходов для конкретного больного, информация о которой затем будет интегрирована с данными систематических обзоров. Вряд ли эти требования будут полностью удовлетворены в ближайшем будущем. Скорее всего, клинические рекомендации не потеряют своего значения как увеличительное стекло для пристального изучения сложившейся ситуации; как веха, отделяющая вчерашнюю отсталость от сегодняшних достижений; как напоминание о недостаточности наших знаний и как повод для спора, в котором, как известно, рождается истина. В любом случае врачи чаще будут принимать правильные клинические решения, если осознают мотивы и движущие силы этого процесса и будут критичнее относиться к предлагаемым их вниманию клиническим рекомендациям. Этому должна способствовать предложенная нами классификация.

ПРИНЯТИЕ КЛИНИЧЕСКОГО РЕШЕНИЯ

Вернемся к нашему клиническому примеру. Оценивая достоверность практических рекомендаций по применению тромболитической терапии при ФП [2], следует начать с вопроса, действительно ли составители рассматривали все соответствующие категории больных, все вмеша-

тельства и клинические исходы. Вы отмечаете, что рекомендации различаются в зависимости от разного исходного риска развития инсульта; однако при этом не учитывается исходный риск возникновения кровотечений. Это упущение может быть связано с тем, что в исследованиях, посвященных изучению прогноза и положенных в основу рекомендаций, данные о факторах риска возникновения кровотечений оказались слишком неоднородными. В случае с конкретной больной вы исключили необходимость применения антиаритмических средств (которые в другом доступном вам анализе принятия решений предлагались в качестве препаратов выбора [42]). В рекомендациях оцениваются все интересующие вмешательства, т.е. прием варфарина или аспирина в фиксированной либо скорректированной по результатам исследования свертывающей системы крови дозе; однако ничего не говорится о предложенном вашим эксцентричным коллегой клопидогреле и сходном с ним тиклопидине.

В рекомендациях описаны основные возможные клинические исходы, включая эмболический инсульт, геморрагический инсульт, желудочно-кишечное и другие тяжелые кровотечения, но не уделяется должного внимания необходимости регулярного исследования свертывающей системы крови или возникновению кровоподтеков при лечении варфарином.

Переходя к оценке качества отбора и обобщения данных, вы обнаруживаете, что критерии включения, использовавшиеся авторами рекомендаций, вполне приемлемы, а поиск в литературе, как указано на сайте Центра анализа и синтеза информации, был исчерпывающим. Методы обобщения данных не оговариваются особо, однако из текста понятно, что они основаны на подсчете и сравнении абсолютных и относительных значений частоты развития благоприятных и неблагоприятных клинических исходов, и положительно сказались на степени обоснованности рекомендаций.

Составители считают, что предпочтения больных должны играть решающую роль в процессе принятия решения, однако данные об относительной приемлемости развития инсульта и кровотечения, которые легли в основу рекомендаций, не представлены. Предпочтения составителей руководства склоняются к приему варфарина в скорректированных дозах при высоком риске развития инсульта и аспирина при низком риске. Поскольку при высоком риске развития инсульта кровотечение может возникнуть даже на фоне приема варфарина, а при низком риске инсульта вероятность его развития снижается на фоне приема антикоагулянтов, в данных рекомендациях приемлемость предупре-

ждения инсульта косвенно сопоставлялась с приемлемостью возникновения тяжелых и незначительных кровотечений, а также с неудобствами, связанными с применением варфарина.

Если, как в данном случае, процесс оценки приемлемости клинических исходов не описан подробно, читатель должен самостоятельно рассмотреть информацию о профессиональной принадлежности и возможной материальной заинтересованности экспертов, которые принимали участие в разработке клинических рекомендаций. Все авторы рекомендаций были квалифицированными специалистами; в группу составителей не включались врачи общей практики и больные. Разработку данных рекомендаций, которые были опубликованы в виде приложения к журналу *Chest* [2], финансировала фармацевтическая компания Dupont, производящая варфарин. Это заставляет о многом задуматься, поскольку финансирующая организация может повлиять на проведение научного исследования. Если составители рекомендаций не представляют исчерпывающих сведений о приемлемости клинических исходов (а это бывает весьма часто), риск возникновения систематической ошибки, связанной с конфликтом интересов, особенно велик.

Авторы использовали оригинальную схему, на основании которой мы разработали описанную выше модификацию, и при подготовке рекомендаций рассматривали только РКИ с однородными результатами; поэтому данные рекомендации отнесены к категории А (см. табл. 2). Рекомендации по профилактике инсульта при высоком (прием варфарина) и низком риске развития этого осложнения (прием аспирина) были классифицированы как соответствующие категории 1. Это значит, что авторы считают соотношение преимуществ и недостатков в обоих случаях очевидным. В нашем клиническом примере риск развития инсульта у больной следует рассматривать как промежуточный. В соответствии с клиническими рекомендациями в данном случае можно применять как варфарин, так и аспирин. В целом публикация отвечает большинству критериев достоверности, и вы склоняетесь к тому, что рекомендации авторов заслуживают доверия.

При анализе принятия решений [5] назначение варфарина сопоставляли только с отсутствием лечения. Сравнение с аспирином не проводилось, потому что эффективность этого препарата не доказана (в других мета-анализах размер эффекта аспирина едва достигал уровня клинической значимости). Применение других антиагрегантов не рассматривалось. К оцениваемым клиническим исходам авторы отнесли неудобства, связанные с мониторингом терапии антикоагулянтами; тяжелое кровотечение; инсульт легкой формы; тяжелый инсульт и стоимость

вмешательства. Незначительное кровотечение среди них даже не упоминается.

Исследователи описали свою поисковую стратегию очень ясно. Они ограничились поиском статей, опубликованных на сайтах в Интернете, однако в указанных границах этот поиск был исчерпывающим. Ясно описаны и убедительно обоснованы критерии отбора данных. Отмечена проблематичность одного из ключевых решений: отбирать для расчета показателей риска данные Фремингемского исследования или РКИ, посвященных лечению больных с ФП.

Чтобы оценить приемлемость клинических исходов, авторы опросили 57 проживающих дома пожилых людей (средний возраст 73 года), применив при обработке данных элементы теории игр (см. главу 2В2 на компакт-диске). По шкале от 0 (смерть) до 1,0 (полное здоровье) приемлемость клинических исходов составила 0,986 для терапии варфарином без осложнений, контролируемой врачом общей практики; 0,880 для тяжелого кровотечения; 0,675 для легких форм инсульта и 0 для тяжелого инсульта.

Исследователи провели анализ чувствительности и показали, что применявшаяся ими модель чувствительна к изменению предположений о приемлемости клинических исходов в группе варфарина. Если значение приемлемости относится к верхнему квартилю (т.е. приближается к 1,0; что указывает на полную приемлемость клинических исходов), то, согласно результатам анализа, почти всем больным следует назначать варфарин. Если же значение приемлемости относится к нижнему квартилю, результаты анализа указывают на необходимость отказа от приема варфарина у большинства больных.

При оценке по критериям достоверности, представленным в табл. 1, рассматриваемый анализ принятия решений заслуживает высокой оценки. Показатели приемлемости и риска, которые рассчитывались с использованием медианы предпочтений больных, а также самых точных оценок риска и степени его снижении (анализ *основного случая*), почти целиком соответствуют аналогичным показателям у больной с ФП из нашего основного примера. Исследователи представили таблицы, с помощью которых в разных клинических ситуациях можно принять оптимальное решение. Поместив в них данные о характеристиках нашей больной, получаем ответ: лечение не будет эффективным. Тем не менее этот клинический случай, скорее всего, соответствует ячейке таблицы, расположенной на границе между отсутствием эффекта и наличием четкого эффекта. Анализ чувствительности позволяет предположить, что при одинаковой ценности жизни

вне зависимости от приема варфарина польза от применения этого препарата перевесит его недостатки

Рассмотрев практические рекомендации и анализ принятия решений, а также убедившись в их высоком методологическом качестве, вы с большим основанием можете принимать клинические решения и давать соответствующие советы своим коллегам. Однако вы несколько обескуражены тем фактом, что лучшее решение для многих больных (в том числе и для вашей больной с ФП) в значительной мере зависит от их собственных предпочтений. Вы решаете более подробно обсудить возможные вмешательства и их последствия с больной во время следующего визита (см. главу 2F на компакт-диске).

Литература

- 1 *Eddy DMA* Manual for Assessing Health Practices and Designing Practice Policies The Explicit Approach Philadelphia American College of Physicians, 1992
- 2 *Laupras A , Albers G , Dalen J , Dunn M I , Jacobson A K , Singer D E* Antithrombotic therapy in atrial fibrillation *Chest* 1998,114 579S-589S
- 3 *Shanefell T M , Ma\o-Smith M F , Rothwangl J* Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature *JAMA* 1999,281 1900-1905
- 4 *Gnlli R , Magnm N , Penna A , Mura G , Liberati A* Practice guidelines developed by specialty societies the need for a critical appraisal *Lancet* 2000,355 103-106
- 5 *Thompson R , Parkin D , Eccles M , Sudlow M , Robinson A* Decision analysis and guidelines for anticoagulant therapy to prevent stroke in patients with atrial fibrillation *Lancet* 2000,355 956-962
- 6 *Vandenbroucke-Grauh C M JE , Vandenbroucke J P* Effect of selective decontamination of the digestive tract on respiratory tract infections and mortality in the intensive care unit *Lancet* 1991,338 859-862
- 7 *Selective Decontamination of the Digestive Tract Trialists' Collaborative Group* Meta-analysis of randomised controlled trials of selective decontamination of the digestive tract *BMJ* 1993,307 525-532
- 8 *Heyland D K , Cook D J , Jaeschke R , Griffith L , Lee H N , Guyatt G H* Selective decontamination of the digestive tract *Chest* 1994,105 1221-1229
- 9 *Kollef M H* The role of selective digestive tract decontamination on mortality and respiratory tract infections *Chest* 1994,105 1101-1108
- 10 *Glatziou PP , Irwig L M* An evidence based approach to individualising treatment *BMJ* 1995 311 1366-1358
- 11 *Sinclair J C , Cook R , Guyatt G H , Pauker S G , Cook D J* When should an effective treatment be used? Derivation of the threshold number needed to treat and the minimum event rate for treatment *J Clin Epidemiol* In press
- 12 *Smith G D , Egger M* Who benefits from medical interventions? *BMJ* 1994,308 72-74
- 13 *Amman E M , Lau J , Kupelnick B , Modeller F , Chalmers T C* A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts treatments for myocardial infarction *JAMA* 1992 268 240-248
- 14 *Keene RL* Decision analysis an overview *Operations Res* 1982,40 803-838
- 15 *Eckman M H , Lexine H J , Pauker S G* Decision analytic and cost-effectiveness issues concerning anticoagulant prophylaxis in heart disease *Chest* 1992,102 538S-549S

- 16 *Kassirer JP, Moskowitz A J, Lau J, Pauker S G* Decision analysis a progress report *Ann Intern Med* 1987;106:275-291
- 17 *Eddy D M* Clinical decision making from theory to practice Designing a practice policy Standards guidelines and options *JAMA* 1990;263:3077, 3081, 3084
- 18 *Wong J B, Salem D N, Pauker S G* You're never too old *N Engl J Med* 1993;328:971-974
- 19 *Krahn M D, Mahoney J E, Eckman M H, Trachtenberg J, Pauker S G, Detsky A S* Screening for prostate cancer a decision analytic view *JAMA* 1994;272:773-780
- 20 *Krahn M, Naylor C D, Basinski A S, Detsky A S* Comparison of an aggressive (US) and a less aggressive (Canadian) policy for cholesterol screening and treatment *Ann Intern Med* 1991;115:248-255
- 21 *Gael V* Decision analysis applications and limitations *CMAJ* 1992;147:413-417
- 22 *Drummond M F, Richardson WS, O'Brien B, Levine M, Heyland D K, for the Evidence-Based Medicine Working Group* Users' Guides to the Medical Literature XIII How to use an article on economic analysis of clinical practice A Are the results of the study valid? *JAMA* 1997;277:1552-1557
- 23 *O'Brien B J, Heyland D K, Richardson WS, Levine M, Drummond M F, for the Evidence-Based Medicine Working Group* Users' Guides to the Medical Literature XIII How to use an article on economic analysis of clinical practice B What are the results and will they help me in caring for my patients? *JAMA* 1997;277:1802-1806
- 24 *Institute of Medicine* Clinical Practice Guidelines Directions for a New Program Washington, DC National Academy Press 1990
- 25 *AMA/Specialty Society Practice Parameters Partnership* Attributes to Guide the Development of Practice Parameters Chicago American Medical Association, 1990
- 26 *American College of Physicians* Clinical Efficacy Assessment Project Procedural Manual Philadelphia American College of Physicians, 1986
- 27 *Gottlieb L K, Margolis C Z, Schoenbaum S C* Clinical practice guidelines at an HMO development and implementation in a quality improvement model *QRB* 1990;16:80-86
- 28 *Lohr K N, Field MJ* A provisional instrument for assessing clinical practice guidelines 1991 Unpublished
- 29 *Woolf S H* Expert Panel on Preventive Services analytic methodology 1991 Unpublished
- 30 *Park R E, Fink A, Brook R H, et al* Physicians' rating of appropriate indications for six medical and surgical procedures *Am J Public Health* 1986;76:766-772
- 31 *Barza M, Pauker S G* The decision to biopsy, treat, or wait in suspected herpes encephalitis *Ann Intern Med* 1980;92:641-649
- 32 *Matchar D B, McCrory D C, Barnett H J M, Feussner J R* Medical treatment for stroke prevention *Ann Intern Med* 1994;121:41-53
- 33 *American College of Physicians* Guidelines for medical treatment for stroke prevention *Ann Intern Med* 1994;121:54-55
- 34 *North American Symptomatic Carotid Endarterectomy Trial Collaborators* Beneficial effect of carotid endarterectomy in symptomatic patients with high-grade carotid stenosis *N Engl J Med* 1991;325:445-453
- 35 *Grady D, Rubin S M, Petitti D B, et al* Hormone therapy to prevent disease and prolong life in postmenopausal women *Ann Intern Med* 1992;117:1016-1037
- 36 *Canadian Task Force on the Periodic Health Examination* The periodic health examination *CMAJ* 1979;121:1193-1254
- 37 *Woolf S H, Battista RN, Anderson G M, Logan A G, Wang E* Assessing the clinical effectiveness of preventive maneuvers analytic principles and systematic methods in reviewing evidence and developing clinical practice recommendations *J Clin Epidemiol* 1990;43:891-905
- 38 *Sackett D L* Rules of evidence and clinical recommendations on the use of antithrombotic agents *Arch Wern Med* 1986;146:464-466
- 39 *Wagner S R, Reidinger M S, Conner L, et al* Practice guidelines and reminders to reduce duration of hospital stay for patients with chest pain *Ann Intern Med* 1994;120:257-263

40. *Hulley S., Grady D., Bush T., Furberg C., Herrington D., Riggs B., Vininghoff E.* Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. Heart and Estrogen/progestin Replacement Study (HERS) Research Group. *JAMA* 1998;280:605-613
41. *Bayoumi A.M., Redelmeier D.A.* Preventing *Mycobacterium avium* complex in patients who are using protease inhibitors: a cost-effectiveness analysis. *AIDS* 1995;9:1503-1512.
42. *Catherwood E., Fitzpatrick W.D., Greenberg M.L., et al.* Cost-effectiveness of cardioversion and antiarrhythmic therapy in non-valvular atrial fibrillation. *Ann Intern Med* 1999;130:625-636.

ЧАСТЬ • 2

Более детальные сведения: практическое применение принципов доказательной медицины и обучение ее основам

- 1 • ЛЕЧЕНИЕ И РИСК: ПРИЧИНЫ НЕДОСТОВЕРНОСТИ РЕЗУЛЬТАТОВ — СИСТЕМАТИЧЕСКИЕ *И* СЛУЧАЙНЫЕ ОШИБКИ
- 2 • ЛЕЧЕНИЕ И ПОНИМАНИЕ РЕЗУЛЬТАТОВ **ИССЛЕДОВАНИЯ: ПРОВЕРКА ГИПОТЕЗ**
- 3 • ЛЕЧЕНИЕ И ПОНИМАНИЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ: ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ
- 4 • ЛЕЧЕНИЕ И ПОНИМАНИЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ; ИЗМЕРЕНИЕ КОРРЕЛЯЦИИ

В печатном варианте второй части книги более детально обсуждаются основные положения, изложенные в первой части. Специально для тех, кто хотел бы полнее понять методы статистической обработки данных, мы подробно остановимся на некоторых понятиях, таких как систематическая и случайная ошибка, проверка исходной гипотезы, доверительные интервалы и оценка взаимосвязи между явлениями.

В электронном варианте книги каждая глава второй части имеет прямую ссылку на соответствующий раздел первой части и наоборот. Во второй части представлены рекомендуемые подходы к изложению наиболее сложных концепций доказательной медицины; они могут оказаться особенно полезными для преподавателей.

Вторая часть книги на компакт-диске начинается с раздела, в котором продолжено обсуждение вопросов, рассматриваемых в главе 1 первой части *Введение: философская концепция доказательной медицины*, в том числе о критических замечаниях в адрес доказательной медицины. Следующая глава посвящена проблеме, которая нередко служит причиной получения недостоверных результатов; речь идет о наличии систематических и случайных ошибок. Мы сочли необходимым подчеркнуть тесную связь между исследованиями, посвященными оценке эффективности и безопасности лечения. Данная глава включена и в печатный вариант книги.

В самой большой главе второй части книги рассматриваются способы наиболее эффективного использования результатов исследований при лечении конкретных больных. Обсуждаются случаи, когда результаты рандомизированных контролируемых испытаний отличаются от прогнозируемых или противоречат данным, полученным в ходе обсервационных исследований. Кроме того, для целого ряда медицинских вмешательств приводятся значения такого показателя, как ЧБНЛ (т.е. число больных, которых необходимо лечить определенным методом в течение определенного времени, чтобы достичь благоприятного исхода у одного больного или предотвратить один неблагоприятный клинический исход); особое внимание уделено зависимости данного показателя от величины исходного риска.

В этом разделе книги рассматриваются также различные аспекты статистической обработки данных, в частности анализ, проводимый исходя из допущения, что все больные получили предписанное лечение; проверка исходной гипотезы; доверительные интервалы и выбор самого наглядного способа представления размера эффекта (глава 4 *Лечение и понимание результатов исследования: измерение корреляции*). Некоторые из этих глав вошли и в печатный вариант книги. Мы приводим реко-

мендации по практическому использованию результатов рандомизированных контролируемых испытаний при лечении конкретных больных и выборе наилучшего метода лечения в клинических испытаниях с участием отдельно взятого больного (испытания N=1).

Особое внимание уделено случаям, когда вмешательство явно влияет на биологические параметры, но не влияет на клинические исходы, имеющие значение для больного («Использование косвенных критериев оценки»). Подробно рассмотрены клинические исходы, которыми авторы рандомизированных испытаний нередко пренебрегают (например, качество жизни). Предложен подход, позволяющий либо судить о применимости результатов, полученных при изучении какого-либо одного препарата определенного фармакологического класса, ко всем препаратам этого класса, либо сделать вывод о большей эффективности одного из препаратов по сравнению с остальными. Наконец, рассматриваются исследования двух конкретных типов: посвященные оценке эффективности компьютерных систем, облегчающих принятие клинических решений, а также исследования с использованием качественных критериев оценки, в которых особое внимание уделяется мнению больных о качестве жизни и медицинской помощи.

Во второй части книги более детально рассматриваются исследования, посвященные диагностике (в частности, те из них, в которых описываются клинические проявления заболеваний, разрабатываются или проходят проверку правила клинического прогнозирования). Читатель узнает о том, что совпадение мнений (например, мнений врачей, оценивающих определенный результат обследования конкретного больного) может быть обусловлено влиянием случайных факторов, а также о методах, позволяющих нивелировать это влияние. Далее приводятся примеры расчета отношения правдоподобия — показателя, который служит для представления результатов исследований диагностических тестов. В главе 4 *Лечение и понимание результатов исследования: измерение корреляции*, обсуждаются разные методы статистического анализа.

В следующем разделе второй части на компакт-диске более детально обсуждаются вопросы использования систематических обзоров. Почему результаты оригинальных исследований по одному и тому же вопросу могут противоречить друг другу? Когда можно доверять данным, свидетельствующим о том, что лечение тем или иным препаратом оказывает положительный эффект в одной группе больных, но неэффективно в другой? Достоверны ли данные о том, что применение разных доз препарата по-разному влияет на клинический исход? Кроме того, в этом

разделе обсуждаются трудности, возникающие при включении в систематические обзоры неопубликованных материалов, а также выбор оптимальных методов статистического анализа.

Во второй части более подробно рассматривается применение имеющихся данных при лечении конкретного больного или группы больных со схожими характеристиками. Изучение этого раздела поможет оценить достоверность клинических рекомендаций и понимать результаты исследований, посвященных сравнительной экономической оценке эффективности разных лечебных вмешательств. Детально обсуждаются исследования, посвященные оценке эффективности скрининговых обследований и качества медицинской помощи. Рассматривается также одна из важнейших проблем доказательной медицины: каким образом следует в полной мере учитывать предпочтения конкретного больного при принятии клинического решения, которое может повлиять на всю дальнейшую жизнь этого больного.



ЛЕЧЕНИЕ И РИСК:
ПРИЧИНЫ НЕДОСТОВЕРНОСТИ
РЕЗУЛЬТАТОВ —
СИСТЕМАТИЧЕСКИЕ
И СЛУЧАЙНЫЕ ОШИБКИ

Г. Гайятт

В подготовке данной главы принимали также участие члены рабочей группы по разработке и внедрению принципов доказательной медицины Ш. Строе, Д. Кук и П. Уайер

В ЭТОМ РАЗДЕЛЕ

Случайная ошибка

Систематическая ошибка

Оценка влияния систематической **и случайной ошибки**

Методы снижения вероятности **возникновения систематической** ошибки в исследованиях, посвященных **эффективности лечения** и риску

СЛУЧАЙНАЯ ОШИБКА

На каждый клинический вопрос есть правильный ответ, который соответствует истине. Например, применение β -блокаторов определенным образом влияет на уровень смертности при сердечной недостаточности, ингаляции кортикостероидных препаратов — на частоту обострений бронхиальной астмы, а каротидная эндартерэктомия — на риск развития инсульта у больных с преходящими нарушениями мозгового кровообращения. К сожалению, нам никогда не известно, насколько выражены эти влияния. Почему?

Представьте, что подбрасываете идеально сбалансированную монету. Каждый раз вероятность выпадения орла или решки составляет 50%. Допустим, что вы исследователь и не знаете о том, что монета идеально сбалансирована, а точнее не имеете представления о том, насколько хорошо она сбалансирована и хотели бы это выяснить. Сформулировав вопрос — какова истинная вероятность выпадения орла или решки — вы подбрасываете монету 10 раз подряд. Допустим, что орел выпал 8 раз, а решка два. Какой вывод можно сделать? Если подойти к результатам эксперимента формально, может сложиться впечатление, что монета сбалансирована плохо (так как орел выпадает намного чаще, чем решка), и что вероятность выпадения орла при каждом подбрасывании составляет 80%.

Подобное заключение удовлетворит немногих, так как мы знаем, что на самом деле даже у идеально сбалансированной монеты орел и решка не всегда выпадают одинаково часто, т.е. результат зависит от случая. Иными словами, в любом эксперименте возможна *случайная ошибка*. При подбрасывании хорошо сбалансированной монеты орел иногда выпадает 8, реже — 9 и совсем редко — 10 из 10 раз.

А что, если орел и решка выпадали одинаково часто (т.е. по 5 из 10 Раз)? Поскольку при проведении данного теста большую роль играет случай, мы усомнимся в том, что полученные результаты достоверны. Нам известно, что, во-первых, орел может выпасть 8, а решка 2 раза в тех случаях, когда истинная вероятность выпадения орла составляет 0,5, а, во-вторых, обе стороны даже плохо сбалансированной монеты могут выпасть одинаково часто (например, когда истинная вероятность выпадения орла составляет 0,8).

Допустим, что наш маленький эксперимент заинтересовал СПОНСОРОВ, которые выделили средства для проведения более крупного исследования. Это позволило значительно увеличить размеры выборки (монету подбрасывали 1000 раз). После того как орел и решка выпали по

500 раз, можно ли будет утверждать, что полученные данные отражают истину? Нет. Мы знаем, что даже в тех случаях, когда истинная вероятность выпадения орла составляет 51%, подбросив монету 1000 раз, можно получить результаты, аналогичные описанным.

Влияние случайных факторов в медицинских исследованиях. Логические рассуждения, изложенные выше, можно применять при оценке результатов медицинских исследований. Так, данные, полученные в рандомизированном контролируемом испытании (РКИ), свидетельствуют о том, что в группе лечения умирают 10 из 100, а в группе контроля — 20 из 100 больных. Значит ли это, что проводимая терапия действительно снижает уровень смертности на 50%? Возможно, но нельзя быть в этом уверенным, так как достоверно неизвестно, насколько эффективно лечение (если оно вообще эффективно). Приведем реальный пример: в 1 плацебо-контролируемом исследовании эффективности бисопролола при умеренно тяжелой и тяжелой сердечной недостаточности в группе контроля умерли 228 (17%) из 1320 больных, в группе лечения — 156 (12%) из 1327 больных [1]. Несмотря на то что истинное снижение относительного риска (ОР) смерти, скорее всего, близко к величине, полученной в исследовании (34%), полученные данные могут оказаться и ошибочными. В начале главы мы задали вопрос "Почему никогда нельзя быть абсолютно уверенными в размере эффекта лечения, независимо от статистической мощности исследования и того, насколько хорошо оно было спланировано?" Потому что всегда возможна случайная ошибка.

СИСТЕМАТИЧЕСКАЯ ОШИБКА

Что имеют в виду, когда говорят о достоверности или обоснованности результатов исследования? Под *достоверностью* подразумевается степень, в которой результат исследования позволяет правильно ответить на поставленный вопрос или измерить тот или иной показатель. В данной книге достоверность используется в качестве технического термина, с помощью которого описывают вероятность возникновения *систематической ошибки*. В отличие от случайной систематическая ошибка приводит к появлению систематического отклонения от истинного значения в определенную сторону. В исследованиях, посвященных оценке эффективности и безопасности лечения, наличие систематической ошибки сопровождается недооценкой или переоценкой его преимуществ или недостатков.

Систематическая ошибка может быть обусловлена тем, что группы лечения и контроля изначально различались по влиянию факторов, не

„меющих отношения к изучаемому вмешательству. В то же время она может стать следствием различий, возникших в ходе исследования. Клинические исходы у больных, не получающих терапию, могут быть благоприятными или неблагоприятными. Под неблагоприятными клиническими исходами подразумевают осложнения, которые развиваются в процессе исследования (например, инсульт). Именно они часто являются объектами изучения или *оцениваемыми клиническими исходами*.

Вероятность развития оцениваемого (неблагоприятного) клинического исхода зависит от множества факторов. Например, риск развития инсульта будет выше у больных пожилого возраста, мужчин, при наличии выраженных патологических изменений сосудов (атеросклероз) и артериальной гипертонии [2]. Каждая из перечисленных характеристик участников исследования называется *прогностическим фактором*, или *фактором, определяющим клинический исход*. Именно эти факторы определяют вероятность развития у конкретного больного оцениваемого клинического исхода.

В отличие от перечисленных другие характеристики участников, такие как цвет глаз или размер обуви, едва ли способны повлиять на риск развития инсульта. У голубоглазых людей или лиц, носящих 12-й размер обуви, риск развития данного осложнения не выше и не ниже, чем у кареглазых или тех, кто носит 8-й размер.

Различия между группами лечения и контроля, влияющие на прогноз. Систематическая ошибка появляется в том случае, если группы лечения и контроля изначально различаются по какому-либо фактору, способному предопределить развитие оцениваемого клинического исхода. В отличие от различий по цвету глаз и размеру обуви такие факторы риска могут привести к появлению систематической ошибки. Например, если в группе лечения атеросклероз более выраженный, или возраст участников старше, в ней будет выше частота развития неблагоприятных клинических исходов, чем в группе контроля. В данном случае возникнет систематическая ошибка, приводящая к занижению эффективности лечения, а полученные результаты будут неистинными (недостоверными).

Если в контрольную группу будет включено больше мужчин или больных с более высокими средними показателями артериального давления, чем в группу вмешательства, возникнет систематическая ошибка, ведущая к переоценке эффективности лечения. Таким образом, одна из причин появления систематической ошибки заключается в Исходных различиях между группами лечения и контроля, влияющих на прогноз.

Эффект плацебо. Даже в том случае, когда факторы риска в основной и контрольной группах не различаются, эффективность вмешательства может быть оценена неверно. Так, больные, уверенные в получении активного лечения, могут ожидать наступления улучшения, и само это ожидание приведет к положительным переменам в их самочувствии и даже функциональном статусе. Таким образом, *эффект плацебо* может приводить к завышению реальной эффективности лечения.

Дифференцированный подход к проведению вмешательств. Другой причиной появления систематической ошибки может стать дифференцированный подход к проведению дополнительных вмешательств, отличных от тех, которые изучаются в данном исследовании, в группах лечения и контроля. Например, если при проведении испытания нового препарата, снижающего риск развития инсульта, в основную группу будет включено больше больных, которые получают аспирин или клопидогрел, чем в группу контроля, оценка эффективности изучаемого лечения будет завышена. Однако, если в основной группе большая доля больных будет также применять изотонический раствор натрия хлорида в виде глазных капель или антацидные препараты, то это никак не отразится на вероятности появления систематической ошибки, поскольку прием аспирина и клопидогрела снижает риск развития инсульта, а применение таких глазных капель или антацидных препаратов — нет. *Сопутствующее вмешательство* используется в данной книге в качестве технического термина, описывающего применение в одной из групп сравнения дополнительных методов лечения, влияющих на частоту развития оцениваемого клинического исхода.

Следует отметить, что нас не волнуют исходные различия между участниками по цвету глаз или размеру обуви (*исходные характеристики*), так же как и различия по применению глазных капель или антацидных препаратов в течение периода наблюдения. Однако мы обратим самое пристальное внимание на несоответствие по тяжести заболевания и частоте приема аспирина в сравниваемых группах, поскольку эти факторы влияют на риск развития инсульта. Таким образом, результаты исследования могут быть смещены под влиянием систематической ошибки, если в группах сравнения различаются исходные характеристики больных или сопутствующие вмешательства, влияющие на прогноз. Термином *искажающий фактор* обозначают любое влияние, имеющее прогностическое значение (включая характеристики больных и вмешательства), которое в группах лечения и контроля различается. Во избежание систематической ошибки группы должны быть сходными (сравнимыми) до начала исследования и оставаться таковыми на всем его протяжении.

Дифференцированный подход к оценке изучаемого клинического исхода - Причиной систематической ошибки также может стать дифференцированный подход к определению изучаемого клинического исхода. Например, решение о том, перенес ли больной преходящее нарушение мозгового кровообращения или микроинсульт, может быть принято в результате совместного обсуждения. Если же в группе контроля данный клинический исход расценивается как инсульт, а в группе лечения — как преходящее нарушение мозгового кровообращения, эффективность терапии будет переоценена.

Выбывание участников из исследования. Систематическая ошибка может появляться в тех случаях, когда из исследования выбывает большее число больных. Если частота оцениваемого неблагоприятного клинического исхода у выбывших больных и остальных участников неодинакова, полученные результаты могут оказаться ошибочными.

ОЦЕНКА ВЛИЯНИЯ СИСТЕМАТИЧЕСКОЙ И СЛУЧАЙНОЙ ОШИБКИ

При изучении принципов доказательной медицины можно столкнуться с непониманием концепций и путаницей в терминологии. На вопрос, что делает результаты исследования достоверными, студенты часто отвечают: "Большие размеры выборки". Однако маленькая выборка не способствует появлению систематической ошибки (и потому не снижает достоверность полученных данных), но повышает риск возникновения случайной ошибки. Следующее упражнение поможет вам лучше уяснить смысл данных терминов.

Представьте группу исследований со сходной структурой и одинаковым числом участников, набранных в одной и той же популяции. Как и в случае с подбрасыванием монеты, когда частота выпадения орла и решки не всегда одинакова, результаты исследований будут различаться в связи с влиянием случайных факторов.

Теперь представьте 4 такие группы, в две из которых были включены исследования с малой выборкой, в остальные две — с большой.

В две группы были включены РКИ с применением слепого метода по отношению к больным, медицинскому персоналу и исследователям, оценивающим клинические исходы (одна группа исследований с малыми "одна — с большими выборками), в остальные две — наблюдательные исследования, в которых больных включали в ту или иную группу с учетом их пожеланий, мнения лечащего врача или без определенной системы. Риск появления систематических ошибок в наблюдательных исследованиях гораздо выше, чем в РКИ. В рассматриваемом примере

нам будет предоставлена уникальная возможность, которой лишены врачи в реальных условиях — знать об истинном эффекте вмешательства. На рисунке он обозначен черным кружком, а мелкие точки представляют не отдельных больных, а результаты одного исследования. Чем больше расстояние между точкой и кружком, тем менее достоверны результаты конкретного лечения.

Прежде чем продолжить чтение, внимательно изучите рисунок и попробуйте сделать собственные выводы о влиянии структуры исследований и их выборки на полученные данные.

Фрагмент А отображает результаты серии крупных РКИ. В связи с высокой статистической мощностью РКИ полученные данные достоверны, поэтому точки (результаты отдельных РКИ) сгруппированы вокруг кружка. Точки выходят за его пределы из-за возникновения случайной ошибки, однако при больших выборках снижается риск ее появления, поэтому результаты отдельных РКИ приближаются к истинному размеру эффекта вмешательства.

В фрагменте В структура РКИ также является залогом того, что результаты отдельных исследований (точки) будут располагаться вокруг истинного значения (кружка). Однако при малых выборках РКИ повышается риск появления случайных ошибок, поэтому некоторые точки отстоят от черного кружка довольно далеко.

Эксперимент с подбрасыванием монеты поможет понять природу различий между результатами исследований, отображенных на фрагментах А и В. Представьте, что в каждом из РКИ идеально сбалансированную монету подбросили 10 раз; при этом результаты отдельных испытаний могут достаточно сильно отличаться от истинных (т.е. в 70 или даже 80% случаев может выпасть орел или решка). Описанная ситуация соот-

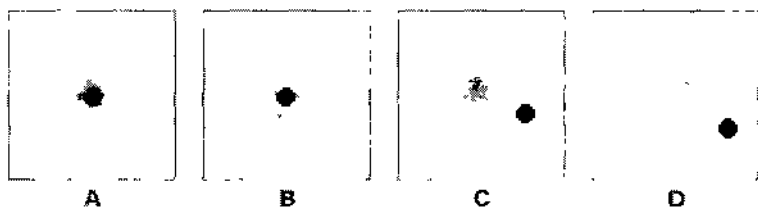


Рисунок. Зависимость частоты возникновения систематической и случайной ошибки от типа исследования и размера выборки.

А. РКИ с большим размером выборки. В. РКИ с малой выборкой. С. Observационные исследования с большим размером выборки. Д. Observационные исследования с малой выборкой

ретствует фрагменту В. Если в каждом РКИ монету подбросили 1000 раз, картина будет соответствовать фрагменту А. При этом вероятность выпадения орла или решки более чем, например, 540 раз (в 54% случаев), достаточно мала. Таким образом, в исследованиях с малой выборкой полученные результаты могут сильно отличаться от истинного значения, а в крупных РКИ, как правило, приближаются к нему.

Фрагменты А и В позволяют понять целесообразность обобщения результатов отдельных исследований. Этот процесс называют *мета-анализом*. Допустим, что информация об эффективности того или иного метода лечения получена в нескольких небольших РКИ. При этом результаты исследований значительно различаются (вследствие влияния случайных факторов). Статистическая мощьность РКИ высока, поэтому точки (результаты отдельных РКИ) располагаются вокруг кружка (истинного размера эффекта лечения). В свою очередь мы можем снизить влияние случайной ошибки и повысить достоверность результатов отдельных РКИ, обобщив их (т.е. приблизив ситуацию фрагмента В к ситуации фрагмента А).

На фрагменте С центр скопления точек (отдельные исследования) далеко отстоит от кружка (истинное значение), поскольку при проведении наблюдательных исследований даже с большой выборкой повышается риск появления систематической ошибки. Вследствие одинаковой структуры ее величина и направленность в отдельных исследованиях одинакова. При большой выборке снижается вероятность появления случайной ошибки, поэтому результаты отдельных исследований различаются мало. Тем не менее они будут недостоверными.

Приведем пример из реальной жизни. В 1 крупном наблюдательном исследовании было показано, что применение препаратов, содержащих витамин Е, снижает смертность от ишемической болезни сердца (ИБС) [5]. Однако полученные результаты не удалось подтвердить в ходе крупного хорошо спланированного РКИ [6].

А вот другой пример. В серии наблюдательных исследований с большой выборкой было выявлено, что проведение заместительной гормональной терапии (ЗГТ) в постменопаузе снижает риск смерти от ИБС на 35% [7]. Вместе с тем в первом же РКИ, в котором участвовали женщины с высоким риском развития сердечно-сосудистых осложнений, не было обнаружено различий по эффективности препаратов для ЗГТ и плацебо [8].

Объяснением в обоих случаях может служить тот факт, что препараты витамина Е и ЗГТ чаще принимают лица с более низким риском развития ИБС, а это приводит к неправильной оценке эффективности лечения.

Ситуация, представленная на фрагменте С, особенно коварна, так как большие выборки исследований внушают врачам уверенность в дос-

товерности полученных результатов. Например, многие врачи до сих пор верят в то, что ЗГТ снижает риск смерти от ИБС.

На фрагменте D, как и на фрагменте C, представлены результаты обсервационных исследований, которые значительно отличаются от истинных (вследствие систематической ошибки). Кроме того, малые выборки стали причиной выраженных различий между данными, полученными в отдельных исследованиях. Возможно, некоторые захотят провести мета-анализ результатов обсервационных исследований. Однако это весьма опасно, так как на основании данных, неточных вследствие выраженного влияния случайной ошибки, будет получен точный, но по-прежнему недостоверный результат.

МЕТОДЫ СНИЖЕНИЯ ВЕРОЯТНОСТИ ВОЗНИКНОВЕНИЯ СИСТЕМАТИЧЕСКОЙ ОШИБКИ В ИССЛЕДОВАНИЯХ, ПОСВЯЩЕННЫХ ЭФФЕКТИВНОСТИ ЛЕЧЕНИЯ И РИСКУ

Как уже отмечалось, систематические ошибки возникают вследствие того, что группы лечения и контроля исходно различаются по тем или иным прогностическим факторам либо данные различия появляются в ходе исследования. Каким образом можно уменьшить вероятность возникновения систематической ошибки? Возможные способы представлены в таблице.

В исследованиях эффективности новых методов лечения часто удается с успехом применять перечисленные методы. При этом можно снизить вероятность включения в группы вмешательства и плацебо больных с различным влиянием прогностических факторов с помощью рандомизации. Эффект плацебо уменьшается, если внешний вид активного препарата и плацебо будет одинаковым. Использование слепого метода по отношению к врачам, назначающим активное лечение или его имитацию, способно устранить возможность избирательного применения сопутствующих вмешательств, а по отношению к исследователям, оценивающим клинические исходы — значительно уменьшить вероятность возникновения систематической ошибки вследствие дифференцированного подхода к оценке клинических исходов.

Как правило, в исследованиях, посвященных вредным воздействиям или осложнениям лечения, методы для снижения вероятности возникновения систематической ошибки реже достигают своей цели, чем при оценке эффективности вмешательств. Исследователям приходится сравнивать данные о больных, подвергнутых изучаемому воздействию или

Таблица. Методы снижения вероятности возникновения систематической ошибки в исследованиях, посвященных оценке эффективности лечения и риска

Причина систематической ошибки	Методы снижения вероятности возникновения систематической ошибки в исследованиях, посвященных эффективности лечения	Методы снижения вероятности возникновения систематической ошибки в исследованиях, посвященных вредным воздействиям или побочным эффектам
<i>Исходные различия</i>		
Различие прогностических факторов в основной и контрольной группах	Рандомизация	Коррекция с учетом различий по прогностическим факторам при статистической обработке данных
<i>Различия, возникающие в процессе исследования</i>		
Эффект плацебо	Применение слепого метода в отношении больных	Выбор оцениваемых клинических исходов, в меньшей степени подверженных влиянию эффекта плацебо (например, смерть)
Сопутствующее вмешательство	Применение слепого метода в отношении врачей	Регистрация различий в проводимой терапии и внесение соответствующих поправок при проведении статистического анализа
Систематическая ошибка вследствие Дифференцированного подхода к оценке клинических исходов	Применение слепого метода в отношении исследователей, оценивающих клинические исходы	Выбор оцениваемых клинических исходов, в меньшей степени зависящих от мнения исследователей (например, смерть)
Выбывание больных из исследования	Наблюдение всех больных до конца исследования	Наблюдение всех больных до конца исследования

вмешательству по их собственному выбору либо под действием обстоятельств. При этом повлиять на возможные различия по клиническим исходам можно только путем внесения при статистическом анализе соответствующих поправок на разное влияние прогностических факторов, использование слепого метода невозможно, и чтобы снизить влияние эффекта плацебо и вероятность появления систематической ошибки,

связанной с различным определением клинических исходов, необходимо выбирать те из них (например, смерть), которые менее подвержены влиянию указанных факторов. Кроме того, риск появления систематических ошибок можно снизить, сведя к минимуму частоту выбывания участников из исследования (см. таблицу).

Описанные выше правила применимы не всегда. В некоторых случаях исследователи находят затруднительным, а то и невозможным, рандомизированно включать больных в группы лечения и контроля, предпочитая проводить обсервационные исследования. При оценке методологического качества исследований с такой структурой приходится использовать критерии достоверности, разработанные для оценки вредных воздействий или осложнений лечения.

Аналогично, если применение эффективного препарата чревато развитием побочных реакций, исследователи могут рандомизированно включать больных в группы лечения и контроля. При оценке методологического качества РКИ эффективности лечения, можно использовать критерии достоверности, разработанные для соответствующих исследований. Независимо от того, рассматриваются вопросы эффективности или безопасности вмешательств, достоверность результатов РКИ почти всегда выше достоверности обсервационных исследований.

Литература

1. CIBIS-II Investigators and Committees. The Cardiac Insufficiency Bisoprolol Study II (CIBIS- II): a randomised trial. *Lancet* 1999;353:9—13
2. Goldstein L.B., Adams R., Becker K., Furberg C.D., Gorelick P.B., Hademenos G., *era/*. Primary prevention of ischaemic stroke: a statement for healthcare professionals from the Stroke Council of the American Heart Association. *Stroke* 2001;32(1):280—299
3. Gubitz G., Sandercock P., *Counsel!* C. Antiplatelet therapy for acute ischaemic stroke (Cochrane Review). In: *The Cochrane Library*; 1, 2001. Oxford. Update Software.
4. CAPRIE Steering Committee. A randomized, blinded, trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE). *Lancet* 1996;348(9038)1329—1339.
5. Knekt P., Reunanen A., Jarvinen R., Seppanen R., Heliövaara M., Aromaa A. Antioxidant vitamin intake and coronary mortality in a longitudinal population study. *Am J Epidemiol* 1994;139:1180—1189.
6. Yusuf S., Dagenais G., Pogue J., Bosch J., Sleight P. Vitamin E supplementation and cardiovascular events in high-risk patients. The Heart Outcomes Prevention Evaluation Study Investigators. *N Engl J Med* 2000;342:154—160.
7. Stampfer M.J., Colditz G.A. Estrogen replacement therapy and coronary heart disease: a quantitative assessment of the epidemiologic evidence. *Prev Med* 1991 ;20:47—63.
8. Hulley S., Grady D., Bush T., *et al.* Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. Heart and Estrogen/progestin Replacement Study (HERS) Research Group. *JAMA* 1998;280:605—613.

2

ЛЕЧЕНИЕ И ПОНИМАНИЕ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ ПРОВЕРКА ГИПОТЕЗ

Г. Гайятт, Р. Йешке, Д. Кук, С. Уолтер

В подготовке данной главы принимали также участие члены рабочей группы по разработке и внедрению принципов доказательной медицины Р. Хатала

В ЭТОМ РАЗДЕЛЕ

Роль случайных факторов

Величина α

Вероятность того, **что** существующие **различия останутся** незамеченными

Пример оценки непрерывных переменных при оценке клинических исходов

Внесение поправок на исходные различия

Одновременная проверка множества гипотез

Проблемы, с которыми приходится сталкиваться **при проверке** гипотез

Как уже отмечалось в предыдущей главе, истинный размер эффекта лечения можно оценить только в ходе эксперимента (см. главу 1 *Лечение и риск: причины недостоверности результатов — систематические и случайные ошибки*). Для более точной оценки истинного эффекта лечения исследователи применяют различные статистические методы. В некоторых исследованиях последние используются главным образом для проверки той или иной гипотезы, начиная с нулевой гипотезы, которую нужно рассмотреть и, возможно, опровергнуть. Как правило, нулевая гипотеза предполагает отсутствие различий между сравниваемыми методами лечения. В рандомизированных контролируемых испытаниях (РКИ), в которых сравнивается эффективность изучаемого метода лечения и плацебо, нулевая гипотеза будет звучать так: истинное различие по частоте наступления определенного клинического исхода в основной и контрольной группах отсутствует (т.е. равно 0). Например, в одном исследовании (804 мужчины с сердечной недостаточностью) эффективности вазодилататоров сравнивали уровень выживаемости среди больных, получавших эналаприл или гидралазин в сочетании с нитратами [1]. В начале исследования было выдвинуто предположение, что эффективность описанных схем лечения одинакова. Исследователи придерживались данной гипотезы до тех пор, пока ее не опровергли полученные результаты. В данном случае нулевая гипотеза подразумевает отсутствие истинных различий по выживаемости среди больных, принимавших эналаприл или получавших комбинированную терапию.

При проверке гипотезы статистический анализ призван ответить на вопрос, соответствуют ли полученные данные нулевой гипотезе. Логика подобного подхода такова: даже если изучаемое лечение в действительности не оказывает ни положительного ни отрицательного влияния на клинические исходы (т.е. его эффективность равна 0), результаты, полученные в основной и контрольной группах, редко совпадают полностью. Чем больше различаются полученные данные, тем менее достоверна нулевая гипотеза. В случае получения значительных различий между группами лечения "контроля исследователи вправе от нее отказаться. В дальнейшем описанная логическая цепочка будет дополнена информацией о роли, которую играют в клинических исследованиях случайные факторы.

Роль случайных факторов

В главе 1 *Лечение и риск: причины недостоверности результатов — систематические и случайные ошибки* рассматривается случай, когда вероятность выпадения орла или решки при каждом подбрасывании идеаль-

но сбалансированной монеты составляет 0,5. Было отмечено, что если подбросить такую монету 10 раз, орел и решка не обязательно выпадут одинаково часто. Иногда это соотношение может составлять 8:2 или даже 9:1. Более того, в редких случаях все 10 раз может выпасть только орел или только решка.

Подобный разброс результатов обусловлен влиянием случайных факторов. Каким образом они действуют, можно продемонстрировать на примере некоторых игр. Иногда при подбрасывании двух костей с идеально ровными гранями (т.е. костей, в которых вероятность выпадения любой цифры от 1 до 6 одинакова) могут выпасть две двойки или две шестерки. Аналогичным образом банкOMET может сдать игроку в покер 5 карт одной масти (к великой радости игрока). Еще реже эти 5 карт могут оказаться не только одинаковой масти, но и последовательного достоинства.

Влияние случайных факторов проявляется не только в вышеописанных игровых ситуациях. При включении в исследование участников из общей популяции необычная распространенность какого-либо хронического заболевания также может оказаться случайной. Кроме того, при сравнении одинаково эффективных препаратов частота наступления того или иного клинического исхода в 2 группах может также оказаться различной под влиянием случайных факторов. По большей части статистическая обработка данных направлена на определение степени, в которой подобные различия можно объяснить влиянием случайных факторов, а в каких случаях необходимо искать другие причины (например, большая или меньшая эффективность того или иного метода лечения). Как будет показано, результаты статистического анализа в большой мере зависят от размера выборки.

ВЕЛИЧИНА P

Исследователь может ошибочно предположить, что между группами лечения и контроля существуют различия, хотя на самом деле они отсутствуют. Согласно статистической терминологии, подобные ошибки называют *ошибкой Грода* (α -ошибкой), а их вероятность соответствует уровню α . Вообразите ситуацию, когда мы не имеем информации о сбалансированности монеты. Предположив, что чаще будет выпадать либо орел, либо решка, мы можем выдвинуть нулевую гипотезу, согласно которой монета сбалансирована и частота выпадения ее сторон будет одинаковой (т.е. вероятность выпадения орла и решки составит 50%). Мы можем проверить данную гипотезу экспериментальным путем, подбросив монету несколько раз. Статистический анализ полученных резуль-

татов должен будет ответить на вопрос, в какой степени они оказались следствием влияния случайных факторов.

Проведем гипотетический эксперимент. Допустим, мы подбросили монету 10 раз и все 10 раз выпал орел. Насколько вероятен подобный исход в том случае, если монета действительно идеально сбалансирована? Большинство людей придут к выводу, что вероятность получения подобного результата под действием случайных факторов очень мала. Мы готовы отвергнуть гипотезу об идеальной сбалансированности монеты (нулевую гипотезу) и сделать вывод о том, что у нее смещен центр тяжести. С помощью статистических методов мы можем более точно оценить вероятность того, что полученные результаты обусловлены влиянием случайных факторов в том случае, если нулевая гипотеза все же верна. Правило перемножения вероятностей независимых событий (когда одно из них никак не влияет на другое) гласит, что вероятность последовательного выпадения орла 10 раз рассчитывается путем перемножения вероятностей его однократного выпадения, т.е. $0,5^{10}$ и т.д. Тогда вероятность данного исхода меньше $1/1000$. В научных статьях эту вероятность отражает величина p (т.е. $p < 0,001$). Каково точное значение величины p ? Если при условии сбалансированности монеты (правильности нулевой гипотезы) повторить эксперимент с ее 10-кратным подбрасыванием много раз, орел случайно выпадет 10 раз подряд менее чем в 0,1 % серий бросков. Вероятность случайного выпадения или орла или решки 10 раз подряд составит 0,002 (2 из 1000 раз).

При проверке гипотез эксперимент не может быть завершен до тех пор, пока не будет принято окончательное решение об отказе от нулевой гипотезы, после чего исследователь делает окончательный вывод, в нашем примере — вывод о несбалансированности монеты. Выбор этапа, на котором будет прекращен эксперимент, зависит от того, насколько мы уверены в сбалансированности монеты в том случае, когда на самом деле у нее смещен центр тяжести (другими словами, насколько мы Уверены в том, что не совершим ошибку I рода). Таким образом, мы выбираем определенное пороговое значение или границу, по одну сторону от которой мы примем нулевую гипотезу, а по другую откажемся от нее. Возвращаясь к эксперименту с выпадением орла 10 раз подряд, следует отметить, что большинство людей отвергнут нулевую гипотезу при вероятности подобного исхода ниже 0,001.

Повторим эксперимент. На этот раз решка выпала 9 раз, а орел — один. Данный результат также нельзя объяснить исключительно влиянием случайных факторов. Величина p при этом составляет 0,02. Таким образом, если монета сбалансирована, а нулевая гипотеза верна, по-

добный (9 орлов и 1 решка или 9 решек и 1 орел) или описанный ранее (10 орлов или 10 решек) исход может быть случайным лишь в 2 случаях из 100.

Должны ли мы отвергнуть нулевую гипотезу в данном случае? Выбор величины p произволен и может быть оспорен. Законы статистики диктуют положительный ответ, так как за точку разделения правдоподобных и неправдоподобных результатов исследования традиционно принимают величину p , равную 0,05. Однако в некоторых случаях это не исключает возможности выбора другого значения p . При $p < 0,05$ полученные результаты расцениваются как *статистически значимые*. Это означает, что они, по всей видимости, обусловлены не только влиянием случайных факторов, а, следовательно, нулевая гипотеза неверна.

Повторим наш эксперимент еще дважды, используя каждый раз другую монету. Первый раз орел выпал 8, а решка — 2 раза. Расчет величины p показывает, что при условии сбалансированности монеты подобное (2/8 или 8/2) или более неравномерное распределение в результате действия случайных факторов встречается в 11 случаях из 100 ($p=0,11$). Таким образом, мы оказываемся по другую сторону от традиционно выбранной точки разделения правдоподобных и неправдоподобных результатов исследования. Если придерживаться общепринятого мнения, нулевую гипотезу отвергать не следует, а полученные данные следует считать *статистически незначимыми*.

Проведем эксперимент в последний раз. Допустим, что решка выпала 7, а орел 3 раза. Опыт говорит нам, что такой исход (хотя и не очень часто) может наступить даже при использовании сбалансированной монеты. Величина p , равная 0,34, подтверждает наше предположение. Полученные результаты могут встречаться в 34 случаях из 100, т.е. нулевая гипотеза оказалась верной.

Несмотря на то что проблемы, затрагиваемые в клинических испытаниях, несколько отличаются от оценки сбалансированности монеты, приведенный пример помогает понять значения величины p , которые приводятся в научных статьях. При сравнении двух методов лечения исследователи задаются вопросом: «Какова вероятность того, что наблюдаемые различия обусловлены только влиянием случайных факторов?» Если мы согласны с точкой разделения $p < 0,05$, следует отвергнуть нулевую гипотезу и сделать вывод о большей (или меньшей) эффективности изучаемого вмешательства, когда при повторении эксперимента такие же или более выраженные различия будут наблюдаться менее чем в 5% случаев.

Вернемся к РКИ, в котором сравнивали эффективность эналаприла или гидралазина в сочетании с нитратами у 804 мужчин с сердечной недостаточностью. В нем продемонстрирован метод оценки гипотезы^c

применением дихотомического клинического исхода — смерти [1]. В течение периода наблюдения от 6 мес до 5,7 года в группе эналаприла умерли 132 из 403 (33%) больных, а в группе комбинированной терапии — 153 из 401 (38%) больных. При сравнении полученных результатов (доли умерших больных) с помощью критерия χ^2 было показано, что если нулевая гипотеза верна (т.е. уровень смертности в обеих группах одинаков), то указанные или более выраженные различия между группами могут встретиться в 11 из 100 случаев ($p=0,11$). Таким образом, приняв за точку разделения величину $p<0,05$, мы не можем отвергнуть нулевую гипотезу и вынуждены считать, что наблюдаемые различия случайны.

ВЕРОЯТНОСТЬ ТОГО, ЧТО СУЩЕСТВУЮЩИЕ РАЗЛИЧИЯ ОСТАНУТСЯ НЕЗАМЕЧЕННЫМИ

Некоторые врачи могут прокомментировать сравнение эффективности эналаприла и гидралазина в сочетании с нитратами, следующим образом: «Нулевую гипотезу отвергнуть нельзя (так как $p>0,05$), однако нельзя утверждать, что эналаприл не более эффективен, чем комбинированная терапия. Полученные результаты все же неубедительны». Подобное заключение заставляет обсудить *ошибку II рода*, которую может совершить исследователь — не заметить существующих различий, т.е. отрицать блбыпую или меньшую эффективность изучаемого лечения.

Вероятность ошибочного заключения об одинаковой эффективности эналаприла и комбинированной терапии может показаться большой. Исследование показало, что прием эналаприла снижает уровень смертности на 5% по сравнению с применением вазодилататоров. Если бы различие по уровню смертности на самом деле составляло 5%, следовало бы сделать заключение, что эналаприл имеет важные преимущества. Несмотря на это, мы вынуждены придерживаться нулевой гипотезы.

Почему же исследователи, наблюдая столь значительные различия по уровню смертности, не могут сделать вывод о том, что эналаприл эффективнее гидралазина в сочетании с нитратами? Потому что число участников исследования было недостаточным для подтверждения истинности этих различий. Вероятность того, что клинически значимые различия останутся незамеченными (вероятность возникновения ошибки II рода), уменьшается по мере увеличения размера выборки. Если эта вероятность достаточно высока, статистическая мощность исследования низка. Чем больше выборка, тем меньше риск совершения ошибки II рода и тем выше его статистическая мощность. В исследование были включены 804 больных, чего,

казалось бы, должно быть достаточно, однако при выборе дихотомического клинического исхода, такого как смерть, для выявления небольшого размера эффекта лечения требуются очень большие выборки. Так, в исследовании эффективности тромболитических препаратов при остром инфаркте миокарда авторы предполагали и наблюдали различие по уровню смертности между группами лечения и контроля, не достигшее 5%. Поэтому в исследование пришлось включить несколько тысяч больных, после этого оно обрело достаточную статистическую мощность.

Если нулевую гипотезу отвергнуть не удалось (т.е. $p > 0,05$), истинная эффективность вмешательства могла быть оценена неверно. В этом случае следует решить, достаточно ли статистическая мощность исследования. При этом чем более выражены (статистически незначимые) различия между группами лечения и контроля, тем выше вероятность того, что истинный эффект терапии остался незамеченным [2]. В другой главе данной книги описывается, каким образом можно определить достаточность размеров выборки (см. главу 3 *Лечение и понимание результатов исследования: доверительные интервалы*).

Структура некоторых исследований предназначена не для того, чтобы доказать преимущества нового лечения, а для того, чтобы продемонстрировать сравнимую клиническую эффективность распространенного метода и более дешевых, простых или безопасных методов. Подобные работы часто называют *исследованиями эквивалентности вмешательств* [3]. При их проведении или оценке полученных результатов особое внимание следует уделять достаточности выборки для выявления небольших, но клинически значимых различий. В противном случае можно сделать ошибочный вывод о сходной эффективности двух методов лечения, в то время как традиционная терапия на самом деле будет иметь преимущества перед более дешевыми, простыми или безопасными вмешательствами.

ПРИМЕР ОЦЕНКИ НЕПРЕРЫВНЫХ ПЕРЕМЕННЫХ ПРИ ОЦЕНКЕ КЛИНИЧЕСКИХ ИСХОДОВ

До этого момента во всех приводимых примерах использовались такие дихотомические критерии оценки или исходы, как положительный и отрицательный результат, орел или решка, смерть или выживание; причем во всех случаях можно было рассчитать частоту наступления того или иного исхода. Однако при сравнении эффективности 2 методов лечения и более часто применяются *непрерывные данные*, имеющие потенциально бесконечное число возможных значений с небольшими разли-

чиями между ними (например, результаты спирометрии, величина сердечного выброса, клиренс креатинина или показатели качества жизни, оцениваемого с помощью специальной анкеты).

Пример применения таких показателей в качестве критерия оценки (в данном случае переносимость физической нагрузки) приводится в исследовании эффективности эналаприла и гидралазина в сочетании с нитратами при сердечной недостаточности. В отличие от влияния на уровень смертности, комбинированная терапия по сравнению с применением эналаприла повышала переносимость физических нагрузок. Период наблюдения составил 6 мес; для анализа полученных результатов использовали χ^2 -критерий. В группе комбинированной терапии переносимость физических нагрузок увеличилась в большей степени, причем полученные данные не зависели от влияния случайных факторов ($p=0,02$).

ВНЕСЕНИЕ ПОПРАВОК НА ИСХОДНЫЕ РАЗЛИЧИЯ

Проводя проверку гипотезы, исследователи часто учитывают исходные различия между группами сравнения, внося *поправки* в ходе соответствующего анализа, т.е. рассчитывая *стандартизованные показатели*. В процессе рандомизации больные распределяются случайным образом; при этом, как правило, формируются сопоставимые группы. Однако в некоторых случаях группы могут существенно различаться по наличию и распределению прогностических факторов, влияющих на изучаемый клинический исход. Например, известно, что в некоторых случаях прогноз хуже у пожилых больных. Однако при сравнении эффективности 2 методов лечения доля таких больных в одной из групп может оказаться больше. Учитывая, что риск развития побочных эффектов у пожилых больных выше, отсутствие поправки на возраст снижает достоверность полученных результатов. Таким образом, при расчете величины p также будет необходимо внести соответствующие поправки. Это позволяет провести анализ данных таким образом, как если бы возрастные различия между группами отсутствовали. Необходимые поправки можно вносить для нескольких переменных одновременно. В данном случае интерпретация величины p не будет ничем отличаться от обычной.

ОДНОВРЕМЕННАЯ ПРОВЕРКА МНОЖЕСТВА ГИПОТЕЗ

Долгое время студенты университетов были объектами для проведения ^{Ния} Различных экспериментов. Следуя традиции, для нашего следующего гипотетического исследования мы выбрали студентов-медиков.

Представьте себе медицинский институт и двух преподавателей, читающих вводный курс медицинской статистики. Один из них пользуется большей популярностью среди учащихся. Декан факультета не может никем заменить менее популярного преподавателя. Тогда, стараясь действовать справедливо, он решает провести рандомизированное распределение 200 студентов 1-го курса по группам; в этом случае у каждого студента будут равные шансы (50%) попасть к тому или иному преподавателю.

Преподаватели использовали это решение для того, чтобы наглядно объяснить студентам некоторые важные принципы медицинской статистики. Они задали вопрос: «Имеются ли различия в характеристиках студентов двух групп, которые нельзя объяснить влиянием случайных факторов?» Из оцениваемых характеристик были выбраны пол, цвет глаз, рост, средний балл, полученный в последний год обучения в колледже, социально-экономический статус и любимая музыка. Для каждого сравнения была сформулирована своя нулевая гипотеза. Например, для анализа групп по половой принадлежности она звучала так: студентов отбирали с одного курса, поэтому число девушек в двух сформированных группах должно быть одинаковым. Поскольку студентов отбирали из одной популяции и распределяли по группам рандомизированно, при всех сравнениях нулевая гипотеза истинна, а во всех случаях, когда она отвергнута, результат будет недостоверным.

Преподаватели определили необходимые характеристики у каждого из студентов. Было установлено, что группы не различаются по распределению 5 из этих характеристик (во всех случаях при оценке существующих различий $p > 0,10$), однако в одной группе глаза были голубыми у 25 из 100, а в другой — у 38 из 100 студентов. Формальный статистический анализ показал, что в том случае, если нулевая гипотеза верна (а она верна), вероятность обнаружения аналогичных или более выраженных различий между группами по доле лиц с голубыми глазами чуть меньше 0,05. Используя традиционную точку разделения ($p < 0,05$), преподаватели вынуждены были отказаться от нулевой гипотезы.

Насколько вероятен тот факт, что при проверке 6 независимых гипотез в 2 группах студентов хотя бы в 1 случае различия окажутся статистически значимыми в силу влияния случайных факторов? Под *независимыми гипотезами* мы подразумеваем, что данные, полученные при проверке одной из них, не влияли на результаты проверки других. Если вероятность получения статистически значимых результатов составляет 0,05, то вероятность получения статистически незначимых результатов равна $1 - 0,05 = 0,95$. При одновременной проверке 2 гипотез вероятность получения статистически незначимых результатов равна $0,95 \times 0,95$

ли $0,95^2$, а при проверке 6 гипотез — $0,95^6$ (т.е. 74%). Таким образом, вероятность получения, по меньшей мере, одного статистически значимого результата при одновременной проверке 6 независимых гипотез составляет $1 - 0,95^6 = 26\%$; иными словами, он будет получен в 1 случае из 4, а не из 20. Если мы хотим сохранить общую границу статистически значимых и незначимых данных, равную 0,05, для каждого из 6 сравнений пороговую величину p следует разделить на 6; таким образом, она составит 0,008.

Приведенный пример имеет двоякий смысл. Во-первых, редкие исходы иногда оказываются случайными. Даже при анализе одной гипотезы результат, соответствующий $p=0,01$, будет получен в 1% наблюдений. Во-вторых, не следует проверять одновременно несколько гипотез, так как полученные при этом результаты могут оказаться ошибочными. В научной литературе имеется множество примеров данного феномена. Так, анализируя результаты 45 испытаний, отчеты о которых были опубликованы в 3 ведущих медицинских журналах, S.J. Rosock et al. обнаружили, что в каждом исследовании применялись в среднем 6 критериев оценки, причем при проведении большинства сравнений оценивалась статистическая значимость различий [2].

В качестве примера, убедительно иллюстрирующего опасность одновременного проведения нескольких сравнений, можно привести РКИ, в котором оценивалось влияние реабилитационных мероприятий на качество жизни больных после инфаркта миокарда. Участников рандомизированно включали в группы стандартной терапии, лечебной физкультуры и медицинских консультаций. Все больные сообщали информацию о своей работе, досуге, сексуальной активности, удовлетворенности своим состоянием, соблюдении рекомендаций врача, качестве досуга и работы, психологическом состоянии, симптомах сердечно-сосудистых заболеваний и общем состоянии здоровья [4]. Три группы были сравнимы практически по всем показателям, однако после 18 мес наблюдения удовлетворенность своим состоянием у больных из группы лечебной физкультуры была выше; необходимость в помощи членов семьи, посвящаемое работе время и частота половых актов в группе проведения консультаций была ниже, чем в двух других группах. Означает ли это, что программы лечебной физкультуры и консультаций необходимо внедрять из-за того положительного (хотя и незначительного) влияния, которые они оказывают на самочувствие больных, или, напротив, от них следует отказаться, так как по большей части клинические исходы между группами не различались? Сами исследователи пришли к выводу, что предложенные ими программы реабилитации не повышают качества жизни больных. Однако их оппоненты могут возразить, что,

если они улучшают хотя бы некоторые клинические исходы, их применение имеет смысл. Очевидно, что использование многочисленных критериев оценки способствует возникновению подобных споров.

Существует ряд статистических методов, позволяющих одновременно проверять несколько гипотез. Один из них был продемонстрирован выше: величину p разделили на число тестов. Можно также выбрать до начала исследования один основной критерий оценки, который и будет определять окончательные выводы. Кроме того, существуют статистические методы (например, обобщение величины эффекта), позволяющие объединять различные клинические исходы в один комбинированный критерий оценки. Более детальное обсуждение статистических методов, которые применяются в исследованиях с множественными критериями оценки, в данной книге не приводится. Однако те, кто заинтересовался данной проблемой, смогут найти необходимую информацию в других источниках [5].

ПРОБЛЕМЫ, С КОТОРЫМИ ПРИХОДИТСЯ СТАЛКИВАТЬСЯ ПРИ ПРОВЕРКЕ ГИПОТЕЗ

Несомненно, что у некоторых врачей возник ряд вопросов, которые остались непонятыми. Почему, например, нужно использовать единственную точку деления, когда ее выбор осуществляется произвольно? Почему на вопрос об эффективности лечения следует отвечать однозначно — да или нет, когда можно рассматривать этот критерий оценки как непрерывную переменную (от «эффективность маловероятна» до «эффективность практически не вызывает сомнения»)?

Врачи, у которых возникли подобные вопросы, стоят на правильном пути. В главе 3 *Лечение и понимание результатов исследования: доверительные интервалы* дается объяснение, почему альтернативные подходы к анализу данных имеют преимущества перед проверкой гипотезы.

Литература

1. *Coñn J.N., Johnson G., Ziesche S., et al.* A comparison of enalapril with hydralazine-isosorbide dinitrate in the treatment of chronic congestive heart failure. *N Engl J Med* 1991;325:303—310.
2. *Detsky A.S., Sackett D.L.* When was a "negative" trial big enough? How many patients you needed depends on what you found. *Arch Intern Med* 1985; **145:709—715**.
3. *Kirshner B.* Methodological standards for assessing therapeutic equivalence. *J Clin Epidemiol* 1991;44:839-849.
4. *Mayou R., MacMahon D., Sleight P., Florencio M.J.* Early rehabilitation after myocardial infarction. *Lancet* 1981 ;2:1399—1401.
5. *Pocock S.J., Geller N.L., Thiatis A.A.* The analysis of multiple endpoints in clinical trials. *Biometrics* 1987;43:487—498.

3**ЛЕЧЕНИЕ И ПОНИМАНИЕ
РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ
ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ**

Г. Гайятт, С. Уолтер, Д. Кук, Р. Йешке

В подготовке данной главы принимали также участие члены рабочей группы по разработке и внедрению принципов доказательной медицины М. Уилсон и М. Стоклер

В ЭТОМ РАЗДЕЛЕ

Как лечить больных с сердечной недостаточностью? Трудности при интерпретации результатов исследования

Решение проблемы: что такое доверительные интервалы?

Использование доверительных интервалов при интерпретации результатов клинических исследований

Интерпретация результатов исследований, в которых не было выявлено различий между основной и контрольной группами

Интерпретация результатов исследований, в которых были выявлены различия между основной и контрольной группами

Был ли размер выборки достаточным?

Вывод

При проверке гипотез определяют вероятность того, что наблюдаемые результаты случайны, если верна нулевая гипотеза, т.е. гипотеза, согласно которой чаще всего подразумевается отсутствие различий между основной и контрольной группами (см. главу 2 *Лечение и понимание результатов исследования: проверка гипотез*). Все больше авторов медицинских исследований и преподавателей медицинских дисциплин понимают ограничения проверки гипотез; теперь становится более популярным альтернативный подход к сравнению данных, полученных в группах сравнения, а именно расчет точечных значений и *доверительных интервалов* (ДИ). Термины и определения, которыми мы будем оперировать, рассматриваются рядом авторов [1–5]; их рассуждения могут дополнить наши. В ходе обсуждения будет использован пример, к которому мы уже прибегали ранее (см. главу 2 *Лечение и понимание результатов исследования: проверка гипотез*).

КАК ЛЕЧИТЬ БОЛЬНЫХ С СЕРДЕЧНОЙ НЕДОСТАТОЧНОСТЬЮ? ТРУДНОСТИ ПРИ ИНТЕРПРЕТАЦИИ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

Авторы двойного слепого рандомизированного контролируемого испытания (РКИ), в котором участвовали 804 мужчины с сердечной недостаточностью, сравнивали применение эналаприла и лечение гидралазином в сочетании с нитратами [6]. За период наблюдения, длительность которого составила от 6 мес до 5,7 года, умерли 132 из 403, или 33%, больных в группе эналаприла, и 153 из 401, или 38%, больных в группе комбинированного лечения. Величина p , характеризующая статистическую значимость различий по смертности между группами, составила 0,11.

Если согласиться с рекомендуемым значением вероятности того, что наблюдаемые различия обусловлены влиянием случайных факторов ($p > 0,05$; см. главу 2 *Лечение и понимание результатов исследования: проверка гипотез*), нулевую гипотезу отвергать нельзя, т.е. можно заключить, что между группами сравнения не было существенных различий. Исследователи провели анализ, в ходе которого помимо выживаемости сравнивали сроки наступления смерти в группах. В результате такого анализа дожития, который в целом более чувствителен, чем оценка различий по частоте развития клинического исхода (см. главу 4 *Лечение и понимание результатов исследования: измерение корреляции*), была рассчитана величина $r = 0,08$ (что указывает на отсутствие статистически значимых различий между группами). Это подтвердило результаты более

простого анализа. Вместе с тем авторы сообщают о том, что различие по уровню смертности между группами через 2 года (который рассматривался в качестве одного из основных критериев оценки) достигло уровня статистической значимости ($p=0,016$).

Подобные противоречия способны запутать любого. На этом этапе врач должен спросить себя, указывают ли результаты исследования на преимущества ингибиторов ангиотензинпревращающего фермента (АПФ) перед гидралазином в сочетании с нитратами либо на отсутствие различий между двумя вмешательствами.

РЕШЕНИЕ ПРОБЛЕМЫ: ЧТО ТАКОЕ ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ?

Как интерпретировать результаты исследования, учитывая ограничения методов проверки гипотез? Можно избрать альтернативный способ, не задаваясь вопросом, соответствуют ли результаты нулевой гипотезе и какова величина p , сформулировав вопросы следующим образом: 1) каково точечное значение размера эффекта, с наибольшей вероятностью отражающее истинное различие между сравниваемыми вмешательствами? и 2) каков приемлемый диапазон, в котором может находиться истинное различие? На второй вопрос можно ответить, определив диапазон ДИ. Прежде чем применять данный подход к решению проблемы о назначении эналаприла или гидралазина в сочетании с нитратами при сердечной недостаточности, мы проиллюстрируем использование ДИ на простом примере с подбрасыванием монеты.

Представим монету, о сбалансированности которой нам ничего неизвестно. Иными словами, действительная вероятность выпадения решки в каждом конкретном случае может составлять 0,5; 1,0 (т.е. во всех случаях должны выпадать решки) или 0 (всегда будут выпадать орлы). Теперь проведем эксперимент.

Сначала мы подбросим монету дважды, и один раз выпадает решка, а другой — орел. Каково самое достоверное предположение о вероятности выпадения решки при любом подбрасывании можно сделать на основании полученных результатов? Соответствует ли этому предположению значение 0,5, которое мы получили (точечное значение)? Каковы правдоподобные рамки, в которых может находиться истинная вероятность выпадения решки? Они довольно широки, и большинство согласится, что этот интервал простирается от $\geq 0,1$ до $\leq 0,9$. Другими словами, после двух испытаний, если один раз выпала решка, а другой — орел,

еще нельзя исключить того, что в действительности вероятность выпадения решки при любом подбрасывании составляет 0,9. Так что после двух испытаний мы пока не приблизились к истине.

Подбросим монету еще восемь раз. Теперь, если суммировать все наши испытания, получается, что пять раз выпала решка, а пять — орел. Наиболее достоверное предположение об истинной вероятности выпадения решки по-прежнему соответствует 0,5 (точечная оценка). Вместе с тем рамки, в которых может находиться истинное значение вероятности, сузились. Уже нельзя предполагать, что она равняется 0,9. Иначе говоря, если истинная вероятность выпадения решки была бы равна 0,9, шансы того, что из 10 случаев решка выпадет только в 5, были бы крайне низкими. Интуитивное ощущение границ, в которых может располагаться истинная вероятность того или иного события, различается у разных людей, однако большинство согласится, что она вряд ли больше 0,8 или меньше 0,2.

После 10 подбрасываний предположения об истинной вероятности, лежащие в диапазоне между 0,2 и 0,8, не равноценны. Самое правдоподобное соответствует точечному значению (0,5), однако близкие к нему значения, например 0,4 или 0,6, также весьма правдоподобны. Чем дальше значение от полученного при точечной оценке, тем менее вероятно, что оно соответствует истинному.

Десять подбрасываний монеты оставили у нас еще некоторые сомнения о свойствах монеты, и мы бросаем ее еще 40 раз. После 50 попыток мы наблюдаем 25 случаев выпадения решки и 25 — орла, и точечная оценка остается равной 0,5. Теперь мы начинаем верить, что центр тяжести монеты не смещен; и диапазон, в котором может располагаться истинная вероятность выпадения решки, теперь сузится еще больше и составит от 0,35 до 0,65. Этот диапазон пока еще достаточно широк, и мы подбрасываем монету еще 50 раз. Если после 100 подбрасываний мы наблюдаем 50 случаев выпадения решки, действительная вероятность этого события вряд ли находится вне диапазона значений от 0,40 до 0,60. Если бы мы подбросили монету 1000 раз, и при этом в 500 случаях наблюдали бы выпадение решки, мы были бы почти уверены, что несбалансированность монеты, даже если таковая действительно присутствует, минимальна.

В данном эксперименте мы прибегли к здравому смыслу, чтобы опередить ДИ значений вероятности вокруг полученного результата (0,5). В любом случае ДИ отражает диапазон, в котором скорее всего лежит истинное значение. Чем меньше размер выборки, тем шире ДИ. Когда выборка увеличивается, сильнее наша уверенность в том, что истинное

значение оцениваемого показателя недалеко от точечной оценки, рассчитанной на основе наблюдений; поэтому ДИ становится меньше.

Представления о здравом смысле у разных людей значительно различаются. К счастью, для оценки ДИ можно использовать статистические методы. Для этого сначала следует определить, каким крайним значениям мы еще можем доверять. В нашем примере соответствующий вопрос будет сформулирован следующим образом: "Каков интервал, внутри которого в 95% случаев находится истинное значение?" В таблице представлены значения 95% ДИ для точечного значения, равного 0,5. Если нет необходимости в такой точности, можно довольствоваться более узким 90% ДИ; в этом случае точечное значение показателя будет находиться в данном диапазоне с 90% вероятностью (см. таблицу).

Из примера с подбрасыванием монеты также видно, каким образом с помощью ДИ можно ответить на вопрос о достаточности размеров выборки исследования для решения его задач. Если вы должны быть уверены в том, что ошибка в подсчетах не превышает 10% (т.е. точечное значение отстоит от верхней или нижней границы ДИ не более чем на 10%), необходимы 100 подбрасываний. Если нужна более точная оценка (точечное значение отстоит от верхней или нижней границы ДИ не более чем на 3%), монету нужно подбросить 1000 раз. Для повышения точности результата требуется только увеличить число наблюдений. В клинических исследованиях это подразумевает включение большего числа больных или более частое измерение изучаемых параметров у каждого участника.

ИСПОЛЬЗОВАНИЕ ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ ПРИ ИНТЕРПРЕТАЦИИ РЕЗУЛЬТАТОВ КЛИНИЧЕСКИХ ИССЛЕДОВАНИЙ

Каким образом знание ДИ помогает нам интерпретировать результаты испытания вазодилататоров при сердечной недостаточности?

Таблица. Разные доверительные интервалы показателя вероятности, равного 0.5, в эксперименте с подбрасыванием монеты

Число подбрасываний	Наблюдаемый результат	95% ДИ	90% ДИ
2	1 решка, 1 орел	От 0,01 до 0,99	От 0,03 до 0,98
10	5 решек, 5 орлов	От 0,19 до 0,81	От 0,22 до 0,78
50	25 решек, 25 орлов	От 0,36 до 0,65	От 0,38 до 0,62
100	50 решек, 50 орлов	От 0,40 до 0,60	От 0,41 до 0,59
1000	500 решек, 500 орлов	От 0,47 до 0,53	От 0,47 до 0,53

Смертность в группе ингибиторов АПФ составила 33%, в группе гидралазина в сочетании с нитратами — 38%, т.е. абсолютное различие по этому показателю равно 5%. Такое различие по смертности представляет собой точечное значение, наше самое правдоподобное предположение о влиянии на смертность ингибиторов АПФ по сравнению с контрольным вмешательством. При этом 95% ДИ значений этого показателя составляет от -1,2 до 12%.

Как теперь можно интерпретировать результат исследования? Наиболее вероятное различие по смертности равно 5%, хотя истинное значение может оказаться в непосредственной близости от -1,2% (т.е. комбинированная терапия гидралазином и нитратами уменьшает смертность на 1,2% по сравнению с применением ингибиторов АПФ) или от 12% (ингибиторы АПФ более эффективны, чем комбинированная терапия). Чем дальше предполагаемая величина от точечного значения (5%), тем менее она вероятна. Можно заключить, что продолжительность жизни у больных, принимающих ингибиторы АПФ, скорее всего выше, чем на фоне приема гидралазина в сочетании с нитратами; но эти различия могут быть несущественными или же, напротив, весьма значительными. Этот способ рассмотрения результатов позволяет не отвечать на вопросы о предпочтительности того или иного метода лечения однозначно ("да" или "нет", как при оценке гипотез) и целесообразности выбора авторами исследования смертности за 2 года в качестве критерия оценки. Теряют также свою актуальность споры о том, подтверждает ли исследование преимущество одного из сравниваемых методов лечения или их равноценность. Можно сделать вывод о том, что при прочих равных условиях ингибитор АПФ более эффективен при сердечной недостаточности, чем гидралазин в сочетании с нитратами, однако доказательства того малоубедительны. На окончательное решение повлияют данные о токсичности и стоимости препарата, а также результаты других исследований (см. главу 11 раздела *От доказательств к практике*). Поскольку в нескольких крупных РКИ было показано снижение смертности на фоне приема ингибиторов АПФ при сердечной недостаточности [7], можно с уверенностью рекомендовать препараты этого фармакологического класса как средство выбора у таких больных.

ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ, В КОТОРЫХ НЕ БЫЛО ^ЯВЛЕНО РАЗЛИЧИЙ МЕЖДУ ОСНОВНОЙ И КОНТРОЛЬНОЙ ГРУППАМИ

В качестве другого примера использования ДИ можно интерпретировать результаты шведского РКИ Swedish Co-operative Stroke Study, в

котором изучалось влияние терапии аспирином на частоту развития повторных инсультов после перенесенного ишемического инсульта [8, 9]. Несмертельный инсульт развился у 18 (7%) участников из группы плацебо (252 больных) и 23 (9%) участников из группы аспирина (253 больных). Точечная оценка в данном случае указывает на повышение частоты развития инсультов в группе аспирина на 2%.

Может показаться, что в этом исследовании, включающем более 500 участников, была выявлена неэффективность применения аспирина для профилактики повторного инсульта. Однако 95% ДИ для точечного значения, равного —2%, составляет от —7 до 3%. Если бы действительно аспирин снижал частоту развития инсультов на 3% по сравнению с плацебо (т.е. 3% больных, у которых развивается инсульт, могли бы его избежать, если бы принимали аспирин), то многие больные стали бы принимать этот препарат. СОР развития инсульта составило бы 43%, а ЧБНЛ — всего 33 (т.е. для предупреждения одного случая инсульта аспирин необходимо назначать 33 больным). Можно прийти к заключению, что результаты данного исследования не исключают наличия клинически значимого положительного эффекта аспирина; в этом смысле оно не было достаточно крупным.

Этот пример показывает, что выборка должна быть большой, если требуется точно оценить эффект вмешательства. Теперь ясно, почему мы рекомендуем врачам по возможности обращаться к систематическим обзорам, в которых обобщаются результаты исследований наивысшего методологического качества [10]. Применительно к нашему примеру такой обзор показывает, что назначение антиагрегантов больным с преходящими нарушениями мозгового кровообращения или перенесенным инсультом снижает ОР развития повторного инсульта примерно на 25% (при 95% ДИ примерно от 19 до 31%) [11]. Ознакомившись с этими данными, многие больные, у которых без лечения вероятность развития инсульта превышает 10% ($\text{ЧБНЛ} \leq 50$) или всего 5% ($\text{ЧБНЛ} \leq 100$), предпочитают принимать аспирин.

Из этого примера также видно, что при оценке результатов исследований, в которых не было выявлено различий между основной и контрольной группами (т.е. не удалось отвергнуть нулевую гипотезу), следует обратить внимание на верхнюю границу ДИ, которая характеризует наибольшую возможную эффективность изучаемого вмешательства. Если верхняя граница ДИ исключает какой-либо существенный эффект лечения, можно смело утверждать, что польза сравниваемых методов одинакова. Если, напротив, ДИ включает значения, соответствующие клинически значимому эффекту изучаемого вмешательства, нельзя отрицать его преимущества.

Подобный подход к интерпретации результатов исследований, в которых не было выявлено различий между основной и контрольной группами

пами, крайне полезен, когда надо определить, нужно ли заменять существующее вмешательство новым, которое дешевле, проще или менее токсично. Результаты *исследований эквивалентности вмешательств* свидетельствуют о целесообразности такой замены, только если обычное лечение не обладает существенными дополнительными преимуществами по сравнению с менее дорогостоящим или более простым. Мы будем уверены, что исключили вероятность наличия существенной дополнительной пользы стандартного лечения, если верхняя граница ДИ вокруг точечного значения размера эффекта ниже порога клинической значимости.

ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ, В КОТОРЫХ БЫЛИ ВЫЯВЛЕНЫ РАЗЛИЧИЯ МЕЖДУ ОСНОВНОЙ И КОНТРОЛЬНОЙ ГРУППАМИ

Какую информацию можно получить, оценивая ДИ при интерпретации результатов исследований, в которых были выявлены различия между основной и контрольной группами (т.е. удалось отвергнуть нулевую гипотезу)? В другом двойном слепом РКИ с участием больных с сердечной недостаточностью сравнивали эффективность эналаприла и плацебо [12]. Из 1285 больных в группе эналаприла умерли или были госпитализированы в связи с сердечной недостаточностью 613 (48%), а из 1284 больных в группе плацебо — 736 (57%). Точечное значение различия по комбинированному показателю смертности и частоты госпитализаций в связи с сердечной недостаточностью составляет 10% при 95% ДИ от 6 до 14%. Наименьший возможный эффект эналаприла в данном случае заключается в уменьшении частоты развития неблагоприятных исходов на 6%, что соответствует ЧБНЛ, равному 17. Если вы считаете такой размер эффекта клинически значимым, выборка достаточна, в противном случае необходимо проведение более крупного исследования (с более узкими ДИ значений размера эффекта).

БЫЛ ЛИ РАЗМЕР ВЫБОРКИ ДОСТАТОЧНЫМ?

До сих пор мы подводили читателя к мысли о том, что оценка ДИ позволяет ответить на вопрос: "Был ли размер выборки достаточным?" Такой подход можно проиллюстрировать с помощью рисунка; на нем изображено распределение результатов 4 типов РКИ, в которых при сравнении двух вмешательств смертность либо снижалась на 5%, либо повышалась на 1%. Вертикальная линия в центре отражает САР, равное 0 (т.е. одинаковую смертность в экспериментальной и контрольной группах). Значения справа от вертикальной линии отражают результаты, соответст-

вующие более высокой смертности в контрольной группе, значения слева от вертикальной линии — более высокой смертности в группе лечения.

На рисунке представлены кривые распределения результатов разных РКИ: крупных и относительно небольших. Для каждой из четырех кривых наивысшая точка отражает истинное изменение смертности. Кривые А и В отображают результаты РКИ, в которых смертность снижалась на 5%, а кривые С и D — результаты РКИ, в которых смертность повышалась на 1%.

Предположим, что мы считаем снижение абсолютного риска смерти на 1% и более клинически значимым результатом. Иными словами, при таких значениях $САР$ преимущества лечения будут перевешивать связанные с ним риск и затраты (см. главу 11 раздела I *От доказательств к практике*). При $САР < 1\%$ лечение лишено смысла (его недостатки превышают преимущества). Например, если изучаемое вмешательство приводит к снижению смертности с 5 до 4% и менее, его можно использовать на практике. Напротив, если смертность снизилась с 5 до 4,5%, от применения этого метода стоит воздержаться. Как можно интерпретировать результаты 4 типов РКИ, учитывая вышеизложенное?

На кривой А более 95% значений $САР$ превышают 1% (все кривые представлены схематично; в действительности вероятность никогда не бывает равна нулю). Рассчитанные на основании результатов РКИ этого метода лечения с таким же размером выборки 95% ДИ в большинстве случаев не включают $САР$, равное 1%. Оценивая результаты данного РКИ, можно быть уверенным, что истинный размер эффекта вмешательства

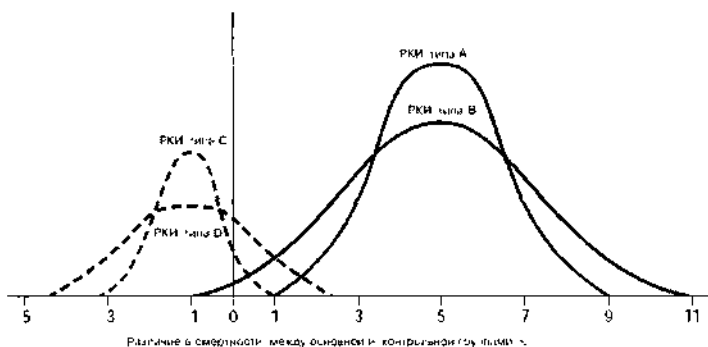


Рисунок. Решение о точности результатов РКИ: распределение результатов четырех типов РКИ, в которых оценивалась эффективность двух вмешательств. А — результаты крупных РКИ, в которых при оценке смертности $САР=5\%$; В — результаты менее крупных РКИ, в которых $САР=5\%$; С — результаты крупных РКИ, в которых $ПАР=1\%$; D — результаты менее крупных РКИ, в которых $ПАР=1\%$. Приводится с разрешения Канадской медицинской ассоциации.

выше порога клинической значимости ($САР=1\%$), т.е. речь идет об исследовании с несомненно положительным результатом. Скорее всего истинное значение $САР$ существенно превышает 1% , поэтому многие больные будут заинтересованы в получении такого лечения. Размер выборки в подобных исследованиях достаточен для того, чтобы сделать вывод о высокой эффективности метода.

Кривая В также отображает результаты РКИ, в которых изучаемое вмешательство снижало смертность на 5% , однако в этих РКИ участвовало меньше больных. Поскольку в некоторых из этих исследований будет отвергнута нулевая гипотеза (об отсутствии различий между группами), 95% ДИ в большинстве случаев будут включать $САР$, равное 1% или менее. При интерпретации результатов таких РКИ остаются сомнения в том, что размер эффекта действительно превышает установленный заранее порог клинической значимости. Подобные РКИ также подтверждают преимущества изучаемого вмешательства, но их результаты менее достоверны, поскольку выборка недостаточна для того, чтобы рекомендовать его практическое применение.

Кривая С отображает результаты РКИ, в которых не удалось исключить нулевую гипотезу, т.е. эффект изучаемого лечения отсутствовал. В среднем смертность в основной группе была на 1% выше, чем в контрольной ($ПАР=1\%$). В большинстве из этих РКИ диапазон 95% ДИ был узким, причем все его значения располагались слева от $+1\%$ (порога клинической значимости). Поскольку верхняя граница ДИ меньше 1% , можно быть уверенным в том, что если эффект лечения присутствует, он очень мал, и вряд ли перевешивает опасности, затраты и неудобства, связанные с вмешательством. Поэтому можно однозначно исключить любые клинически значимые преимущества лечения и отказаться от его проведения — по крайней мере у больных из этой популяции.

Кривая D отображает результаты РКИ того же вмешательства, что и на кривой С (т.е. смертность в основной группе на 1% больше, чем в контрольной), однако в связи с меньшими размерами выборки диапазон ДИ шире и в большинстве РКИ включает в себя значения, превышающие порог клинической значимости (1%). Не исключено (хотя и маловероятно), что истинный эффект изучаемого лечебного вмешательства может соответствовать значениям $САР$, превышающим 1% . Несмотря на то что мы воздерживаемся от использования данного метода (скорее всего он повышает, ^a не снижает смертность), у нас нет полной уверенности в его неэффективности. Большинство РКИ, результаты которых представлены на кривой D, не дают однозначного ответа; для исключения клинически значимых различий между группами необходимы более крупные РКИ.

Вывод

Если в исследовании были выявлены различия между группами сравнения, следует обратить внимание на нижнюю границу ДИ, чтобы определить, был ли достаточным размер выборки. Если эта нижняя граница, которая соответствует наименьшему возможному эффекту лечения при правдоподобности полученных результатов, выше порога клинической значимости, размер выборки достаточен; поэтому проведения дополнительных РКИ не требуется. Если же нижняя граница ДИ не достигает этого порога, размер выборки недостаточен, и необходимо проведение новых, более крупных РКИ.

В случае с исследованием, в котором не было показано различий между группами сравнения, обратите внимание на верхнюю границу ДИ, чтобы определить, достаточное ли количество больных участвовало в исследовании. Если эта верхняя граница, которая соответствует наибольшему возможному эффекту лечения при правдоподобности полученных результатов, не достигает порога клинической значимости, размер выборки достаточен; поэтому проведения дополнительных РКИ не требуется. Если же верхняя граница ДИ выше этого порога, размер выборки недостаточен, и необходимо проведение новых, более крупных РКИ.

Литература

- 1 *Simon R* Confidence intervals for reporting results of clinical trials *Ann Intern Med* 1986,105 429—435
- 2 *Gardner M J, Altman D G, eds* Statistics With Confidence Confidence Intervals and Statistical Guidelines London BMJ Publishing Group, 1989
- 3 *Bulpitt C J* Confidence intervals *Lancet* 1987,1 494-497
- 4 *Pocock S J, Hughes M D* Estimation issues in clinical trials and overviews *Stat Med* 1990,9 657—671
- 5 *Braitman L E* Confidence intervals assess both clinical significance and statistical significance *Ann Intern Med* 1991,114 515-517
- 6 *Colin J N, Johnson G, Ziesche S, et al* A comparison of enalapril with hydralazine-isosorbide dinitrate in the treatment of chronic congestive heart failure *N Engl J Med* 1991 325 303—310
- 7 *Garg R, Yusuf S* Overview of randomized trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure Collaborative Group on ACE Inhibitor Trials *JAMA* 1995 273 1450-1456
- 8 *Button M, Hjemers C, Samuelsson K* High-dose salicylic acid after cerebral infarction a Swedish cooperative study *Stroke* 1997,18 325
- 9 *Sackett D L, Haynes R B, Guyatt G H, Tugwell P* Clinical Epidemiology A Basic Science for Clinical Medicine Boston Little, Brown and Company, 1991 218—220
- 10 *Oxman A D, Guyatt G H* Guidelines for reading literature reviews *CMAJ* 1988,138 697—703
- 11 Antiplatelet Trialists' Collaboration Secondary prevention of vascular disease by prolonged antiplatelet treatment *BMJ* 1988,296 320-331
- 12 *The SOLVD Investigators* Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure *N Engl J Med* 1991,325 293—302

4

ЛЕЧЕНИЕ И ПОНИМАНИЕ
РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ
ИЗМЕРЕНИЕ КОРРЕЛЯЦИИ

Р. Йешке, Г. Гайятт, А. Барратт, С. Уолтер,
Д. Кук, Ф. МакАлистер, Дж. Эттиа

В подготовке данной главы принимала также участие
Ш Строе

В ЭТОМ РАЗДЕЛЕ

Дихотомические и непрерывные данные

Таблица **сопряженности**

Абсолютный риск

Снижение абсолютного **риска**

Относительный риск

Снижение относительного **риска**

Отношение шансов

Что лучше отражает степень снижения абсолютного риска: относительный риск или отношение шансов?

Число больных, которых необходимо лечить

Индекс потенциального вреда

Еще раз о таблице сопряженности

Доверительные интервалы

Анализ выживаемости

Исследования случай—контроль

Какой метод измерения корреляции **наилучший?**

Врачи изучают результаты клинических испытаний, чтобы оценить связь между лечением и клиническим исходом. В ходе исследования такая связь может быть подтверждена или отвергнута; например, полученные результаты могут свидетельствовать о том, что изучаемое вмешательство уменьшает риск развития осложнений или, напротив, не влияет на него.

В настоящей главе основное внимание будет уделено дихотомическим клиническим исходам, таким как смерть, развитие инсульта или инфаркта миокарда. Результат исследований, посвященных влиянию вмешательства на такие клинические исходы, обычно представлен в виде доли участников, у которых был выявлен тот или иной клинический исход, в каждой из групп. Рассмотрим результаты применения трех методов лечения в трех популяциях с различным уровнем смертности (рис. 1). В первой популяции смертность снизилась с 30 до 20%, во второй — с 10 до 6,7%, в третьей — с 1 до 0,67%.

Все три метода лечения уменьшают смертность примерно на $\frac{1}{3}$, однако этот показатель, взятый отдельно, не позволяет достаточно полно определить их эффективность. Мы получим дополнительную информацию о корреляции между вмешательством и клиническим исходом, измерив относительный риск (ОР), снижение относительного риска (СОР), снижение абсолютного риска (САР), отношение шансов (ОШ), число больных, которых необходимо лечить определенным методом в течение определенного времени, чтобы достичь определенного благоприятного

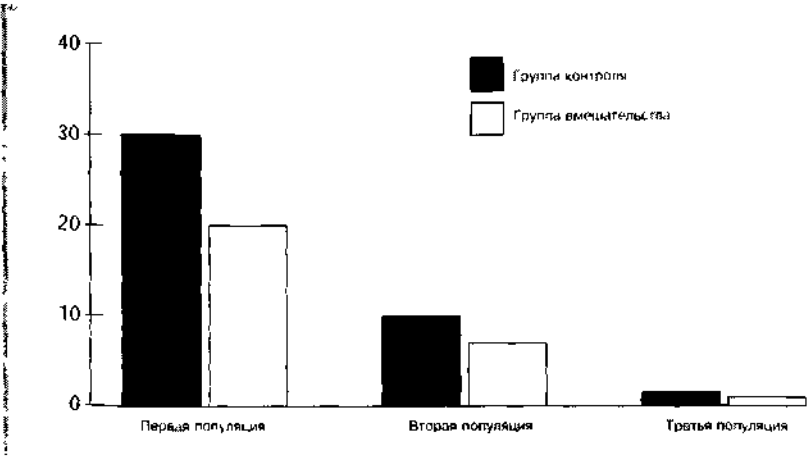


Рис. 1. Одинаковое снижение относительного риска при разном снижении абсолютного риска смерти.

—

эффекта или предотвратить определенный неблагоприятный исход у одного больного (ЧБНЛ), или индекс потенциального вреда (ИПВ).

ДИХОТОМИЧЕСКИЕ И НЕПРЕРЫВНЫЕ ДАННЫЕ

Основой для первичного анализа обычно служат данные о доле участников в группах контроля и вмешательства, у которых выявлен тот или иной клинический исход. В качестве благоприятных или неблагоприятных клинических исходов могут быть выбраны наличие или отсутствие инсульта, инфаркта миокарда, метастазов злокачественных опухолей, смерть или выживание, заживление язвы или исчезновение симптомов. Более того, даже при рассмотрении непрерывных данных авторы нередко используют дихотомические клинические исходы. Например, информацию о переносимости физической нагрузки без возникновения болей в груди, числе приступов стенокардии в месяц, изменении функции легких или частоте обращений в отделение неотложной помощи обычно представляют в виде средних величин в группах сравнения. Эти же данные можно преобразовать в дихотомические, если определить пороговые значения или степень изменения, отражающую существенное улучшение или ухудшение состояния, с последующим подсчетом доли участников, у которых был достигнут означенный порог или отмечались соответствующие изменения. Так, в исследовании, посвященном оценке эффективности приема кортикостероидных препаратов при хронической стабильной обструкции дыхательных путей, в качестве критерия оценки использовали такой показатель, как объем форсированного выдоха за 1-ю секунду (ОФВ¹). Увеличение исходного ОФВ¹ более чем на 20% считали благоприятным клиническим исходом [1]. В другом исследовании оценивали различия в доле больных с хроническими заболеваниями легких, у которых существенно улучшалось качество жизни, связанное с состоянием здоровья [2]. Результат подобных исследований будет зависеть от того, какие изменения изучаемого показателя авторы сочтут существенными (ОШ, о котором пойдет речь ниже, по сравнению с другими показателями в меньшей степени зависит от упомянутого субъективного фактора).

ТАБЛИЦА СОПРЯЖЕННОСТИ

Табл. 1 представляет собой пример четырехпольной таблицы сопряженности, содержащей данные о дихотомическом клиническом ис-

ходе, полученные в клиническом испытании. Предположим, что в рандомизированном контролируемом испытании (РКИ) оценивали смертность при кровотечении из варикозно-расширенных вен пищевода после их эндоскопической перевязки по сравнению с эндоскопической склеротерапией [3]. Средняя продолжительность наблюдения составила 10 мес; к концу испытания в группах перевязки сосудов и склеротерапии умерли 18 из 64 и 29 из 65 больных соответственно (табл. 2).

Таблица 1. Пример четырехпольной таблицы сопряженности

Вмешательство	Клинический исход	
	Да	Нет
Да	a	b
Нет	c	d

Примечание

$$OP = \frac{a/(a+b)}{c/(c+d)}$$

$$COP = \frac{c/(c+d) - a/(a+b)}{c/(c+d)}$$

$$CAP = \frac{c}{c+d} - \frac{a}{a+b}$$

$$ЧБНЛ = \frac{1}{ARP}$$

$$ОШ = \frac{a/b}{c/d} = \frac{ad}{cb}$$

Таблица 2. Сравнительная эффективность эндоскопической склеротерапии и эндоскопической перевязки сосудов пищевода при кровотечении из варикозно-расширенных вен пищевода (приводится по [3])

Вмешательство	Клинический исход		
	смерть	выживание	всего
Перевязка сосудов пищевода	18	46	64
Склеротерапия	29	36	65

Примечание. OP=0,63, COP=0,37, CAP=0,165, ЧБНЛ=6, ОШ=0,49

АБСОЛЮТНЫЙ РИСК

Определение абсолютного риска (АР) служит наиболее простым способом оценки связи между вмешательством и исходом. АР смерти в группе перевязки сосудов пищевода составляет 28% (18 из 64 больных, или $a/a+b$), а в группе склеротерапии — 45% (29 из 65 больных, или $c/c+d$). Риск развития неблагоприятного исхода в контрольной группе обычно называют *исходным риском*, или *частотой развития исхода в контрольной группе*.

СНИЖЕНИЕ АБСОЛЮТНОГО РИСКА

Величины АР можно соотнести путем определения разницы между ними. Этот показатель получил название САР; он рассчитывается по формуле: $[a/(a+c)] - [b/(b+d)]$ (см. табл. 1) и позволяет оценить, насколько реже тот или иной неблагоприятный исход наступает в основной группе по сравнению с контрольной. В нашем примере $САР=0,446-0,281=0,165$ или 16,5%.

ОТНОСИТЕЛЬНЫЙ РИСК

Другой способ оценки связи между величинами АР заключается в определении отношения между ними, т.е. ОР, или отношения риска. Этот показатель позволяет судить, во сколько раз изменилась вероятность клинического исхода в группе изучаемого вмешательства (в нашем примере эндоскопической перевязки сосудов пищевода) по сравнению с контрольной группой (склеротерапии). ОР рассчитывают по формуле: $[a/(a+c)]/[b/(b+d)]$ (см. табл. 1). В рассмотренном примере в группе перевязки умерли 18 из 64 больных, а в группе склеротерапии — 29 из 65, т.е. $ОР=63\%$. Другими словами, АР смерти при перевязке сосудов пищевода примерно на У, ниже, чем при склеротерапии.

СНИЖЕНИЕ ОТНОСИТЕЛЬНОГО РИСКА

Оценить эффективность вмешательства можно и с помощью СОР, которое отражает уменьшение исходного риска в группе изучаемого вмешательства и рассчитывается путем деления САР на АР в контрольной группе (см. табл. 1). В нашем примере для определения СОР необ-

ходимо 16,5% (САР) разделить на 44,6% (АР смерти в группе склеротерапии); таким образом, СОР составляет 0,37. СОР можно подсчитать и другим способом: $1,0 - \text{ОР}$; в рассмотренном случае $1,0 - 0,63 = 0,37$ или 37% - Если не использовать математических терминов, эндоскопическая перевязка сосудов пищевода по сравнению со склеротерапией сопровождается уменьшением риска смерти на 37%.

ОТНОШЕНИЕ ШАНСОВ

Вместо оценки риска возникновения того или иного клинического исхода мы можем определить отношение шанса его развития к шансу того, что он не возникнет. Пожалуй, чаще всего понятием шансов оперируют при прогнозировании исхода спортивных соревнований, таких как скачки, бокс и теннис. В медицинских исследованиях ОШ определяется как отношение доли участников исследования, у которых выявлен тот или иной клинический исход, к доле больных, у которых данный исход не развился. В большинстве случаев шансы и ОР приблизительно равны, поэтому многие авторы, фактически рассчитывая ОШ, представляют полученные результаты в виде ОР. Иногда такое толкование может приводить к ошибочным выводам, поэтому ниже мы подробнее остановимся на том, что же представляет собой данный показатель на самом деле.

Рассмотрим следующий пример. Предположим, что у 20% больных, включенных в испытание, развился инсульт; следовательно, шанс составляет $(1/5)/(4/5)$ или $0,20/0,80 = 0,25$. Поскольку знаменатель в верхней и нижней части выражения одинаков, его можно отбросить, после чего останется доля больных, у которых развился инсульт, деленная на долю участников, у которых инсульт не возник. Для пересчета в ОР следует шансы поделить на сумму $(1 + \text{шансы})$. Например, если шанс возникновения неблагоприятного клинического исхода в послеоперационном периоде равняется 0,5, ОР составит $0,5/1 + 0,5 = 0,33$. В табл. 3 продемонстрирована связь между некоторыми значениями ОР и шансов. Обратите внимание, что разница между значениями ОР и шансов прямо пропорциональна ОР.

В нашем примере для расчета шансов смерти в группе перевязки сосудов пищевода следует разделить 18 (число умерших) на 46 (число выживших), а в группе склеротерапии — 29 на 46. ОШ (см. табл. 2) определяется по формуле $(a / c) / (b / d) = (18 / 46) / (29 / 36) = 0,49$. Подобно тому, как отношение показателей риска получило название ОР, найден-

Таблица 3. Соотношение между ОР и ОШ при разных значениях ОР

ОР, %	ОШ
80	4
60	1,5
50	1,0
40	0,67
33	0,50
25	0,33
20	0,25
10	0,11
5	0,053

Примечание $ОР = ОШ / (1 + ОШ)$, $ОШ = ОР / (1 - ОР)$

ное значение можно было бы назвать относительными шансами, однако эпидемиологи, не склонные к аналогиям, предпочли другой термин, а именно ОШ.

Врачи интуитивно чувствуют, что такое ОР и ОШ. Игроки прекрасно понимают смысл шансов. Но никто, кроме, разве что, некоторых статистиков, ясно не представляет себе, что такое ОШ [4, 5]. Несмотря на это, до недавнего времени ОШ использовали в качестве основного критерия оценки связи между явлениями [6], поскольку в отличие от ОР данный показатель лишен недостатков, связанных с произвольным выбором того или иного клинического исхода (например, смерти или, напротив, выживания) для сравнения риска его развития в группах [7].

Любой врач, наверное, предпочел бы заменить несколько расплывчатое ОШ интуитивно понятным показателем ОР. Вернувшись к нашей таблице сопряженности (см. табл. 1), мы увидим, что для такой замены необходимо, чтобы значение $ОР = [a / (a + B)] / [c / (c + d)]$ было близко к значению $ОШ = (a / B) / (c / d)$. Это условие будет соблюдено, если a будет значительно меньше, чем B , а c значительно меньше, чем d ; другими словами, если частота изучаемого исхода в обеих группах будет небольшой. В табл. 3 показано, что по мере уменьшения ОР его величина приближается к значению шансов. В большинстве РКИ частота развития изучаемого клинического исхода достаточно мала, поэтому значения ОР и ОШ практически совпадают. Аналогично показатели ОР и ОШ окажутся почти одинаковыми, если размер эффекта лечения невелик (т.е. ОР и ОШ приблизительно равны 1), но при выраженном эффекте эти показатели различаются.

Когда частота развития клинического исхода высокая, а эффект вмешательства выраженный, можно воспользоваться специальными формулами для преобразования ОШ в ОР [8, 9]. К счастью, необходимость в таком пересчете возникает нечасто. Чтобы понять причины этого феномена, рассмотрим мета-анализ, посвященный сравнению эффективности перевязки сосудов и склеротерапии при варикозном расширении вен пищевода [10]. Результаты мета-анализа свидетельствуют о том, что риск возникновения рецидива кровотечения в группе склеротерапии составляет 47% (как и в большинстве испытаний). ОШ рецидива кровотечения при перевязке сосудов равен 0,52, что подтверждает наличие достаточно выраженного эффекта лечения. Несмотря на довольно высокую частоту изучаемого клинического исхода и выраженный эффект вмешательства, ОР в данном примере составляет 0,60; таким образом, его отличие от показателя ОШ настолько мало — и это очень важно с практической точки зрения — что выбор того или иного способа представления результатов не должен повлиять на принятие клинического решения.

ЧТО ЛУЧШЕ ОТРАЖАЕТ СТЕПЕНЬ СНИЖЕНИЯ АБСОЛЮТНОГО РИСКА: ОТНОСИТЕЛЬНЫЙ РИСК ИЛИ ОТНОШЕНИЕ ШАНСОВ?

Итак, мы установили, что различия между ОШ и ОР очень редко имеют практическое значение. Чтобы лучше понять взаимосвязь между ОР, ОШ и САР, несколько видоизменим нашу таблицу сопряженности (см. табл. 2). Предположим, что смертность в обеих группах уменьшилась приблизительно на 50%, т.е. в группе перевязки сосудов умерли 9 из 64 больных, а в группе склеротерапии — 14 из 65. АР смерти при этом снизился с 28 до 14% и с 44,6 до 23,3% соответственно. ОР, как и раньше, составляет $14 / 22,3 = 0,63$. ОШ равняется $(9 / 55) / (14 / 51)$ или 0,60, что несколько отличается от прежнего значения (0,49) и приближается к показателю ОР. Значение же САР изменилось кардинально — с 16,5 до 8%. Таким образом, уменьшение смертности в обеих группах в 2 раза не повлияло на ОР, привело к небольшому увеличению ОШ и уменьшению САР в 2 раза. Данный пример (см. табл. 1) показывает, что ОР может быть одинаковым при различных показателях САР; и в то время как ОР не дает никакой информации о риске развития неблагоприятного исхода в контрольной группе (в нашем примере в группе менее эффективного лечения), САР в значительной степени зависит от исходного АР.

Таким образом, $OR=0,67$ может свидетельствовать о снижении смертности на фоне терапии как с 1 до 0,67%, так и с 30 до 20% (см. табл. 1). Теперь допустим, что частота развития тяжелых побочных эффектов, связанных с лечением, составляет 10% (например, на фоне химиотерапии при злокачественных опухолях). При такой частоте осложнений мы едва ли станем рекомендовать лечение всем без исключения больным, если обусловленное им снижение смертности равно 0,33% (с 1 до 0,67%). В то же время, если лечение приводит к уменьшению смертности с 30 до 20%, его проведение можно признать целесообразным, поскольку в такой ситуации положительный эффект будет отмечен у 10 больных из 100, а осложнения — у 1 из 100 (весьма неплохое соотношение риска и пользы вмешательства).

Показатель $СОР$ аналогичен $ОР$ и не отражает изменение риска по сравнению с исходным уровнем, т.е. в контрольной группе. В нашем примере $СОР$ не изменится, даже если частота развития неблагоприятного исхода в обеих группах уменьшится наполовину.

ЧИСЛО БОЛЬНЫХ, КОТОРЫХ НЕОБХОДИМО ЛЕЧИТЬ

Эффект вмешательства можно выразить и при помощи подсчета числа больных, которых необходимо лечить, чтобы предотвратить определенный неблагоприятный исход у одного больного (ЧБНЛ) [11]. В табл. 2 показано, что смертность в группе перевязки сосудов пищевода составляет 28,1%, а в группе склеротерапии — 44,6%. При достаточной точности этой оценки мы можем сказать, что перевязка сосудов у 100 больных позволяет предотвратить 15 или 16 дополнительных смертей по сравнению со склеротерапией (при таком размере выборки этот показатель равен $САР$, т.е. разнице между частотой развития клинического исхода в контрольной и основной группах). Если проведение лечения 100 больным способствует увеличению числа выживших на 16, скольких больных необходимо лечить, чтобы избежать одной смерти? Для ответа на этот вопрос разделим 100 на 16 ($САР$ в %) и получим ЧБНЛ, приблизительно равное 6. Данное значение можно определить и другим способом, а именно как величину, обратную $САР$ (в виде дроби), по формуле $1 / САР$ (см. табл. 1). И $САР$, и ЧБНЛ зависят от изменения $АР$ развития неблагоприятного клинического исхода на фоне лечения, что не удивительно, поскольку ЧБНЛ обратно пропорционально $САР$. ЧБНЛ можно также рассчитать по номограмме, включающей значения исходного $АР$ и $СОР$ (рис. 2) [12].

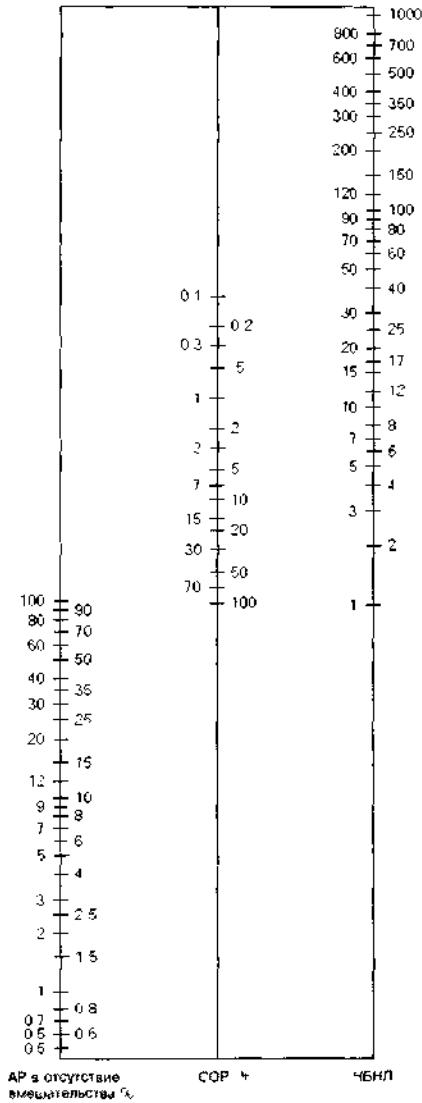


Рис. 2. Номограмма для расчета ЧБНЛ (приводится с разрешения ВМЖ Publishing Group).

ЧБНЛ обратно пропорционально доли больных в контрольной группе, у которых развился неблагоприятный клинический исход. При увеличении этого показателя в 2 раза ЧБНЛ уменьшится в 2 раза, при уменьшении в 4 раза — увеличится в 4 раза. В нашем примере при снижении исходного уровня смертности в 2 раза при неизменном СОР перевязка сосудов пищевода 100 больным приведет к дополнительному предотвращению 8 (22—14) смертей, а ЧБНЛ увеличится вдвое (до 12).

Кроме того, ЧБНЛ обратно пропорционально СОР. Более эффективное вмешательство, вызывающее увеличение СОР в 2 раза, будет характеризоваться двукратным уменьшением ЧБНЛ. Если СОР при одном из вариантов лечения в 4 раза ниже, чем при другом, ЧБНЛ при первом вмешательстве будет в 4 раза выше. В табл. 4 представлены данные, иллюстрирующие эту взаимосвязь.

Значения САР и обратного ему ЧБНЛ отражают изменение исходного риска, однако сами по себе не несут информации об АР в контрольной группе. Так, САР равно 5% (соответственно ЧБНЛ=20) при снижении смертности с 10 до 5% (СОР=50%) и с 50 до 45% (СОР=10%).

ИНДЕКС ПОТЕНЦИАЛЬНОГО ВРЕДА

ИПВ рассчитывается аналогично ЧБНЛ. Если 5 из 100 больных, получающих Р-блокаторы, жаловались на повышенную утомляемость, то 1 побочный эффект наблюдается при лечении 20 больных; иными словами, ИПВ составляет 20.

Таблица 4. Взаимосвязь между исходным АР, СОР и ЧБНЛ

Частота развития клинического исхода в контрольной группе	Частота развития клинического исхода в группе вмешательства	ОР, %	СОР, %	САР	ЧБНЛ
0,02	0,01	50	50	0,01	100
0,4	0,2	50	50	0,2	5
0,04	0,02	50	50	0,02	50
0,04	0,03	75	25	0,01	100
0,4	0,3	75	25	0,1	10
0,01	0,005	50	50	0,005	200

Примечание. ОР равняется отношению частоты развития клинического исхода в основной и контрольной группах; СОР = 1 - ОР; САР равно разнице между частотой развития клинического исхода в основной и контрольной группах; ЧБНЛ = 1 / САР.

При обсуждении данного вопроса мы не упомянули о подходах к рассмотрению отчетов об исследовании, авторы которых представили полученные результаты в виде ОШ, а не ОР. Как уже отмечалось, при низкой частоте развития клинического исхода об ОР можно судить по ОШ. Чем выше ОР, тем меньше достоверность такого суждения. В табл. 5 и 6

Таблица 5. Связь между ЧБНЛ и ОШ (приводится по [18])

Частота развития клинического исхода в контрольной группе	ОШ в группе вмешательства									
	0,5	0,55	0,6	0,65	0,7	0,75	0,8	0,85	0,9	
0,05	41	46	52	59	69	83	104	139	209	
0,1	21	24	27	31	36	43	54	73	110	
0,2	11	13	14	17	20	24	30	40	61	
0,3	8	9	10	12	14	18	22	30	46	
0,4	7	8	9	10	12	15	19	26	40	
0,5	6	7	8	9	11	14	18	25	38	
0,7	6	7	9	10	13	16	20	28	44	
0,9	12	15	18	22	27	34	46	64	101	

Примечание. Формула для расчета ЧБНЛ: $ЧИК = \frac{1 - CER(1 - OR)}{CER(1 - CER)(1 - OR)}$,

где ЧИК — частота развития клинического исхода в контрольной группе.

Таблица 6. Связь между ИПВ и ОШ (приводится по [18])

Частота развития клинического исхода в контрольной группе	ОШ в группе вмешательства									
	1,1	1,2	1,3	1,4	1,5	2	2,5	3	3,5	
0,05	212	106	71	54	43	22	15	12	9	
0,1	112	57	38	29	23	12	9	7	6	
0,2	64	33	22	17	14	8	5	4	4	
0,3	49	25	17	13	11	6	5	4	3	
0,4	43	23	16	12	10	6	4	4	3	
0,5	42	22	15	12	10	6	5	4	4	
0,7	51	27	19	15	13	8	7	6	5	
0,9	121	66	47	38	32	21	17	16	14	

Примечание. Формула для расчета ИПВ: $ЧИК = \frac{CER(1 - OR) + 1}{CER(OR - 1)(1 - CER)}$,

где ЧИК — частота развития клинического исхода в контрольной группе.

приведены данные, которые помогают довольно точно оценить **ЧБНЛ и ИПВ** по значениям **АР** в контрольной группе и **ОШ**.

Еще раз о таблице сопряженности.

Какой бы способ представления размера эффекта вмешательства мы ни выбрали, следует иметь в виду, что таблица сопряженности отражает результаты, полученные в определенный момент времени. Таким образом, оценивая показатели **ОР, САР, СОР, ОШ** и **ЧБНЛ** или **ИПВ**, необходимо указывать длительность наблюдения. Например, следует сообщить о том, что при перевязке сосудов пищевода по сравнению со склеротерапией **САР=17%**, а **ЧБНЛ=6** при средней продолжительности наблюдения, равной 10 мес. Результат будет недостоверным, если продолжительность наблюдения недостаточна (тогда клинический исход не успевает развиваться) или, напротив, чрезмерна (так, при использовании в качестве критерия оценки смертности нужно помнить, что через 100 лет все участники исследования уже умрут).

ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ

До сих пор мы рассматривали точечные значения размера эффекта вмешательств (перевязки сосудов пищевода или склеротерапии). Однако оценка результатов любого эксперимента приближительна. Истинный эффект может быть больше или меньше того, который наблюдали в данном исследовании. **ДИ** показывает, насколько истинное значение того или иного показателя может отличаться от полученного в исследовании (глава 3 *Лечение и понимание результатов исследования: доверительные интервалы*). Статистические программы позволяют рассчитать **ДИ** для каждого из рассмотренных нами показателей.

АНАЛИЗ ВЫЖИВАЕМОСТИ

Как уже отмечалось, данные, включенные в таблицу сопряженности, характеризуют эффект вмешательства за определенный период времени. Такой подход к оценке эффективности вполне приемлем при определении риска развития клинического исхода на протяжении относительно небольшого периода и одинаковой продолжительности наблюдения у всех участников. Однако при проведении длительных испытаний авторов интересует не только общий **АР** клинического исхода, но и распределение клинических исходов во времени. Например, исследование может быть предпринято для того, чтобы получить ответ на во-

прос, позволяли ли лечение при смертельных заболеваниях (тяжелая сердечная недостаточность или неоперабельный рак легкого) увеличить продолжительность жизни.

Для оценки частоты развития какого-либо клинического исхода в зависимости от времени необходимо составить несколько таблиц сопряженности, отражающих результаты, которые были получены в различные сроки. В табл. 2 представлены данные, характеризующие эффект от лечения спустя 10 мес. Можно составить аналогичные таблицы, включающие информацию о всех больных, которые остаются под наблюдением через 1 нед, 1 или 3 мес после начала испытания, другими словами, спустя любой выбранный нами период. Метод, позволяющий оценить частоту развития клинического исхода в зависимости от времени его возникновения, получил название анализа выживаемости. Несмотря на название метода, он применим не только при оценке смертности; подобным образом можно рассчитать вероятность любого дихотомического клинического исхода.

Кривая дожития, составленная для группы больных, отражает частоту развития того или иного исхода в различные периоды времени, прошедшего с момента начала исследования [13]. На рис. 3 представлена кривая дожития из отчета об испытании различных методов лечения при кровотечении из варикозно-расширенных вен пищевода. Поскольку примерно в половине случаев продолжительность наблюдения была больше рассчитанной средней величины (286 дней), кривая выходит за соответствующие пределы. Начиная с определенного момента, значения становятся все более и более приблизительными, так как для точной оценки остается слишком мало больных. Расширение ДИ вокруг точечных значений отражает снижение достоверности результатов.

Даже если истинные значения ОР или СОР в каждый выбранный момент времени одинаковы, влияние случайных факторов приводит к возникновению различий. Поэтому в идеале следует рассчитывать средневзвешенное СОР на основании средних значений и числа больных, данные о которых были включены в анализ в каждый из выбранных моментов времени. Статистические методы позволяют сделать это. Средневзвешенное значение ОР получило название *отношения риска*.

Предположив истинность нулевой гипотезы (т.е. отсутствие различий между группами), мы можем определить значение p , которое отражает вероятность того, что различие в результатах объясняется случайностью. Статистические методы (чаще всего используется регрессионная модель Кокса) позволяют осуществить стандартизацию по характеристикам участников в обеих группах на момент начала испытания (см.

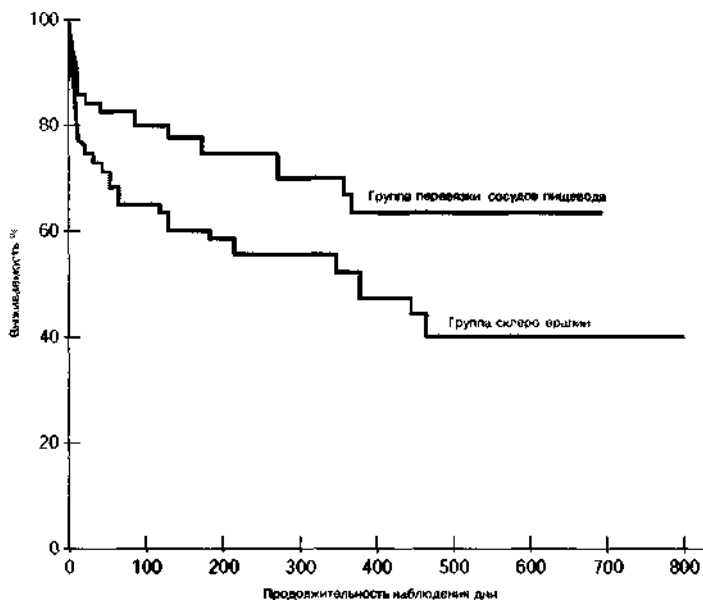


Рис. 3. Кривые дожития, построенные для групп перевязки сосудов пищевода и склеротерапии.

главу 3 *Лечение и понимание результатов исследования: доверительные интервалы*). Если больные в одной из групп старше (и, следовательно, риск развития неблагоприятных клинических исходов у них выше) или, напротив, тяжесть заболевания у них меньше (т.е. риск ниже), авторы могут выбрать статистический метод для коррекции с учетом этих различий. Это позволяет получить такие же результаты, как если бы влияние факторов риска на момент начала испытания в обеих группах было одинаковым.

Другой способ заключается в нанесении на графики точек, соответствующих определенной одинаковой доле участников в каждой группе, у которых развился изучаемый клинический исход. При этом расстояние между точками отражает увеличение срока с начала исследования до наступления клинического исхода в связи с проводимым лечением. Например, в исследовании [14] использование ингибиторов ангиотензин-превращающего фермента (АПФ) при инфаркте задней стенки левого желудочка приводило к снижению смертности на 25%, что на самом деле выражалось увеличением продолжительности жизни в группе

лечения на несколько месяцев. Представленный в такой форме эффект вмешательства выглядит куда менее впечатляющим.

ИССЛЕДОВАНИЯ СЛУЧАЙ—КОНТРОЛЬ

Все данные, рассмотренные ранее в качестве примеров, были получены в ходе проспективных РКИ с формированием 2 групп, в одной из которых применяли то или иное вмешательство, в то время как другая служила для сравнения. Регистрация изучаемых клинических исходов осуществлялась в ходе последующего наблюдения. Методологически такие работы напоминают обсервационные исследования, называемые *проспективными когортными исследованиями*, с той разницей, что в последних отсутствует контрольная группа. По данным РКИ и проспективных когортных исследований можно рассчитать такие показатели, как AP, CAP и OP.

В *исследованиях случай—контроль* группы формируются в зависимости от выявления изучаемых клинических исходов, а не по факту медицинского вмешательства или вредного воздействия. Иными словами, в одну из групп включают участников, у которых на момент начала ретроспективного исследования развился тот или иной клинический исход, например инсульт, инфаркт миокарда или рак, а в другую — тех, у кого данный клинический исход не был зарегистрирован. Задача авторов при этом заключается в выявлении факторов, которые чаще встречаются в одной из групп.

В одном из исследований случай—контроль изучалась связь между использованием солнечных ламп или горизонтальных соляриев и риском развития меланомы кожи [15]. В основную группу включили 583 больных меланомой, в контрольную группу — 608 участников. Группы практически не различались по возрасту, полу и месту жительства. В табл. 7 представлены результаты, полученные у мужчин.

Если бы данные, приведенные в табл. 7, были получены в РКИ, анализ результатов можно было бы начать с определения риска развития

Таблица 7. Связь между риском развития меланомы кожи и использованием солнечных ламп или горизонтальных соляриев (по данным исследования случай—контроль [11])

Воздействие	Основная группа (число заболевших)	Контрольная группа
Да	67	41
Нет	210	242

заболевания в основной и контрольной группах. В данном случае подобный подход недопустим, поскольку число участников контрольной группы выбиралось авторами произвольно. Для оценки ОР необходимо составить представление о популяции, в которой действуют факторы риска, а в исследовании случай—контроль необходимо получить необходимую для этого информацию.

Единственным способом измерения корреляции в таких исследованиях является оценка ОШ. Можно задать вопрос, был ли одинаковым шанс подвергнуться искусственному ультрафиолетовому облучению у больных с меланомой и лиц без этого заболевания. В рассматриваемом исследовании воздействию подверглись 67 из 210 и 41 из 242 участников соответственно. Таким образом, $OШ = (67 / 210) / (41 / 242) = 1,88$ (при 95% ДИ от 1,20 до 2,98), т.е. полученный результат свидетельствует о наличии корреляции между искусственным ультрафиолетовым облучением и развитием меланомы. Диапазон ДИ не включает в себя* 1,0, что не позволяет связать полученный результат с влиянием случайных факторов.

Даже если полученный результат неслучаен, это вовсе не означает, что причиной развития меланомы обязательно было использование солнечных ламп и горизонтальных соляриев. К числу других факторов, способных привести к этому заболеванию или ложному заключению о наличии связи между искусственным ультрафиолетовым облучением, можно отнести более высокую вероятность сообщения о воздействии среди лиц с изучаемым заболеванием (систематическая ошибка, связанная с воспоминанием о воздействии), более длительным пребыванием на солнце этих участников, различие по типу кожи. В ходе исследования авторы изучили многие из этих причин. Для подтверждения того, что посещение солярия приводит к увеличению риска развития меланомы, необходимо проведение дополнительных исследований.

КАКОЙ МЕТОД ИЗМЕРЕНИЯ КОРРЕЛЯЦИИ НАИЛУЧШИЙ?

Какой метод измерения корреляции в наибольшей степени соответствует требованиям научно обоснованного подхода к медицинской практике? На наш взгляд, выбор такого метода важен, поскольку один и тот же результат, представленный в разной форме, может послужить основой для принятия различных клинических решений [16—20]. Например, в исследовании L. Fogrow et al. [16] было показано, что врачи реже принимали решение о проведении вмешательства, если его эффективность

была описана в виде абсолютных, а не относительных изменений. С. D. Nayloret al. [17] также приводят данные, свидетельствующие о том, что воочи ниже оценивают эффективность лечения, если в статье было указано САР вместо СОР. Более того, то же вмешательство признавали еще менее эффективным, если авторы статьи приводили ЧБНЛ. Сотрудникам фармацевтических компаний известно об этих особенностях восприятия врачами результатов исследований, поэтому данные, как правило, представляют в виде СОР в основной группе.

Оказалось, что больные, как и врачи, обращают внимание на форму представления результатов [12, 21—23]. В одном из исследований, в котором участвовали лица с потенциально опасными для жизни заболеваниями, было продемонстрировано, что один и тот же метод лечения вызывает большее доверие больных, если данные о его эффективности были представлены в виде СОР, а не САР [12].

Учитывая перечисленное, мы можем посоветовать изучить все имеющиеся первичные данные (в форме таблиц сопряженности или анализа выживаемости) и полученные результаты в виде относительных и абсолютных показателей. После этого, если вы сумеете рассчитать САР и обратное ему ЧБНЛ, то сможете обоснованно принять решение о целесообразности лечения. Доказательный подход подразумевает рассмотрение всей доступной информации с целью оценки возможного риска и пользы вмешательства в каждом конкретном случае.

Литература

- 1 Callahan C M, Dittus R S, Kalz B P Oral corticosteroid therapy for patients with stable chronic obstructive pulmonary disease a meta-analysis *Ann Intern Med* 1991,114 216—223
- 2 Guyatt G H, Jumper E F, Walter S D, Griffith L E, Goldstein R S Interpreting treatment effects in randomised trials *BMJ* 1998 316 690—693
- 3 Stiegmann G V, Goff J S, Mithaletz-Onod PA, et al Endoscopic sclerotherapy as compared with endoscopic ligation for bleeding esophageal varices *N Engl J Med* 1992,326 1527—1532
- 4 Sinclair J C, Bracken M B Clinically useful measures of effect in binary analyses of randomized trials *J Chn Epidemiol* 1994,47 881—889
- 5 Sackett D L Down with odds ratios¹ *Evid Based Med* 1996,1 164—166
- 6 Laird N M, Mosteller F Some statistical methods for combining experimental results *Int J Technol Assess Health Care* 1990,6 5—30
- 7 Walters D Choice of effect measure for epidemiological data *J Chn Epidemiol* 2000 53 931—939
- 8 Davies H T, Crombie I K, Tavakoli M When can odds ratios mislead² *BMJ* 1998,316 989—991
- 9 Zhang J, Yu K F What's the relative risk³ A method of correcting the odds ratio in cohort studies of common outcomes *JAMA* 1998,280 1690—1691
- 10 Laine L, Cook D Endoscopic ligation compared with sclerotherapy for treatment of esophageal variceal bleeding a meta-analysis *Ann Intern Med* 1995 123 280—287
- 11 Loupapaou A, Sackett D L, Roberts R S An assessment of clinically useful measures of the consequences of

- treatment N Engl J Med 1988,318 1728-1733
- 12 *Chatellier G , Zapletal E, Lemattre D , Menard J, Degoulet P* The number needed to treat a clinically useful nomogram in its proper context BMJ 1996,312 426-429
 - 13 *ColdmanAJ, Elwood J M* Examining survival data CMAJ 1979,121 1065-1068, 1071
 - 14 *Tan L B, Murphy R* Shifts in mortality curves saving or extending lives¹ Lancet 1999,354 1378-1381
 - 15 *Walter S D , Marrett L D , From L , Hertzman C, Shannon H S , Roy P* The association of cutaneous malignant melanoma with the use of sunbeds and sunlamps Am J Epidemiol 1990,131 232-243
 - 16 *Forrow L , Taylor WC, Arnold R M* Absolutely relative how research results are summarized can affect treatment decisions Am J Med 1992,92 121-124
 - 17 *Naylor C D , Chen E, itrauss B* Measured enthusiasm does the method of reporting trial results alter perceptions of therapeutic effectiveness⁹ Ann Intern Med 1992,117 916-921
 - 18 *Hux J E, Levinton C M, Naylor C D* Prescribing propensity influence of life-expectancy gains and drug costs J Gen Intern Med 1994,9 195-201
 - 19 *Redelmeier D A , TverskyA* Discrepancy between medical decisions for individual patients and for groups N Engl J Med 1990,322 1162-1164
 - 20 *Bobbio M, Demicheis B , Giustetto G* Completeness of reporting trial results effect on physicians' willingness to prescribe Lancet 1994,343 1209-1211
 - 21 *Malenka D J, Baron J A , Johansen S, Wahrenberger J W', Ross J M* The framing effect of relative and absolute risk J Gen Intern Med 1993,8 543-548
 - 22 *McNedBJ Pauker S G, Sox H C Jr TverskyA* On the elicitation of preferences for alternative therapies N Engl J Med 1982,306 1259-1262
 - 23 *Hux J E Naylor C D* Communicating the benefits of chronic preventive therapy does the format of efficacy data determine patients' acceptance of treatment⁹ Med Decis Making 1995,15 152-157

ПРИЛОЖЕНИЕ

СПОСОБЫ РАСЧЕТА ОСНОВНЫХ ПОКАЗАТЕЛЕЙ

Р. Леунг

В ЭТОМ РАЗДЕЛЕ

Лечение и риск

Диагностика

Степень согласованности мнений, рассчитываемая с учетом влияния фактора случайности: коэффициент согласия k

Число больных, которых необходимо лечить определенным методом в течение определенного времени, чтобы достичь одного благоприятного исхода или предотвратить один неблагоприятный клинический исход

Доверительные интервалы

ЛЕЧЕНИЕ И РИСК

(См главы 4 *Исследования, посвященные эффективности лечения*, 5 *Оценка риска первой части*, главу 4 второй части *Лечение и понимание результатов исследования измерение корреляции*)

		Клинический исход	
		Да	Нет
Вмешательство	Да	a	b
	Нет	c	d

$$\text{Частота возникновения исхода в контрольной группе (ЧИК)} = \frac{c}{(c+d)}$$

$$\text{Частота возникновения исхода в группе лечения (ЧИЛ)} = \frac{a}{(a+b)}$$

$$\text{Относительный риск (ОР)} = \frac{a/(a+b)}{c/(c+d)}$$

$$\text{Снижение относительного риска (COR)} = 1 - \text{ОР} = \frac{c/(c+d) - a/(a+b)}{c/(c+d)}$$

$$\text{Снижение абсолютного риска (САР)} = \frac{c}{c+d} - \frac{a}{a+b}$$

$$\text{Число больных, которых необходимо лечить (ЧБНЛ)} = \frac{1}{\text{САР}}$$

$$\text{Отношение шансов (ОШ)} = \frac{a/b}{c/d} = \frac{ad}{cb}$$

Расчет ЧБНЛ с использованием значений ЧИК и ОШ:

$$\text{ЧБНЛ} = \frac{1 - \text{ЧИК}(1 - \text{ОШ})}{\text{ЧИК}(1 - \text{ЧИК})(1 - \text{ОШ})}$$

Расчет индекса потенциального вреда (ИПВ) с использованием значений ЧИК и ОШ:

$$\text{ИПВ} = \frac{1 + \text{ЧИК}(1 - \text{ОШ})}{\text{ЧИК}(1 - \text{ЧИК})(1 - \text{ОШ})}$$

ДИАГНОСТИКА

(См. главу 8 *Диагностические методы*)

Результат «золотого стандарта»

		Результат «золотого стандарта»	
		Положительный	Отрицательный
Результат изучаемого теста	Положительный	a	b
	Отрицательный	c	d

Частота истинноположительного результата изучаемого диагностического теста = a

Частота истинноотрицательного результата изучаемого диагностического теста = b

Частота ложноположительного результата изучаемого диагностического теста = c

Частота ложноотрицательного результата изучаемого диагностического теста = d

$$\text{Чувствительность} = \frac{a}{a+c}$$

$$\text{Специфичность} = \frac{d}{b+d}$$

$$\text{Отношение правдоподобия положительного результата (ОП+)} = \frac{a/(a+c)}{b/(b+d)}$$

$$\text{Отношение правдоподобия отрицательного результата (ОП-)} = \frac{c/(a+c)}{d/(b+d)}$$

$$\text{Прогностическая ценность положительного результата} = \frac{a}{a+b}$$

$$\text{Прогностическая ценность отрицательного результата} = \frac{d}{c+d}$$

$$\text{Диагностическая точность} = \frac{a+d}{a+b+c+d}$$

$$\text{Претестовая вероятность (распространенность)} = \frac{a+c}{a+b+c+d}$$

$$\text{Претестовые шансы} = \frac{\text{распространенность}}{1 - \text{распространенность}} = \frac{a+c}{b+d}$$

Послетестовые шансы = претестовые шансы \times ОП

Послетестовая вероятность = послетестовые шансы / (1 + послетестовые шансы)

СТЕПЕНЬ СОГЛАСОВАННОСТИ МНЕНИЙ, РАССЧИТЫВАЕМАЯ С УЧЕТОМ ВЛИЯНИЯ ФАКТОРА СЛУЧАЙНОСТИ: КОЭФФИЦИЕНТ СОГЛАСИЯ К

(См. главу 2С на компакт-диске)

		Мнение исследователя В	
		Да	Нет
Мнение исследователя А	Да	a	b
	Нет	c	d

Коэффициент согласия

$$k \text{ (каппа)} = \frac{\text{выявленное согласие} - \text{ожидаемое согласие}}{1 - \text{ожидаемое согласие}}, \text{ где:}$$

$$\text{выявленное согласие} = \frac{a + d}{a + b + c + d},$$

$$\text{ожидаемое согласие} = \frac{(a + b)(a + c)}{a + b + c + d} + \frac{(c + d)(b + d)}{a + b + c + d}$$

$$\text{Отношение шансов (ОШ)} = \frac{ad}{bc}$$

$$\text{Коэффициент } \Phi \text{ (фи)} = \frac{\sqrt{\text{ОШ}} - 1}{\sqrt{\text{ОШ}} + 1} + \frac{\sqrt{ab} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}}$$

ЧИСЛО БОЛЬНЫХ, КОТОРЫХ НЕОБХОДИМО ЛЕЧИТЬ ОПРЕДЕЛЕННЫМ МЕТОДОМ В ТЕЧЕНИЕ ОПРЕДЕЛЕННОГО ВРЕМЕНИ, ЧТОБЫ ДОСТИЧЬ ОДНОГО БЛАГОПРИЯТНОГО ИСХОДА ИЛИ ПРЕДОТВРАТИТЬ ОДИН НЕБЛАГОПРИЯТНЫЙ КЛИНИЧЕСКИЙ ИСХОД

(См. главу 2F на компакт-диске)

$$\text{ЧБНЛ} = \frac{\text{затраты}_{\text{лки}} + \text{затраты}_{\text{пки}}}{\text{затраты}_{\text{л}} + \sum (\text{затраты}_{\text{лпэ}})(\text{частота}_{\text{лэ}}) + \sum (\text{затраты}_{\text{ппэ}})(\text{частота}_{\text{пэ}})},$$

где

ЧБНЛ — значение ЧБНЛ, при котором затраты на профилактику будут эффективными;

*затраты*_{лки} — затраты на лечение одного изучаемого клинического исхода;

*затраты*_{лпэ} — затраты на лечение одного побочного эффекта вмешательства,

*затраты*_л — общие затраты на лечение одного больного,

*частота*_{лэ} — частота развития побочных эффектов у пролеченных больных;

*затраты*_{пки} — затраты на предотвращение одного изучаемого клинического исхода;

*затраты*_{ппэ} — затраты на предотвращение одного побочного эффекта вмешательства

ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ

(См главу 3 второй части *Лечение и понимание результатов исследования: доверительные интервалы*)

При проведении исследования в формате 2x2 выборку в целом можно представить следующим образом.

	Показатель 1	Показатель 2	Общее число наблюдений
Группа 1	a	b	n
Группа 2	c	d	m

Для данной таблицы можно рассчитать следующие доверительные интервалы [1]:

Математическое действие при оценке точечного значения	Точечное значение	Доверительные интервалы	Примеры рассчитываемых параметров
Простое деление	$\frac{a}{n}$	$\frac{a}{n} \pm z \sqrt{\frac{a(n-a)}{n^3}}$	ЧИК, ЧИЛ, чувствительность, специфичность, прогностическая ценность положительного результата, прогностическая ценность отрицательного результата
Вычитание одного результата простого деления из другого	$\frac{a}{n} - \frac{c}{m}$	$(\frac{a}{n} - \frac{c}{m}) \pm z \sqrt{\frac{a(n-a)}{n^3} + \frac{c(m-c)}{m^3}}$	САР
Деление одного результата простого деления на другой	$\frac{a/n}{c/m}$	$\frac{a/n}{c/m} e^{\pm z \sqrt{\frac{1}{a} \frac{1}{n} + \frac{1}{c} \frac{1}{m}}}$	ОР, ОП+, ОП-
Деление отношения между значениями показателя в одной группе на аналогичное отношение в другой группе	$\frac{a/b}{c/d}$	$\frac{a/b}{c/d} e^{\pm z \sqrt{\frac{1}{a} \frac{1}{b} + \frac{1}{c} \frac{1}{d}}}$	ОШ

Примечание $z = 1.96$ для 95% доверительного интервала

Литература

СЛОВАРЬ ТЕРМИНОВ*

Термин	Определение	См. также
Автокорреляция	Наблюдается, если результаты предшествующих наблюдений влияют на результаты последующих. Например, при хроническом заболевании вероятность хорошего самочувствия после относительно благополучных суток выше, чем после неблагоприятных.	
Алгоритм лечебный или диагностический	Ясное подробное описание последовательных действий при оказании помощи больным в определенных клинических ситуациях.	
Альтернативный диагноз	См. <i>Дифференциальный диагноз</i>	
Анализ выгоды-затрат	Разновидность анализа затрат, при котором учитываются стоимость того или иного вмешательства и возможные последствия его проведения (включая увеличение продолжительности и качества жизни), результат выражается в денежном эквиваленте.	
Анализ выживаемости	Анализ, при котором предметом изучения служит не только частота развития того или иного клинического исхода, но и срок его возникновения.	
Анализ данных в зависимости от назначенного лечения (анализ, проводимый исходя из допущения, что *се больные получили предписанное лечение)	Анализ, в который включаются данные о клинических исходах у всех больных, в случайном порядке отнесенных к той или иной группе, независимо от того, получали ли участники запланированное вмешательство или нет. Анализ позволяет сохранить преимущество рандомизации, которое заключается в том, что неизвестные факторы, способные повлиять на результат, примерно одинаково представлены в сравниваемых группах.	
Анализ затрат	Сравнение двух вмешательств, исходя из их стоимости, позволяет получить данные толь-	

* Син — синоним

€_М — основной термин

^x также — термины, связанные с основным

**которые из приводимых ниже терминов используются только в электронной версии книги (При-

*«<• ред)

Термин	Определение	См. также
	ко о затратах на профилактику или лечение (но не о клинических исходах)	
Анализ минимизации затрат	Разновидность анализа затрат, при котором эффективность сравниваемых вмешательств одинакова, поэтому целью исследования является сопоставление их стоимости	
Анализ полезности затрат	Разновидность анализа эффективности затрат, при котором эффект вмешательства выражен в виде приемлемости (предпочтительности) клинических исходов для больного. Обычно рассматривается стоимость улучшения качества жизни на один условный год	Число сохраненных лет жизни с поправкой на ее качество
Анализ принятия решений	Систематический подход к выбору оптимального клинического решения в условиях неопределенности. Он включает в себя перечисление как можно большего числа альтернативных клинических исходов, определение их вероятности, приемлемости и, наконец, количественную оценку относительной целесообразности выбора каждого из возможных вмешательств	
Анализ чувствительности	Позволяет определить значительность изменений, которым подвергнутся результаты анализа принятия решений при изменении вероятности развития клинических исходов, предпочтений или структуры принятия решения. Анализ чувствительности может заключаться в проведении повторного анализа принятия решений после изменения одного или более параметров	
Анализ эффективности затрат	Разновидность анализа затрат, при котором стоимость вмешательства выражается в относительных единицах, например, на одну сохраненную жизнь или при снижении АД на 1 мм рт ст	
Байесовский анализ	Анализ, при котором на основании исходной информации о вероятности какого-либо события (претестовая вероятность) и данных теста уточняют значение вероятности (послетестовая вероятность)	
Величина/)	При условии правильности нулевой гипотезы вероятность того, что при повторении эксперимента будут наблюдаться различия между	Вероятность

Термин	Определение	См. также
	группами, равные наблюдаемым или даже превосходящие их	
Вероятность	Показатель, отражающий правдоподобие наличия определенного заболевания или состояния (в исследованиях диагностических тестов) или последующего развития клинического исхода (в исследованиях методов лечения)	Величина <i>p</i>
Вероятность ошибки I рода (а-ошибки)	Вероятность ошибочного заключения о существовании различия между двумя методами лечения, когда в действительности это различие отсутствует. Обычно исследователи выбирают допустимую величину данного показателя на этапе планирования размеров выборки	
Верхнее пороговое значение вероятности наличия заболевания, превышение которого указывает на необходимость немедленного начала лечения	Если в результате проведения диагностического теста показана вероятность наличия заболевания выше данной величины, необходимо немедленное начало лечения	Нижнее пороговое значение вероятности наличия заболевания, превышение которого указывает на необходимость дальнейшего обследования
Вмешивающийся (искажающий) фактор	Фактор, искажающий результат исследования в связи с тем, что он также влияет на вероятность развития изучаемого клинического исхода. Нередко влияние вмешивающихся факторов в сравниваемых группах различно. Вероятность искажения результатов в РКИ вследствие вмешивающихся факторов ниже, чем в наблюдательных исследованиях	
Внешняя Достоверность	Метод измерения обладает внешней достоверностью, если он выглядит способным измерять именно то явление, которое должен регистрировать	
Воздействие	Положительно или отрицательно влияет на состояние здоровья участников исследования	
	Любые изменения состояния здоровья в определенной группе, которые каким-либо образом могут быть связаны с вредным или полезным воздействием либо медицинским вмешатель-	

Возможные клинические исходы

	ством, включая изменение продолжительности и качества жизни в связи с лечением того или иного заболевания при его наличии либо в его отсутствие, отказом от проведения лечения вследствие правильного или ошибочного заключения об отсутствии заболевания
Вопросы, подразумевающие ответы в свободной форме; частично структурированные и структурированные вопросы	Вопросы первого типа не требуют соблюдения определенной формы ответа; ответ на структурированный вопрос должен быть построен по заданному плану
Воспроизводимость результатов исследований	Возможность воспроизведения полученных результатов в других исследованиях
Вред	Неблагоприятные последствия, связанные с воздействием того или иного фактора
Выборка по принципу «снежного кома»	Целевая выборка, которая формируется для репрезентативного представления самых разных характеристик, включая типичные; нетипичные случаи; крайние случаи; случаи, отражающие важные социально-политические аспекты; случаи, взаимосвязанные с другими ситуациями или состояниями
Выбывшие участники	Участники исследования, чье состояние на момент завершения наблюдения неизвестно
Гетерогенность (неоднородность)	Различия между больными или результатами исследований
Градиент зависимости эффекта от дозы	См. <i>Зависимость эффекта от дозы</i>
Деонтологический подход к медицинским решениям	Деонтологический подход к распределительной справедливости подразумевает, что единственной задачей врача служит удовлетворение потребностей конкретного больного. Противоречит утилитарному подходу

Термин	Определение	См. также
	ством, включая изменение продолжительности и качества жизни в связи с лечением того или иного заболевания при его наличии либо в его отсутствие, отказом от проведения лечения вследствие правильного или ошибочного заключения об отсутствии заболевания	
«опросы, охватывающие ответы в свободной форме; частично структурированные и структурированные опросы	Вопросы первого типа не требуют соблюдения определенной формы ответа; ответ на структурированный вопрос должен быть построен по заданному плану	
«воспроизводимость результатов исследований	Возможность воспроизведения полученных результатов в других исследованиях	
«риск»	Неблагоприятные последствия, связанные с воздействием того или иного фактора	
«выборка по принципу «снежного кома»	Целевая выборка, которая формируется для репрезентативного представления самых разных характеристик, включая типичные; нетипичные случаи; крайние случаи; случаи, отражающие важные социально-политические аспекты; случаи, взаимосвязанные с другими ситуациями или состояниями	
«бывшие участники»	Участники исследования, чье состояние на момент завершения наблюдения неизвестно	
«гетерогенность (однородность)»	Различия между больными или результатами исследований	
«зависимость эффекта от дозы»	См. <i>Зависимость эффекта от дозы</i>	
«деонтологический подход к клиническим решениям»	Деонтологический подход к распределительной справедливости подразумевает, что единственной задачей врача служит удовлетворение потребностей конкретного больного. Противоположен утилитарному подходу	

Термин	Определение	См. также
Дерево решений	В большинстве случаев поиск оптимального решения осуществляется с построением дерева решения В статьях, посвященных принятию клинических решений, обычно приводятся одна или несколько диаграмм, в которых наглядно продемонстрирована структура анализа	
Дискриминантный анализ	Статистический метод, аналогичный логистическому регрессионному анализу, используемый для определения факторов, связанных с наличием либо отсутствием того или иного клинического исхода	
Дисперсия	Технический термин, используемый при статистической оценке вариабельности результатов исследования	
Дифференциальная диагностика	Выбор одного из нескольких диагнозов, каждый из которых может объяснить состояние больного	
Дихотомическая переменная	Переменная, которая может принимать только одно из двух значений, например беременная или небеременная женщина, смерть или выживание, развитие инсульта или его отсутствие	
Дихотомический клинический исход	Признак, который может иметь только два значения (присутствует или отсутствует, да или нет), например рецидив рака, инфаркт миокарда, смерть	
Доверительный интервал (ДИ)	Диапазон, в пределах которого может лежать истинное значение показателя в популяции, из которой сформирована выборка исследования	
Доказательная медицина (ДМ)	Продуманное, строгое и последовательное применение самой достоверной научной информации в процесс принятия клинических решений у конкретных больных Для осуществления принципов доказательной медицины (ДМ) на практике врач должен обладать богатым клиническим опытом, учитывать жизненные ценности и предпочтения больных, а также использовать самые надежные данные систематических исследований	
Долговременные тенденции	Изменение вероятности развития клинического исхода с течением времени независимо от влияния известных прогностических факторов	

Термин	Определение	См. также
Достоверность	Применительно к исследованиям, посвященным диагностике и лечению, результат можно считать достоверным, если он не подвержен влиянию систематических ошибок. При изучении качества жизни, связанного со здоровьем, достоверность означает соответствие проведенного измерения запланированному	
Достоверность конструкта	Конструкт обобщает представления об изучаемых явлениях, основанные на предшествующем опыте. От правильности этих представлений зависят предположения о результатах анкетирования при условии достоверности его методов. При оценке достоверности конструкта для измерения одного и того же явления используют различные шкалы и определяют логическую связь между результатами анкетирования и характеристиками отдельных больных или групп больных	
Жизненные ценности больного	Лежат в основе предпочтений конкретного больного	
Заболеваемость	Число новых случаев заболевания, выявленных на протяжении определенного времени; исчисляется в процентах от общего числа лиц, у которых может возникнуть это заболевание	Распространенность
Зависимая переменная	В регрессионном анализе выделяют прогностический фактор (независимую переменную) и изучаемый клинический исход (зависимую переменную)	
Зависимость эффекта от дозы	Увеличение риска развития того или иного клинического исхода по мере увеличения интенсивности либо длительности предполагаемого вредного воздействия <i>Син.: градиент зависимости эффекта от дозы</i>	
Загрязнение	Непреднамеренное применение вмешательства, которое было назначено участникам одной из групп, больными, включенными в другую группу	
Закон перемножения вероятностей	Правило расчета общей вероятности независимых (не влияющих друг на друга) событий. Пример: при 10 последовательных подбрасываниях монеты	

Термин	Определение	См. также
	ваниях монеты вероятность того, что 10 раз подряд выпадет решка, рассчитывается путем возведения вероятности одного события (1/2) в десятую степень, т.е. равна $0,5^{10}$ (0,5-0,5-0,5 и т.д.)	
Затраты на здравоохранение	Прямые и непрямые затраты на нужды здравоохранения, когда ресурсы не могут быть использованы на другие полезные цели (альтернативные издержки)	
Здоровье	Состояние физического, психического и социального благополучия; под здоровьем нельзя понимать только отсутствие заболевания и физических недостатков (определение ВОЗ)	
«Золотой стандарт»	См. <i>Стандартный диагностический метод</i>	
Здравоохранение, основанное на принципах доказательной медицины	См. <i>Доказательная медицина</i>	
Избирательный скрининг	Проводится в отсутствие симптомов, но при наличии одного или более факторов риска развития искомого заболевания, например указаний на заболевания ближайших родственников, особенностей образа жизни или принадлежности обследуемого к популяции с высокой распространенностью соответствующего заболевания	Скрининг
Исследуемое заболевание	В исследованиях диагностических тестов так называют заболевание, наличие которого подтверждают или отвергают у участников (например, туберкулез, рак легких или железодефицитная анемия)	
Исследуемый клинический исход	В исследованиях эффективности лечения так называют клинический исход, который выявляют при наблюдении и стремятся предотвратить (например, инфаркт миокарда, инсульт или смерть) или, наоборот, достигнуть (заживление язвы) с помощью вмешательства	Когортное исследование
Индекс капта (к), или средне-	Отражает вероятность того, что согласие между экспертами объясняется не только СЛУЧЗИ-	

термин	Определение	См. также
«эшелонное значение индекса аппа	ностью Может принимать значения от $-1,0$ до $+1,0$	
Индекс потенци- ального вреда ИПВ)	Число больных, которым необходимо проводить данное вмешательство в течение определенного периода времени для того, чтобы развился побочный эффект. Обратно пропорционален повышению абсолютного риска (ПАР)	Повышение абсолютного риска (ПАР)
Информацион- ные пособия, физические блечить риятие цинических >ешений	Доходчиво предоставляют больным подробную количественную информацию о сравнительных преимуществах и недостатках возможных медицинских вмешательств	
Информирован- ное согласие	Согласие потенциального участника исследования на проведение изучаемых вмешательств, полученное после того, как он будет полностью информирован обо всех аспектах исследования	
Испытание методов лечения	В ходе такого испытания больному проводят то или иное вмешательство, через заранее определенный период времени оценивают его эффект, после чего в зависимости от полученных результатов рекомендуют продолжение терапии либо ее прекращение	
Исследование до—после	Исследование, в ходе которого авторы сравнивают состояние участников до и после изучаемого вмешательства	Перекрестное испытание
Исследование или испытание с положительным результатом	Исследования, в которых были получены данные, свидетельствующие о разной эффективности изучаемого и контрольного вмешательства	
Исследование случай— контроль	Исследование, организованное для выявления связи между каким-либо фактором риска и клиническим исходом. В таком исследовании сравнивают долю участников, испытавших вредное воздействие, в двух группах, в одной из которых развился, а в другой не отмечался изучаемый клинический исход	
Исследования (испытания)	Проводят для того, чтобы подтвердить или опровергнуть одинаковую клиническую эффек-	

Термин	Определение	См. также
эквивалентности вмешательств	тивность традиционных или стандартных вмешательств и более дешевых, простых либо безопасных методов	
Исследования I фазы	См <i>Структура исследования</i>	
Исследования II фазы	См <i>Структура исследования</i>	
Исследования III фазы	См <i>Структура исследования</i>	
Исследования IV фазы	См <i>Структура исследования</i>	
Исследования с отрицательным результатом	Исследования, в которых были получены данные, свидетельствующие о том, что изучаемое вмешательство не более эффективно, чем контрольное	
Исходная дата	Время наступления определенного события, после которого начинается наблюдение за участником исследования с целью выявления изучаемого клинического исхода	
Исходная когорта	Группа лиц, наблюдение за которыми началось в одно и то же время, сразу после развития у них того или иного заболевания (постановки диагноза) либо после первого воздействия предполагаемого фактора риска	Когорта, Когортное исследование
Исходный риск	Риск развития неблагоприятного исхода в контрольной группе См <i>число исходов в группе контроля (ЧИК)</i>	
Качественное исследование	Исследование, посвященное социальным и эмоциональным аспектам медицинской помощи, а также личным переживаниям потребителей и производителей медицинских услуг	Количественное исследование
Качество жизни, связанное со здоровьем	Мера самочувствия больного и того значения, которое он придает своему состоянию здоровья	
Качество медицинской помощи	Степень соответствия системы здравоохранения техническим и гуманитарным стандартам высококачественной медицинской помощи	Профиль здоровья

Термин	Определение	См. также
Квалифицированная медицинская помощь	<i>См. Специализированная (квалифицированная) медицинская помощь</i>	
Клинические (практические) руководства	Разрабатываемые по определенной методике рекомендации, которые должны помочь врачам и больным принимать клинические решения в различных ситуациях. Представляют собой свод утверждений, примерных указаний или принципов оказания медицинской помощи, отражающих общепринятые либо новые правила, подходы к лечению и профилактике. В практических руководствах приводятся показания к проведению вмешательств или отдельных процедур при решении конкретной клинической задачи. Разрабатываются государственными агентствами, медицинскими центрами, другими организациями, такими как профессиональные сообщества, специально создаваемые рабочие группы или советы экспертов.	
Клинические исходы	Варианты изменения состояния здоровья участников исследования вследствие воздействия изучаемого воздействия или медицинского вмешательства.	Цель лечения; Условие прекращения исследования
Клинический признак	Любой признак того или иного заболевания, выявленный при обследовании больного врачом. Служит объективным свидетельством в пользу наличия заболевания.	
Когорта	Группа лиц со схожей характеристикой или характеристиками. За такой группой обычно наблюдают в течение определенного времени для определения частоты развития того или иного заболевания или осложнения (прогноз).	Когортное исследование
Когортное исследование (когортное аналитическое исследование)	Проспективное исследование факторов, которые могут послужить причиной развития того или иного заболевания. В ходе когортного исследования из лиц без изучаемого клинического исхода формируются две группы, в одной из которых участники подвергаются воздействию вредного фактора, а в другой — нет. При последующем наблюдении сравнивают частоту развития клинического исхода. <i>Син...проспективное исследование</i>	Когорта; Исходная когорта

Термин	Определение	См. также
Количественное исследование	Предпринимается для проверки заранее сформулированной гипотезы, описывающей зависимость между двумя переменными или более. В количественном исследовании получают данные, которые могут быть включены в статистический анализ.	Качественное исследование
Команда «explode» ("расширить поиск")	Используется при поиске в базе данных Medline, позволяя найти все статьи, индексированные с использованием медицинских предметных рубрик, а также более узких ключевых понятий.	
Компьютерные системы для облегчения принятия решений	Компьютерные программы, созданные специально для помощи в принятии клинического решения у конкретного больного.	
Контролируемое испытание	См. <i>Рандомизированное контролируемое испытание</i>	
Контроль над расходованием медицинских ресурсов	Оценка специальной комиссией целесообразности госпитализации, длительности пребывания в стационаре и объемов стационарной медицинской помощи для определения необходимости в такого рода медицинских услугах и выработке оптимальных методов расходования ресурсов здравоохранения.	
Контрольная группа	Группа, в которой участники не получают изучаемого вмешательства либо получают стандартное или самое эффективное по имеющимся данным лечение.	
Концепция	Положения, из которых складываются теории.	
Корреляция	Величина (сила) взаимосвязи между различными переменными или явлениями.	
Косвенные критерии оценки	Используются при измерении показателей, которые сами по себе не представляют большого значения для больных, но связаны с важными клиническими исходами (как, например, минеральная плотность костной ткани с риском возникновения переломов, уровень холестерина в крови с риском развития инфаркта миокарда, уровень АД с риском развития инсульта).	

Термин	Определение	См. также
Коэффициент корреляции	Числовое выражение силы взаимосвязи между двумя переменными в числовом выражении, которое может принимать значения от -1,0 до 41,0	
Кривая дожития	Графическое представление доли больных, остающихся в живых (или тех, у кого не развился тот или иной клинический исход) на определенный момент времени, от начала (100% участников) до конца наблюдения Син.: <i>кривая Каплана—Мейера</i>	
Кривая Каплана—Мейера	См. <i>Кривая дожития</i>	
Критерии включения	Критерии, в соответствии с которыми осуществляется отбор участников в исследование из популяции	
Критерии исключения	Критерии, в соответствии с которыми больные не включаются в выборку исследования	
Критерии качества медицинских услуг	Применяются для оценки качества вмешательства безотносительно того, необходимо ли оно. Например, результаты пробного определения уровня холестерина в крови укладываются в допустимый диапазон вокруг известного значения этого показателя в образце	
Критерий знаков	Статистический метод проверки нулевой гипотезы об одинаковом правдоподобии любого из дихотомических клинических исходов	
Критерий Ф(фи-статистика)	Показатель (метод оценки) степени согласия между исследователями, независимой от случайных факторов. Рассчитывается по формуле: $\frac{\sqrt{OШ} - 1}{\sqrt{OШ} + 1}$, где OШ — отношение шансов	
Критическая оценка клинического решения	Способ компьютерной поддержки работы врача, состоящий в предложении альтернативных решению врача вариантов	
Линейный регрессионный анализ	Разновидность регрессионного анализа, при котором зависимая или изучаемая переменная представлена непрерывными данными, и пред-	

Термин	Определение	См. также
	полагается линейная связь между зависимой и независимой переменными	
Логистическая регрессия	Разновидность регрессионного анализа, при котором зависимая или изучаемая переменная представлена дихотомическими данными, и используется логарифмическая модель	
Логические (булевы) операторы	Команды, используемые для сужения параметров поиска в базах данных, например, для объединения нескольких терминов (AND/OR) или исключения некоторых из них (NOT)	
Ложноотрицательный результат	В исследованиях методов лечения ложноотрицательным результатом считается заключение о неэффективности вмешательства, в то время как в действительности оно эффективно. В исследованиях диагностических методов ложноотрицательным результатом теста считается заключение об отсутствии у больного соответствующего заболевания, в то время как в действительности оно имеется	
Ложноположительный результат	В исследованиях методов лечения ложноположительным результатом считается заключение об эффективности вмешательства, в то время как в действительности оно неэффективно. В исследованиях диагностических методов ложноположительным результатом теста считается заключение о наличии у больного соответствующего заболевания, в то время как в действительности оно отсутствует	
Медиана продолжительности жизни	Время, в течение которого остаются в живых половина участников исследования	
Медицинские работники	Врачи-неспециалисты, интерны и врачи-специалисты, младшие и квалифицированные медицинские сестры, помощники врача и другие лица, участвующие в оказании медицинской помощи	Профессиональные медицинские работники
Мера приемлемости клинического исхода	Метод измерения, в котором учтены все аспекты качества жизни, связанного со здоровьем, а также предпочтения и жизненные ценности больных. Крайние значения этого параметра представлены смертью и полным здоровьем	Качество жизни, связанное со здоровьем

Термин	Определение	См. также
Мета-анализ	Обзор, в котором при помощи количественного метода оценки обобщены данные нескольких исследований, а итоговый результат представлен в виде одного средневзвешенного показателя	
Метод бутстреп	Статистический метод, позволяющий рассчитать такие показатели, как стандартная ошибка и доверительный интервал путем повторного извлечения выборок из имеющегося набора данных	
Минимальная приемлемость	Нижний порог приемлемости (предпочтений или соответствия приемлемости жизненным ценностям больного) того или иного клинического исхода по мере увеличения вероятности его развития	
Многофакторное (многомерное) уравнение регрессии	Разновидность регрессионного анализа, при котором строится математическая модель, позволяющая объяснить или предсказать величину зависимой либо изучаемой переменной путем одновременного учета влияния всех независимых переменных или прогностических факторов	Многофакторный анализ
Многофакторный (многомерный) анализ	Анализ, при котором одновременно учитывается влияние нескольких независимых переменных	Многофакторное уравнение регрессии
Мозаичный подход к сбору информации	Используется в качественных исследованиях для сопоставления данных, порученных из разных источников	
Недополученная выгода	Равна материальным и прочим благам (включая улучшение здоровья населения), которые могли быть получены при проведении альтернативного вмешательства	
Нежелание рисковать	Характерно для людей, которые предпочитают постоянный исход лотерее, даже когда она сулит более благоприятный вариант. Такой человек предпочтет гарантированное получение 10 долларов США, а не равную вероятность либо не получить ничего, либо выиграть 30 долларов США	
Независимая непемная	Причинный или прогностический фактор, который может быть связан с определенным кли-	

Термин	Определение	См. также
Независимая связь	<p>ническим исходом Данный термин обычно используется при регрессионном анализе</p> <p>Связь между двумя событиями, значимость которой не исчезает после поправки на все другие предполагаемые факторы, способные повлиять на прогноз</p>	
Непрерывная переменная	Переменная, которая может принимать любые значения с небольшими различиями между ними	
Непрямые расходы и выгоды	Влияние проведения альтернативных вмешательств на другие клинические исходы (помимо изучаемого) и другие аспекты оказания медицинской помощи	
Нерандомизированное испытание	Исследование, в котором распределение участников в группы происходит по воле исследователя или заранее намеченному плану, а не в случайном порядке	Рандомизированное контролируемое испытание (РКИ)
Нижнее пороговое значение вероятности наличия заболевания, превышение которого указывает на необходимость дальнейшего обследования	Если в результате проведения диагностического теста вероятность наличия заболевания опускается ниже данной величины, можно отвергнуть рассматриваемый диагноз и отказаться от дальнейшего обследования	
Нижнее пороговое значение вероятности наличия заболевания, превышение которого указывает на необходимость дальнейшего обследования	Если в результате проведения диагностического теста показана вероятность наличия заболевания ниже данной величины, необходимо дальнейшее обследование	Верхнее пороговое значение вероятности наличия заболевания, превышение которого указывает на необходимость немедленного начала лечения
Номограмма	Система шкал, облегчающая расчет вероятности	
Нулевая гипотеза	Начальная гипотеза, для подтверждения или опровержения которой проводится статистический анализ	

Термин	Определение	См. также
Обеспечение высокого качества медицинской помощи	Все процедуры методы или концепции, используемые при сборе, обработке и анализе данных с целью поддержания или повышения качества медицинской помощи	
Обзор	Общий термин, которым обозначают поиск и синтез результатов двух и более исследований по одной теме	
Обобщаемость	Возможность распространить результаты исследования на большую группу лиц со сходными характеристиками	
Обратное правило утраивания	Правило приблизительного расчета если какое-либо событие происходит один раз в x дней то для выявления хотя бы одного случая с вероятностью 95% необходимо проводить наблюдение в течение $3x$ дней	
Обсервационное исследование	Исследование, в котором распределение больных в группы вмешательства и контроля зависит от предпочтений больного или врача	
Обусловленная вероятность события	Вероятность одного события при условии наличия другого Например, вероятность события А при наличии события В выражается как p (ASB)	
Одномоментное исследование	Обследование группы лиц в определенный момент времени или в течение определенного периода времени Влияние изучаемого фактора и клинический исход определяются одновременно	
Однофакторный регрессионный анализ(простая регрессия)	Описывается уравнением с единственной независимой переменной	Регрессия
Ожидаемая частота развития клинического исхода	Вероятность наступления изучаемого клинического исхода или достижения условия прекращения исследования	
Описание серии последовательных случаев	См Постеодовательная выборка	

Термин	Определение	См. также
Описание серии случаев	Исследование одного и того же вмешательства у отдельных последовательно включенных больных без группы контроля. Например, СОСУДИСТЫЙ хирург может описать результаты реваскуляризации сонных артерии у 100 больных с ишемией головного мозга	Последовательная выборка
Описание случаев	Вид исследования, ограниченный данными об отдельных больных	
Описательный обзор	Обзор, в ходе которого оцениваются и совместно рассматриваются все доступные отчеты (как опубликованные, так и неопубликованные) об оригинальных исследованиях, посвященных определенной теме	Систематический обзор
Оригинальные (первичные) исследования	Исследования, в которых получают первичные данные в отличие от систематических обзоров, принимаемых для совместного рассмотрения результатов нескольких оригинальных исследований	
Основная гипотеза (рабочий диагноз)	Наиболее убедительное объяснение состояния больного	
Относительные шансы	Отношение шансов развития определенного клинического исхода в основной и контрольной группах Син <i>отношение шансов</i>	
Относительный риск (ОР)	Отношение вероятности развития определенного клинического исхода у лиц, подвергшихся воздействию какого-либо фактора, к вероятности его возникновения у лиц, не подвергшихся такому) воздействию	Снижение относительного риска (COR), риск, отношение риска
Отношение правдоподобия	Применительно к результатам скринингового или диагностического теста (включая симптомы и другие клинические проявления заболевания) представляет собой отношение вероятности получения определенного результата теста у больных с изучаемым заболеванием к вероятности получения такого же результата у лиц без данного заболевания Син <i>правдоподобие</i>	
Отношение риска	Отношение вероятности развития определенного клинического исхода у лиц, подвергшихся	

Гермин	Определение	См. также
	<p>ся воздействию какого-либо фактора, к вероятности его возникновения у лиц, не подвергшихся такому воздействию</p> <p>Син.: <i>относительный риск (ОР)</i></p>	
Отношение рисков	<p>Относительный риск, рассчитанный за определенный период времени, например, при проведении анализа выживаемости. Представляет собой взвешенный показатель относительного риска за весь период наблюдения</p>	
Отношение шансов (ОШ)	<p>Отношение шансов развития клинического исхода в группе вмешательства и контрольной группе</p> <p>Син.: <i>относительные шансы</i></p>	
Оценка мнений участников исследования	<p>Обсуждение предварительных результатов исследования с его участниками для того, чтобы установить, насколько верно было представлено их мнение, не возникло ли значительного искажения фактов, одинаково ли относятся к сформулированным выводам участники с различным прогнозом</p>	
Оценка согласия мнений, рассчитанная с поправкой на случайное согласие	<p>Рассчитанная с помощью специальных методов вероятность согласованности мнений, исключающая возможность их случайного совпадения</p>	
Паллиативное лечение	<p>Паллиативное лечение включает медицинские вмешательства, проводимые при низкой вероятности излечения больного. В <i>Stedman's Medical Dictionary</i> (27-е издание) определяется как мероприятия, направленные на уменьшение выраженности симптомов без устранения заболевания. Нередко паллиативное лечение носит комплексный характер с вовлечением членов семьи и близких</p>	
Первичная медицинская помощь	<p>Медицинская помощь при первом обращении к врачу по поводу конкретного заболевания или состояния. Обычно к врачам первичного звена относят врачей общей практики, семейных врачей, терапевтов или педиатров первичного звена. Первичную помощь могут оказывать и другие медицинские работники, в частности спе-</p>	

Термин	Определение	См. также
	<p>специально обученные медицинские сестры и парамедицинский персонал. Обычно врачи общей практики, семейные врачи, специально обученные медицинские сестры и парамедицинский персонал ограничиваются только оказанием первичной помощи, однако они могут сочетать ее с элементами специализированной (квалифицированной) медицинской помощи при наличии специальной подготовки. Таким образом, медицинская помощь классифицируется по характеру контакта (первичный или по направлению), нежели по квалификации персонала</p>	
Перекрестное испытание	<p>Исследование, в котором все участники последовательно получают как изучаемое, так и контрольное вмешательство</p>	Испытание до-- после
Период вымывания	<p>В клинических испытаниях таким термином обозначается срок, который должен пройти после отмены лечения, прежде чем оно перестанет оказывать эффект</p>	
Период наблюдения	<p>Срок, в течение которого исследователь наблюдает за участниками и регистрирует клинические исходы</p>	
Плацебо (имитация вмешательства)	<p>Вмешательство, не оказывающее биологическое действие (обычно прием таблеток или капсул без биологически активных ингредиентов)</p>	
Повышение абсолютного риска (ПАР)	<p>Различие по величине абсолютного риска (выраженное в процентах или доле больных, у которых отмечен изучаемый клинический исход) между основной и контрольной группами. Обычно используется при изучении вредных воздействий</p>	Снижение абсолютного риска (САР); индекс потенциального вреда (ИПВ)
Подбор пар	<p>Метод, используемый для достижения однородности основной и контрольной групп по влиянию факторов, которые сами по себе не являются объектом изучения, но тем не менее могут повлиять на оценку полученных результатов (вмешивающиеся факторы). Например, в исследованиях случай—контроль подбор пар может осуществляться по таким показателям, как возраст, пол или другие характеристики участников</p>	

Термин	Определение	См. также
Полный период наблюдения	См. <i>Период наблюдения</i>	
Популяционное исследование	Обсервационное или описательное исследование без преднамеренного вмешательства, в ходе которого участников систематически обследуют для уточнения отсутствия или наличия (либо выраженности) изучаемого признака	
Порог информационной достаточности	Соответствует объему знаний, полученных при сборе, анализе информации и формулировании выводов, который обеспечивает создание методологически стройной системы взглядов. Дальнейшие исследования в этом направлении практически не влияют на положения выстроенной концепции	
Порядок действий при побочной реакции	Алгоритм действий, направленных на выявление или уменьшение выраженности побочных реакций	
Последовательная выборка	Формируется при последовательном включении всех больных с данным заболеванием (состоянием), обратившихся за определенный период времени в медицинские учреждения, на базе которых проводится исследование Син.: <i>Описание серии последовательных случаев</i>	
Последовательные тесты	Диагностические тесты, выполняющиеся последовательно, а не одновременно	
Послетестовая вероятность	Вероятность наличия изучаемого заболевания, рассчитанная после получения результатов диагностического теста	
Послетестовые шансы	Шансы наличия изучаемого заболевания, рассчитанные после получения результатов диагностического теста	
Постмаркетинговое (обсервационное) исследование	См. <i>Исследования</i>	
Правдоподобие	См. <i>Отношение правдоподобия</i>	
Правила клинического	Основаны на комплексном изучении факторов, определяющих вероятность того или иного ди-	

Термин	Определение	См. также
прогнозирования (принятия клинических решений)	агноза в настоящем и клинического исхода в будущем В некоторых случаях, когда эта вероятность достаточно высока или, напротив, мала, предлагается алгоритм обследования или лечения	
Практика, основанная на принципах доказательной медицины	См <i>Доказательная медицина</i>	
Предварительный отбор исследований	При поиске в медицинской литературе обеспечивает включение в обзор только исследований высокого методологического качества	
Предполагаемое изменение затрат	Прогнозируемое изменение затрат, связанное со снижением или повышением частоты развития клинических исходов в результате проводимой терапии	
Предпочтения больных	Относительная ценность, которую больные придают тому или иному состоянию здоровья	
Претестовая вероятность	Вероятность наличия изучаемого заболевания, рассчитанная до получения результатов диагностического теста	
Претестовые шансы	Шансы наличия изучаемого заболевания, рассчитанные до получения результатов диагностического теста	
Приемлемость	Учитывает предпочтения больных, оцененные при помощи методов, в основе которых лежат современные представления об этом вопросе Предпочтения больных связаны с нежеланием или нежеланием (страхом перед развитием) определенного клинического исхода Теория приемлемости основывается на положениях, характеризующих процесс принятия решения, когда его последствия неизвестны Чаше всего применяются такие методы измерения приемлемости, как стандартизация риска и временной компромисс	
Причинный фактор	Фактор, непосредственно вызывающий клинический исход	
Проверка теорий	Согласование новых данных с уже существующими теориями	

Термин	Определение	См. также
Прогноз	Возможные клинические исходы и вероятность их развития при том или ином заболевании	
Прогностическая ценность результата диагностического теста	Прогностическая ценность положительного результата диагностического исследования равна вероятности наличия заболевания при положительном результате теста Прогностическая ценность отрицательного результата диагностического исследования равна вероятности отсутствия заболевания при отрицательном результате теста	
Прогностические факторы	Характеристики больных или участников исследования, определяющие более высокую или более низкую вероятность развития благоприятных либо неблагоприятных клинических исходов	
Прогностическое исследование	Исследование, в котором наблюдение за участниками начинается одновременно и проводится с целью определения частоты и сроков развития определенных клинических исходов	
Продолжительность жизни с поправкой на ее качество	Число лет предстоящей жизни с поправкой на ожидаемое качество жизни в эти годы	
Продольное исследование	См <i>Когортное исследование</i>	
Просеивание данных (проведение множественных сравнений)	Поиск различий по частоте развития клинических исходов между группами или подгруппами больных без предварительной формулировки гипотезы	
Проспективное исследование	См <i>Когортное исследование</i>	
Профессиональные медицинские работники	Все лица, имеющие сертификацию для деятельности, относящейся к сохранению здоровья, включая врачей-неспециалистов, медицинских сестер, врачей-специалистов, физиотерапевтов, фармацевтов, специалистов по профессиональным заболеваниям и респираторной терапии, консультантов по общим вопросам (например, по здоровому образу жизни)	Медицинские работники

Термин	Определение	См. также
Профилактика	Одно либо несколько вмешательств направленных на предотвращение того или иного заболевания Первичная профилактика проводится в его отсутствие, вторичная — при наличии заболевания для предотвращения его прогрессирования и развития связанных с ним осложнений Нередко вторичная профилактика неотличима от лечения Примером первичной профилактики может служить вакцинация против коклюша, вторичной — вмешательств, проводимые у женщин с остеопорозом и переломами позвонков в анамнезе для восстановления минеральной плотности костной ткани во избежание развития повторных переломов	
Профиль здоровья	Тематические разделы анкеты, которые предназначены для использования в общей популяции (с участием здоровых, больных с различными заболеваниями и состояниями любой тяжести) и отражают важнейшие аспекты качества жизни, связанного со здоровьем	Качество жизни, связанное со здоровьем
Прямые затраты	Затраты, непосредственно связанные с производством медицинской услуги (работа врача и медицинской сестры, расходные материалы и т п)	
Пусковые события	События, при возникновении которых задействуется компьютерная система для облегчения принятия решения	Порядок действий при побочной реакции
Размер эффекта	Разница по частоте развития клинических исходов между группами вмешательства и контроля, деленное на показатель вариабельности, чаще на стандартное отклонение	
Рандомизация	Отнесение участников к той или иной группе, которое определяет случай, обычно при помощи таблицы случайных чисел Следует отличать от систематического распределения (например, по четным или нечетным дням месяца), а также от распределения по выбору исследователя	Случайный, случайная выборка, случайная ошибка
Рандомизированное испытание	Исследование диагностического теста, метода профилактики или лечения в котором участники в случайном порядке распределяются в основную и контрольную группы после чего за ними осуществляется наблюдение для определения эффекта вмешательства	Нерандомизированное испытание

Термин	Определение	См. также
--------	-------------	-----------

Син *контролируемое испытание, рандомизированное контролируемое испытание*

Рандомизированное контролируемое испытание См *Рандомизированное испытание*

Распространенность Доля лиц с тем или иным заболеванием в определенный момент времени. Распространенность, рассчитанная по данным, которые были получены в исследованиях высокого методологического качества, позволяет определить претестовую вероятность наличия изучаемого заболевания.

Реабилитация Комплекс мероприятий, которые направлены на восстановление функции, нарушенной вследствие болезни или травмы.

Регрессионная модель Кокса При регрессионном анализе позволяет учесть исходные различия между экспериментальной и контрольной группами применительно к выживаемости. Скорректированный анализ

Регрессия Статистический метод, заключающийся в построении на основе данных о независимых переменных (прогностических факторах) математической модели, позволяющей предсказать значения зависимой переменной.

Рекурсивное разбиение Метод оценки влияния предполагаемых прогностических факторов на вероятность того или иного клинического исхода. Из выборки формируют подгруппы с различными характеристиками (возраст, пол и т.д.) и сравнивают между собой для выявления различий по частоте выявления изучаемого клинического исхода.

Риск Характеристика причинно-следственной связи между вредным воздействием и клиническим исходом или изучаемым критерием оценки (включая заболеваемость, побочные эффекты, токсичность). Абсолютный риск (АР), снижение абсолютного риска (САР), относительный риск (ОР), снижение относительного риска (СОР)

РКИ с участием отдельно взятого Исследование с единственным участником, которое проводится для изучения эффекта вме-

Термин	Определение	См. также
больного (РКИ «п=1»)	шатательства или влияния воздействия у конкретного больного	
Симптом	Любое отклонение от нормальной функции, деятельности или ощущения, о котором сообщает больной. Служит субъективным свидетельством наличия заболевания	
Синдром	Совокупность признаков или симптомов, а также отклонений физиологических параметров от нормы	
Синдромный диагноз	В отсутствие стандартных диагностических методов уверенность врача в правильности поставленного диагноза существенно снижается. В таких случаях диагностические критерии основываются на наборе клинических признаков и симптомов заболевания	Синдром
Систематическая ошибка	<p>Систематическое отклонение результатов от истинных значений. Систематическая ошибка может быть обусловлена</p> <ul style="list-style-type: none"> a) назначением препаратов с учетом прогноза врачи по-разному подходят к назначению терапии при различном прогнозе. В таких случаях сравнение результатов в основной и контрольной группах приведет к неправильной оценке эффекта лечения, b) методами регистрации данных, возникает при использовании в сравниваемых группах различных подходов к регистрации данных (с помощью компьютера и вручную), c) выявлением определенного исхода, возникает при более тщательном выявлении изучаемого клинического исхода в одной из сравниваемых групп, d) использованием изучаемого диагностического метода в комплексном обследовании, результаты которого рассматриваются как «золотой стандарт» для оценки этого метода; e) более тщательным опросом участников одной из сравниваемых групп; f) предпочтительной публикацией положительных результатов, возникает, если вероятность публикации полученных данных зависит от того, выявляют ли они клиническую и статистическую значимость эффекта вмешательства, 	

Термин

Определение

См. также

g) **влиянием клинического исхода на воспоминания о воздействии**, возникает вследствие того, что участники, у которых возникли изучаемые клинические исходы, чаще вспоминают о соответствующем вредном воздействии, чем участники из контрольной группы; независимо от реальных наличия, длительности и интенсивности воздействия;

h) **влиянием ожидаемого исхода на особенности исследования**, синоним систематической ошибки, связанной с выявлением определенного исхода; возникает при более тщательном выявлении клинического исхода в одной из сравниваемых групп;

i) **подтверждением диагноза**, возникает, если результаты диагностического теста влияют на включение участника в группу вмешательства

Син.: *систематическая ошибка, связанная с порядком обследования больных*

Систематическая ошибка, связанная с более тщательным опросом участников одной из сравниваемых групп

См. *Систематическая ошибка*

Систематическая ошибка, связанная с влиянием клинического исхода на воспоминания о воздействии

См. *Систематическая ошибка*

Систематическая ошибка, связанная с влиянием ожидаемого исхода на особенности исследования

См. *Систематическая ошибка*

Систематическая ошибка, связанная с

См. *Систематическая ошибка*

Термин	Определение	См. также
<p>выявлением определенного клинического исхода</p>		
<p>Систематическая ошибка, связан- ная с использо- ванием изучаемого диагностического метода в комплексном обследовании, результаты которого рассматриваются как «золотой стандарт» для оценки этого метода</p>	См. <i>Систематическая ошибка</i>	
<p>Систематическая ошибка, связан- ная с методами регистрации данных</p>	См. <i>Систематическая ошибка</i>	
<p>Систематическая ошибка, связан- ная с назначени- ем препаратов с учетом прогноза (channeling effect or bias)</p>	См. <i>Систематическая ошибка</i>	
<p>Систематическая ошибка, связан- ная с подтвер- ждением диагноза</p>	См. <i>Систематическая ошибка</i>	
<p>Систематическая ошибка, связан- ная с порядком обследования больных</p>	См. <i>Систематическая ошибка</i>	
<p>Систематическая ошибка, связан-</p>	См. <i>Систематическая ошибка</i>	

Термин	Определение	См. также
ная с предпочтительной публикацией положительных результатов исследований		
Систематический обзор	Исследование, в ходе которого проводится критический анализ и оценка (а не просто совместное рассмотрение) результатов других исследований для того, чтобы ответить на заранее сформулированный ясный клинический вопрос при помощи методов, позволяющих свести к минимуму возможность появления систематической ошибки	Описательный обзор
Системы напоминания	Предназначены для напоминания врачам о важных мерах, которые необходимо предпринять во избежание того или иного осложнения	Системы тревоги
Системы тревоги	Используются для анализа аналоговой или цифровой информации о параметрах жизнедеятельности больного; при этом выдается сообщение (сигнал тревоги), если необходимо вмешательство врача	Системы напоминания
Скорректированный анализ	Позволяет получить результаты, в которых учтены различия по прогностическим факторам между группами. Например, если средний возраст участников в группе вмешательства, а следовательно, и риск развития побочных эффектов выше, чем в контрольной группе, при проведении скорректированного анализа будет показан больший терапевтический эффект, чем в его отсутствие	Регрессионная модель Кокса
Скрининг	Мероприятия, направленные на выявление лиц с повышенным риском развития неблагоприятных клинических исходов, которые возможно предотвратить. Проводятся в отсутствие симптомов или факторов риска развития (за исключением возраст- <i>i</i> или пола) выявляемого заболевания	Избирательный скрининг
Скрытое отнесение участников к той или иной группе	См. <i>Соккрытие порядка рандомизации</i>	

Термин	Определение	См. также
Слепой метод (ослепление, маскирование вмешательства)	При использовании простого слепого метода участник исследования не знает, к какой из сравниваемых групп он отнесен. Метод может применяться к больным; врачам; исследователям, регистрирующим или оценивающим клинические исходы, а также статистикам и авторам статьи. Во избежание неясности в исследованиях, в которых критерием оценки служит острота зрения, обычно используется термин "маскирование вмешательства"	Скрытие порядка рандомизации
Случайная выборка	Выборка, при формировании которой вероятность включения единиц наблюдения (чаще всего отдельных участников) в исследование независима, постоянна и обычно одинакова для каждой из них. Будет ли больной включен в состав участников, определяет случай, например, при помощи таблицы случайных чисел	Случайный; случайная ошибка
Случайная ошибка	Приводит к тому, что истинный размер эффекта вмешательства никогда точно неизвестен. Результаты конкретного исследования представляют собой лишь часть всех возможных значений, которые могли быть получены при изучении вмешательства в общей популяции. Таким образом, среднее значение, полученное в каком-либо исследовании, всегда несколько отличается от истинного значения в общей популяции. Чем больше величина случайной ошибки, тем меньше достоверность полученных результатов	Случайный; случайная выборка
Случайный	Термин обозначает процесс, при котором знание о произошедшем событии не позволяет предсказать его наступление в будущем. Так, вероятность распределения конкретного больного в одну из 2 групп постоянна и обычно составляет 0,5, однако заранее никогда неизвестно, в какую группу попадет тот или иной участник	Рандомизация; случайная ошибка
Смертность	Частота наступления смертельного исхода в выборке	
Снижение абсолютного риска (САР)	Различие по величине абсолютного риска (выраженное в процентах или доле больных, у которых отмечен изучаемый клинический исход)	Повышение абсолютного риска (ПАР); число

Термин	Определение	См.также
	<p>между группами вмешательства (число исходов в группе лечения, ЧИЛ) и контроля (число исходов в группе контроля, ЧИК). Используется при изучении медицинских вмешательств</p>	<p>больных, которых необходимо лечить (ЧБНЛ); риск Относительный риск (ОР); Эффект лечения</p>
<p>Снижение относительного риска (СОР)</p>	<p>Снижение исходного риска, связанное с вмешательством; рассчитывается путем деления снижения абсолютного риска (САР) на величину абсолютного риска в контрольной группе (ЧИК)</p>	
<p>Совокупность концепций</p>	<p>Система, обуславливающая взаимодействие между концепциями</p>	
<p>Скрытие результата рандомизации</p>	<p>Имеет место в случае, когда лицо, ответственное за включение больных в исследование, не знает, в какую из групп будет распределен следующий участник. При отказе от сокрытия порядка рандомизации больных с более благоприятным прогнозом могут чаще включать в группу активного лечения, что приведет к завышению положительного эффекта лечения или даже к ложному заключению о его эффективности</p> <p><i>Син.: скрытое отнесение участников к той или иной группе</i></p>	<p>Слепой метод</p>
<p>Соответствие оказываемой медицинской помощи общепринятым стандартам</p>	<p>Степень, в которой медицинские учреждения и работники (врачи, медицинские сестры и т.д.) соблюдают стандарты диагностического обследования, мониторинга, лечения, профилактики и использования ресурсов, что обеспечивает оказание высококачественной медицинской помощи</p>	
<p>Соотношение цен и затрат</p>	<p>При наличии систематических расхождений между ценами и затратами на проведение сравнимых вмешательств сопоставить их фактическую стоимость можно, определив соотношение цен и затрат</p>	
<p>Сопутствующие вмешательства</p>	<p>Дополнительные вмешательства, частота проведения которых может быть различной в группах лечения и контроля, что нередко приводит к возникновению систематической ошибки</p>	

Термин	Определение	См. также
Сопутствующие заболевания	Заболевания, имеющиеся у участников исследования помимо изучаемого	
Состояние здоровья	Характеризует физическое состояние того или иного лица или группы лиц на протяжении определенного периода времени (обычно оценивается одновременно)	
Сохраненный год жизни с поправкой на ее качество	Единица измерения продолжительности жизни, в которой учтено ухудшение состояния здоровья и связанное с ним снижение качества жизни. Например, если больная прожила 10 лет, а качество ее жизни снизилось на 50% в связи с хроническим заболеванием легких, показатель выживаемости будет эквивалентен 5 годам жизни с поправкой на ее качество	
Специализированная (квалифицированная) медицинская помощь	Медицинская помощь, которую оказывает врач с узкой специализацией, после направления к нему больного другим медицинским работником. Принято выделять два уровня специализированной (квалифицированной) медицинской помощи: вторичная (квалифицированная) и третичная (специализированная). Вторичную медицинскую помощь обычно оказывает специалист более широкого профиля, например общий хирург, терапевт или акушер-гинеколог Син.: <i>квалифицированная медицинская помощь</i>	Первичная помощь
Специфичность диагностического теста	Доля лиц без выявляемого заболевания от всех обследованных с отрицательным результатом диагностического теста (в качестве теста могут рассматриваться клинические проявления болезни)	Чувствительность диагностического теста; подтверждение предполагаемого диагноза положительным результатом высокоспецифичного диагностического теста
Стандарт для сравнения	См. <i>Стандартный диагностический (скрининговый) метод</i>	
Стандартизация риска	Способ прямого определения предпочтений больных или приемлемости клинических исходов, при котором сам больной оценивает качество собственной жизни по шкале от 0 до 1,0 (0 — смерть, 1,0 — полное здоровье), после чего	

Термин	Определение	См. также
	<p>делает гипотетический выбор между жизнью в своем нынешнем состоянии в течение x лет (мес) и стандартным риском лечения, связанного с вероятностью P (от 0 до 0,99) полного выздоровления на x лет (мес) и вероятностью 1-Я немедленной смерти</p>	
Стандартная ошибка	<p>Стандартное отклонение оценки, полученной в исследовании, от истинного значения в популяции (т.е., стандартная ошибка средней, полученной в исследовании, представляет собой стандартное отклонение рассчитанного среднего значения показателя от такового в популяции)</p>	
Стандартный диагностический метод	<p>Диагностический метод с высокой, доказанной во многих исследованиях точностью, служащий стандартом, с которым сравнивают новые диагностические (скрининговые) тесты. Стандартный метод может быть сложным и включать несколько исследований; может потребоваться наблюдение за изменением состояния больного или достижение согласия экспертов, как это нередко бывает в психиатрических исследованиях</p> <p>Син.: «золотой стандарт», стандарт для сравнения</p>	
Стандарты	<p>Условно установленные нижний допустимый, наивысший достижимый уровни либо приемлемый диапазон качества медицинской помощи или ее результативности</p>	
Статистическая значимость	<p>Считается подтвержденной при опровержении нулевой гипотезы. Иными словами, при условии истинности последней вероятность наблюдаемого различия не достигает оговоренного порогового уровня (обычно 0,05)</p>	
Статистическая мощность чувствительность)	<p>При сравнении эффективности двух вмешательств' возможность выявления различия между группами при условии, что оно существует в действительности</p>	
Статистический вывод	<p>Заключение об истинных значениях тех или иных показателей, сделанное на основе применения статистических методов, которые используются для того, чтобы 1) сделать предпо-</p>	

Термин	Определение	См. также
	<p>ложение о точечном значении показателя в популяции на основании данных, полученных в выборке, и 2) проверить статистические гипотезы</p>	
Степени свободы	<p>Технический термин, с помощью которого в биостатистике характеризуют статистическую мощность (чувствительность) анализа. Она тем больше, чем больше степеней свободы</p>	
Структура исследования	<p>Способ организации исследования лекарственных препаратов.</p> <p>a. Исследования I фазы: направлены на изучение механизма действия препарата и выявление ранних тяжелых побочных эффектов; обычно проводятся на здоровых добровольцах.</p> <p>b. Исследования II фазы: направлены на получение предварительных данных об эффективности препарата; включают больных с соответствующими заболеваниями.</p> <p>c. Исследования III фазы: рандомизированные контролируемые испытания, проводимые для точного определения размера эффекта препарата.</p> <p>d. Исследования IV фазы (постмаркетинговые наблюдательные исследования): проводятся после доказательства эффективности препарата и поступления его на фармацевтический рынок; обычно целью таких исследований служит изучение частоты развития редких побочных эффектов</p>	
Теория	<p>Совокупность и взаимосвязь концепций</p>	
Тест хи-квадрат	<p>Статистический метод, используемый для оценки распределения дихотомических клинических исходов в сравниваемых группах; нулевая гипотеза предполагает, что это распределение однородно</p>	
Точечная оценка	<p>Наиболее правдоподобная оценка истинного размера эффекта вмешательства</p>	
Требование к минимальному числу исследователей, участвующих в сборе	<p>Подразумевает участие в сборе и анализе данных двух исследователей или более с последующим согласованием мнения</p>	

Термин	Определение	См. также
информация о клинических исходах		
Третичная медицинская помощь	<i>См. Специализированная (квалифицированная) медицинская помощь</i>	
Удобная выборка	Лица или группа лиц, отобранные для участия в исследовании, исходя из удобства для исследователя или благодаря тому, что они находились в определенном месте в определенное время	Клинический исход; цель лечения
Условие прекращения исследования	Развитие изучаемого клинического исхода или определенного числа клинических исходов, при котором прекращают наблюдать участника испытания или когортного исследования (например, развитие инсульта или регистрация более высокой смертности в контрольной группе)	
Утилитарный подход к медицинским решениям	Утилитарный подход к распределительной справедливости подразумевает, что в каждом конкретном случае врач должен принимать клиническое решение, учитывая все факторы (в том числе социальные), и принятое решение должно принести наибольшую пользу наибольшему числу лиц. При таком подходе врач должен учитывать, как повлияет на других использование ресурсов для данного больного. Подобное мнение противоположно деонтологическому подходу	
Участники, у которых выявлено изучаемое заболевание	Применимо к исследованиям диагностических тестов	
Участники, у которых изучаемое заболевание не выявлено	Применимо к исследованиям диагностических тестов	
Учреждение первичной медицинской помощи	Предназначено для оказания исключительно первичной медицинской помощи; при необходимости в специализированной (квалифицированной) медицинской помощи больных на-	

Термин	Определение	См. также
	<p>правляют в другие лечебно-профилактические учреждения. Некоторые медицинские центры оказывают наряду с первичной специализированную медицинскую помощь. Таким образом, медицинская помощь классифицируется по характеру контакта (первичный или по направлению), нежели по типу учреждения</p>	
Факторы риска	<p>Нередко проводят различие между прогностическими факторами и факторами риска. Под последними подразумеваются характеристики больных, в наибольшей степени связанные с развитием заболевания</p>	
Физическое состояние	<p>Широкий термин, включающий в себя такие понятия, как наличие того или иного заболевания, расстройств, синдромов и симптомов</p>	Состояние здоровья
Фокусные группы	<p>Группы, в которые обычно включают от 4 до 8 человек со сходным состоянием или опытом для изучения их реакции на определенную ситуацию или воздействие</p>	
Функции правдоподобия	<p>Создаются на основе статистической модели и полученных в ходе исследования данных, что позволяет получить значения вероятности этих данных для различных значений неизвестных параметров модели. Значения параметров, соответствующие наибольшей вероятности, представляют собой показатели наибольшего правдоподобия этих параметров</p>	
Характеристическая кривая	<p>Метод графического представления мощности диагностического теста, при использовании которого на вертикальную ось проецируется вероятность истинноположительного результата теста (чувствительность), на горизонтальную — вероятность ложноположительного результата (1—специфичность). Чем точнее тест, тем ближе площадь под кривой к 1,0; если же результаты теста объясняются случайностью, площадь под кривой будет равна 0,5</p>	
Целевая выборка	<p>В качественных исследованиях последовательное или рандомизированное включение участников, которое обычно применяется в количественных исследованиях, часто заменяют це-</p>	

Термин	Определение	См. также
	ленаправленным отбором небольшого числа больных с определенными характеристиками	
Цель лечения	Разрешение или снижение выраженности тех или иных проявлений заболевания (симптомов, клинических признаков или отклонений физиологических параметров от нормальных значений)	Условие прекращения исследования; клинические исходы
Центр третичной медицинской помощи	Медицинский центр, куда поступают больные после оказания им первичной и вторичной (квалифицированной) медицинской помощи, и где созданы условия для проведения специализированных диагностических исследований и лечебных вмешательств, недоступных в других лечебно-профилактических учреждениях. Как правило, в таких центрах оказывают широкий спектр медицинских услуг, включая элементы первичной и вторичной медицинской помощи, поэтому данный термин отражает скорее возможности учреждения, а не уровень медицинской помощи в рамках конкретного исследования	Специализированная (квалифицированная) медицинская помощь; первичная медицинская помощь
Циклы наблюдения	Обычно используются в обсервационных исследованиях, что позволяет изучить изменение частоты развития клинических исходов на протяжении нескольких отрезков времени и определить, связано ли это изменение с вмешательством или воздействием	
Частота развития клинического исхода	Доля больных, у которых выявлен изучаемый клинический исход. Принято разделять частоту развития клинического исхода в группе лечения (ЧИЛ) и группе контроля (ЧИК) Син.: <i>число исходов в группе лечения (ЧИЛ)</i>	Эффект лечения; Исходный риск
Число больных, которых необходимо лечить (ЧБНЛ)	Число больных, которым необходимо проводить данное вмешательство в течение определенного периода времени для того, чтобы предотвратить один неблагоприятный клинический исход. При обсуждении данного показателя необходимо указывать тип вмешательства, его длительность и неблагоприятный клинический исход, который удалось предотвратить. Обрато пропорционален снижению абсолютного риска (САР)	Снижение абсолютного риска (САР)

Термин	Определение	См. также
Число исходов в группе контроля (ЧИК)	См. <i>Частота развития клинического исхода; исходный риск</i>	
Число исходов в группе лечения (ЧИЛ)	См. <i>Частота развития клинического исхода</i>	
Число сохраненных лет жизни с учетом трудоспособности	Продолжительность жизни с поправкой на нетрудоспособность	
Чувствительность диагностического теста	Доля лиц с выявляемым заболеванием от всех обследованных с положительным результатом диагностического теста (в качестве теста могут рассматриваться клинические проявления болезни)	Анализ чувствительности; специфичность диагностического теста, исключение предполагаемого диагноза при получении отрицательного результата высокочувствительного диагностического теста
Шансы	Отношение вероятности развития события к вероятности того, что оно не наступит	
Шкалы Ликерта	Шкалы, обычно состоящие из 3–9 величин, включающих крайние степени отношения или ощущения (например, от "совершенно не согласен" до "полностью согласен"). Используются при изучении взглядов участников исследования на ту или иную проблему	Визуальные аналоговые шкалы
Экономическая оценка	Сравнительный анализ альтернативных методов профилактики или лечения с учетом затрат и влияния на клинические исходы	
Экономический анализ	Совокупность количественных методов, используемых при сравнении двух или более вмешательств либо программ с точки зрения необходимых ресурсов и ожидаемого влияния на клинические исходы	
Экспериментальное лечение	Вмешательство (нередко новое или видоизмененное), которое сравнивают в исследовании	

Гермин	Определение	См. также
	с существующим стандартом или отсутствием активного лечения	
Эффект контрольного списка	Влияние, которое оказывает на стиль работы врача необходимость регистрации данных или действий с использованием специальных структурированных форм	
Эффект лечения	Результаты сравнительных клинических исследований могут быть представлены в виде различных показателей, отражающих эффект лечения, например снижения абсолютного риска (САР), снижения относительного риска (СОР); отношения шансов (ОШ); числа больных, которых необходимо лечить (ЧБНЛ), и размера эффекта. Целесообразность использования того или иного показателя и то, будут ли при этом рассчитывать вероятность, средние или медианы, зависит от типа изучаемого клинического исхода. Так, САР, СОР и ЧБНЛ применяются при изучении дихотомических клинических исходов, а размер эффекта — при анализе непрерывных данных	Снижение абсолютного риска (САР); снижение относительного риска (СОР), отношение шансов (ОШ); число больных, которых необходимо лечить (ЧБНЛ)
Эффект газдирателя	Улучшение результата какой-либо деятельности в условиях, когда выполняющий ее знает, что за ним осуществляется наблюдение	
Эффект •братной связи	Влияние проводимой в лечебном учреждении оценки качества медицинских услуг на стиль практики	
Эффект плацебо	Эффект вмешательства, не зависящий от его биологического действия	
Эффект тоннеля	Психологический феномен, который объясняет необходимость учета при проведении экономического анализа более узких взглядов на проблемы, связанные с дополнительными расходами при изменении подходов к лечению. Такие расходы требуют изменения финансирования, решения о котором принимают ответственные лица или организации	
>эффект Хоторна	Улучшение результата какой-либо деятельности, когда выполняющий ее знает, что за ним осуществляется наблюдение	

Термин	Определение	См. также
Эффективность	Техническая эффективность есть соотношение между затратами на проведение вмешательства и его результатом (увеличение продолжительности жизни с поправкой на ее качество) Вмешательство, обеспечивающее более значительное увеличение продолжительности жизни с поправкой на ее качество при аналогичных или меньших затратах, более эффективно. При анализе технической эффективности применяют анализ минимизации, эффективности и полезности затрат Концепция распределительной эффективности подразумевает, что охрана здоровья не является единственной целью общества, поэтому различные задачи лечения должны быть соотнесены между собой и с затратами Обычно для этого используется анализ выгоды затрат	
Эффекты, общие для препаратов одного фармакологического класса	Аналогичные эффекты, свойственные всем или почти всем препаратам какого-либо фармакологического класса (например, (3-блокаторам, антагонистам кальция, ингибиторам ангиотензинпревращающего фермента)	Чувствительность диагностического теста
SnNout	Исключение предполагаемого диагноза при получении отрицательного результата высокочувствительного диагностического теста	Специфичность диагностического теста
SpPin	Подтверждение предполагаемого диагноза положительным результатом высокоспецифичного диагностического теста	

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Анализ в подгруппах
 - критерии 227, 228
 - недостатки 95, 215—217
- Анализ выживаемости 322—325
- Анализ данных с внесением поправок на определенные исходные факторы 293
- Анализ исходя из допущения, что все больные получили предписанное лечение 80
- Анализ принятия клинических решений
 - определение вероятностей 244
 - применение 242
 - структура дерева решений 242, 243
 - сумма вероятностей 235, 236
- Анализ принятия решений
 - и анализ принятия клинических решений 242
 - и дерево решений 242, 243
 - определение 242
- Анализ чувствительности
 - и предположения о вероятности клинических исходов 260
 - однофакторный по сравнению с многофакторным 259, 260
 - определение 225
- База данных Best Evidence
 - описание 39, 40
 - стратегия использования 40—45
- База данных Medline
 - описание 38, 39
 - стратегия использования 53—55
- База данных Up-to-date
 - описание 50
 - стратегия использования 51, 52
- Вероятность
 - и гипотеза 288—291
 - и эффект лечения 275
 - определение 275, 288
- Вероятность наличия искомого заболевания
 - единицы измерения 157, 158
 - значимость 154, 155
- Вмешивающиеся факторы
 - влияние 114—116
 - неизвестные 117
 - определение 278
- Выбор тактики, профилактики и лечения 252, 253
- Выбытие участников из исследования 83—85, 120, 121
- Гипотеза
 - ограничения при проверке 296
 - проверка 287—291
- Градиент зависимости эффекта от дозы 124

- Дерево решений
и вероятности 244
структура 242—244
- Диагноз 136—138
- Дифференциальная диагностика
диагностические подходы 138, 139
критерии 151—153
систематические исследования 141
- Доверительные интервалы
и исследования с отрицательным результатом 303—305, 307
и исследования с положительным результатом 305, 307
и кривые дожития 202, 203
и нижнее пороговое значение вероятности наличия заболевания, превышение которого указывает на необходимость дальнейшего обследования 155, 156
и относительный риск (ОР) 124, 125
и размер выборки 89, 90, 301, 302
и систематические обзоры 226
и снижение относительного риска (COP) 87—89
интерпретация границ 90—93
определение 98, 99, 301, 302
применение 98, 99
- Доказательная медицина
и формирование политики здравоохранения 15, 16
иерархия доказательности данных литературы 14, 15
определение 16
- Доказательные (научно обоснованные) данные
иерархия доказательности 8—11
определение 8
- Достоверность
и диагностические тесты 165—169
и дифференциальная диагностика 148—154
и лечение 75—87
и обобщение доказательных данных 213—222
и побочные эффекты 108—123
и прогноз 195—200
и рекомендации по лечению 247
определение 276
- Жизненные ценности больных
и принятие клинических решений 6—8
- Значение p 288—290
- Исследуемый клинический исход 199
- Индекс потенциального вреда (ИПВ) 126, 127
расчет 320—322
- Исследование с отрицательным результатом
и доверительные интервалы 303—305
определение 299, 300
- Исследование с положительным результатом
и доверительные интервалы 307, 308
- Исследование случай—контроль
и отношение шансов 123, 124

- недостатки 117
- определение 116
- исследования (испытания) эквивалентности вмешательств 291, 292
- сходный риск 314
- клинические вопросы
 - виды 28—32
 - диагностический процесс 136—138
 - и диагностические тесты 164
 - и дифференциальный диагноз 148
 - лечение 74, 75
 - обобщение доказательств 209—211
 - общие и частные 19
 - от доказательств к практике 236, 237
 - побочные эффекты вмешательства 108, 109
 - прогноз 192—194
 - соответствие целям поиска 213—215
 - формулирование 22—28
- клинические рекомендации см *Рекомендации по лечению*
- клиническое руководство 247
- огорточное исследование 113—116
- экранный библиотечный
 - определение 43
 - стратегия использования 44—49
- эвиденциальные критерии оценки 95—97
- кривая Каштана—Мейера см. *Кривые дожития*
- кривые дожития 200—202
- эжное заключение об одинаковой эффективности методов
 - и статистическая мощность исследования 292
 - определение 291
- эжное заключение о разной эффективности методов
 - определение 288
- эвиденциальная литература
 - изучение 21, 22
 - источники 33—35
 - критическая оценка 66, 67
 - несистематические исследования 210
 - оригинальные (первичные) исследования 210
 - оценка источников 35—37
 - оценка методологического качества 219, 220
 - поиск в ней 218, 219
 - при поиске без ограничений 53—61
 - при поиске предварительно оцененных данных 39—53
- эта-анализ
 - и систематическая ошибка 281
 - определение 210
- эдели переходных состояний 247
- эноственные сравнения
 - недостатки 294, 295
 - определение 293, 294
 - статистические подходы 296

Непрерывная переменная 292, 293

Нижнее пороговое значение вероятности наличия заболевания, превышение которого указывает на необходимость дальнейшего обследования
выбор значения 142
и верхнее пороговое значение вероятности наличия заболевания, превышение которого указывает на необходимость немедленного начала лечения 139—141
и доверительные интервалы 155, 156

Нулевая гипотеза 288—291

Обсервационные исследования 325
виды 112, 113
и предпочтения больных 12, 13
и прогноз 194, 195
и систематический обзор 219, 220
недостатки 68

Однородность групп сравнения по прогностическим факторам 80—83

Описание клинического случая 118

Описание серии случаев 118

Основные и побочные эффекты лечения
определение 85
уменьшение вероятности возникновения систематической ошибки 283, 284

Отнесение больных к той или иной группе
и анализ в группах 79, 80
и слепой метод 82, 83

Относительные шансы 315, 316

Относительный риск (ОР)
и доверительные интервалы 123, 124
и когортные исследования 123, 124
и снижение относительного риска 317, 318
определение 314

Отношение правдоподобия
номограмма для расчета 173
определение 169, 170
применение 172

Отношение рисков см. *Относительный риск*

Отношение шансов (ОШ)
и исследования случай—контроль 123
определение 315
статистические преимущества 316, 317

Параллельное (сопутствующее) вмешательство 278

Период наблюдения
длительность 153, 154

Послетестовая вероятность 173, 174, 181

Правдоподобие
и изменение прогноза с течением времени 200, 201
и оценка прогноза 202, 203

Правила клинического прогнозирования 184

Правила принятия решений см. *Правила клинического прогнозирования*

Практическое руководство см. *Клиническое руководство*

Предотвращение вредного воздействия 127, 128

Претестовая вероятность 137, 138, 140

- Лрменимость результатов исследовании
 - и изменения в распространенности заболевания 157
 - и обобщение доказательств 227—230
 - посвященных диагностическим методам 180—187
 - посвященных дифференциальной диагностике 156—158
 - посвященных побочным эффектам вмешательств 125—127
 - посвященных прогнозу 204, 205
 - посвященных лечению 93 — 102
- Принятие клинических решений
 - и побочные эффекты вмешательства 127, 128
 - и предпочтения больных 6, 7
 - процесс 12
 - Принятие решения см *Принятие к тнических решений*
- Прогностические факторы
 - коррекция с их учетом 117, 118
 - определение 192, 193
 - и однородность групп сравнения 195, 196
 - факторы риска 193
- Проспективное когортное исследование см *Обсервационное исследование*
- Размер выборки
 - и доверительные интервалы 301, 302
 - и случайная ошибка 279, 280
 - планирование 275,276
- Размер эффекта 225
- Рандомизация
 - и обсервационные исследования 78, 79
 - и прогностические факторы 79
 - и скрытое отнесение участников к той или иной группе 78, 79
 - определение 77, 78
- Рандомизированные контролируемые испытания (РКИ)
 - и исследования с отрицательным результатом 299, 300
 - методов лечения 67, 68
 - преимущества и недостатки 111 — 113
- Регрессионная модель Кокса 323
- Рекомендации по лечению
 - доказательные данные по сравнению с предпочтениями 239, 240
 - иерархия 255—259
 - принятие клинических решении 238
 - разработка 251—254
 - рассмотрение различий между ними 239
- Репрезентативная выборка 149 150
- РКИ с участием отдельно взятого больного (РКИ "п=1") 93, 94
- Систематическая ошибка
 - виды 119, 120, 168, 169,218,219
 - выявление 119
 - и выбор критериев оценки 276—279
 - и нерепрезентативная выборка 195 196
 - и отбор исследований для включения в систематический обзор 217
 - и различия по прогностическим факторам 277
 - и эффект плацебо 278

обустроенная

- более тщательным опросом участников одной из сравниваемых групп 119
- выбыванием участников из исследования 2⁹
- назначением препаратов с учетом прогноза 119

определение 276

связанная с

- влиянием клинического исхода на воспоминания о воздействии 119
- влиянием ожидаемого исхода на особенности исследования 120
- предпочтительной публикацией положительных результатов исследований 218,219
- с порядком обследования больных 168, 169

Систематический обзор

- и доверительные интервалы 226
- и количественная оценка 224—226
- и обзорные исследования 219, 220
- и оригинальные статьи 219, 220
- определение 210
- по сравнению с мета-анализом 210
- по сравнению с несистематическим обзором 210
- проведение 210—212

Слепой метод

- в исследованиях, посвященных лечению 82, 83
- и отнесение участников к той или иной группе 82
- в исследованиях, посвященных побочным эффектам вмешательств 283, 284
- и статистический анализ 83
- при сравнении точности изучаемого диагностического теста и "золотого стандарта" 167, 168

Случайная ошибка

- и систематическая ошибка 279, 280
- определение 275

Случайные факторы 287, 288

Снижение абсолютного риска (CAR)

- в группах риска 314
- определение 314
- по сравнению с числом больных, которых необходимо лечить (ЧБНЛ) 317—320
- по сравнению со снижением относительного риска (COP) 326, 327

Снижение относительного риска (COP)

- и доверительные интервалы 124
- и относительный риск (OR) 317, 318
- и снижение абсолютного риска (CAR) 326, 327
- и число больных, которых необходимо лечить (ЧБНЛ) 317—320
- определение 314

Специфичность 176

Справочник Clinical Evidence

- описание 51, 52
- стратегия использования 52—54

Стандартный диагностический метод 167, 168

Статистическая значимость 290

Структура исследования 75, 76, 194, 195

Суррогатные клинические исходы см *Косвенные критерии оценки*

Тест на гетерогенность (неоднородность) 223

Точечная оценка

- и доверительные интервалы 88—91
- и истинный эффект лечения 223
- определение 88

Хи-квадрат 291

Число больных, которых необходимо лечить (ЧБНЛ)

- и снижение абсолютного риска (САР) 318—320
- и снижение относительного риска 318—320
- определение 97
- расчет 97—99

Число исходов в контрольной группе см. *Исходный риск*

Чувствительность 175, 176

Эффект лечения

- вероятность 275
- гомогенность (однородность) и гетерогенность (неоднородность) 222, 223
- определение риска 87, 88
- точность оценки 88—91

Эффект плацебо

- и систематическая ошибка 278, 282—284
- определение 278