

Издательство Медиа Сфера

Общество специалистов
доказательной медицины

ФГБУ «Поликлиника №1»
Управления делами Президента РФ

«Доказательная гастроэнтерология» — научно-практический рецензируемый медицинский журнал.

Выходит 4 раза в год.
Основан в 2011 году.

«*Dokazatel'naâ gastroenterologiâ*» (Evidence-based Gastroenterology) is a quarterly peer-reviewed medical journal published by MEDIA SPHERA Publishing Group. Founded in 2011.

Журнал представлен в следующих международных базах данных и информационно-справочных изданиях: РИНЦ (Российский индекс научного цитирования), Ulrich's Periodicals Directory, Google Scholar.

Издательство Медиа Сфера:

127238 Москва,
Дмитровское ш., д. 46, корп. 2, этаж 4
Тел.: (495) 482-4329
Факс: (495) 482-4312
Отдел рекламы:
Тел.: (495) 482-0604
Отдел подписки:
Тел.: (495) 482-5336
Факс: (495) 482-4312
E-mail: mediasph@mediasphera.ru
www.mediasphera.ru
Адрес для корреспонденции:
127238 Москва, а/я 54, Медиа Сфера

Адрес редакции:

127238 Москва,
Дмитровское ш., д. 46, корп. 2, этаж 4
Зав. редакцией О.Д. Изотова
Тел.: (495) 482-4329
E-mail: gastro@mediasphera.ru

Оригинал-макет изготовлен
Издательством Медиа Сфера
Компьютерный набор и верстка:
Г.В. Кременчуцкая, М.Л. Калужнин
Корректоры: А.К. Балихина,
Г.И. Федоровская

Индексы по каталогу агентства «Роспечать»
84557 — для индивидуальных подписчиков
84558 — для предприятий и организаций

Формат 60×90 1/8; тираж 1500 экз.
Усл. печ. л. 000
Заказ

Отпечатано в ООО «Типография Мосполиграф»

ДОКАЗАТЕЛЬНАЯ ГАСТРОЭНТЕРОЛОГИЯ

Том 1

2.2012

НАУЧНО - ПРАКТИЧЕСКИЙ ЖУРНАЛ

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Главный редактор Е.Л. Никонов
Ответственный секретарь К.В. Пюрвеева

А.Г. Автандилов
А.А. Будзинский
С.А. Булгаков
В.В. Власов
Е.С. Вьючнова
Н.И. Громова
К.Г. Гуревич
И.А. Егорова
И.О. Иваников
С.В. Кашин
Л.В. Кудрявцева
И.В. Маев
И.А. Морозов
А.М. Нечипай
Л.М. Рошаль
В.Г. Румянцев
А.В. Фарбер
А.В. Чжао
А.П. Эттингер
Н.Д. Юшук
Э.П. Яковенко

Редакция не несет ответственности за содержание рекламных материалов. Точка зрения авторов может не совпадать с мнением редакции. К публикации принимаются только статьи, подготовленные в соответствии с правилами для авторов. Направляя статью в редакцию, авторы принимают условия договора публичной оферты. С правилами для авторов и договором публичной оферты можно ознакомиться на сайте: www.mediasphera.ru. Полное или частичное воспроизведение материалов, опубликованных в журнале, допускается только с письменного разрешения издателя — издательства «Медиа Сфера».

ОРИГИНАЛЬНЫЕ СТАТЬИ

Громова Н.И.
Сравнительная эффективность противовирусной терапии больных хроническим гепатитом С пегилированными интерферонами альфа-2а и альфа-2b в сочетании с рибавирином

НАУЧНЫЕ ОБЗОРЫ

Кашин С.В., Зеевальд Ш., Иваников И.О., Никонов Е.Л.
Пищевод Баррета: современные возможности диагностики, терапии и снижения риска развития рака

Danese S., Colombel J.-F., Reinisch W., Rutgeerts P.J.
Изменение стратегий терапии после 10-летнего клинического применения инфликсимаба в лечении болезни Крона

КЛИНИЧЕСКИЕ РУКОВОДСТВА И РЕКОМЕНДАЦИИ

Malfertheiner P., Megraud F., O'Morain C.A., Atherton J., Axon A.T.R., Bazzoli F., Gensini G.F., Gisbert J.P., Graham D.Y., Rokkas T., El-Omar E.M., Kuipers E.J.
Рекомендации по лечению инфекции *Helicobacter pylori* (Маастрихт—IV/Флоренция)

ДИСКУССИЯ

Морозов И.А., Никонов Е.Л., Миллер Д.А.
Структурно-функциональное обоснование оптимальных условий проведения эрадикации инфекции *Helicobacter pylori* в желудке

В ПОМОЩЬ ГАСТРОЭНТЕРОЛОГУ

Фарбер А.В., Никонов Е.Л.
Лечебно-диагностический алгоритм при гастроэзофагеальной рефлюксной болезни

Корнеева О.Н., Драпкина О.М.
Профилактика желудочно-кишечных осложнений при приеме антитромбоцитарных препаратов

КЛИНИЧЕСКИЙ СЛУЧАЙ

Белятко Е.А., Шифрин О.С.
Инфликсимаб в лечении неспецифического язвенного колита: случаи клинического течения заболевания после родов

ОБРАЗОВАТЕЛЬНЫЙ ПРАКТИКУМ

Кашин С.В.
Современные аспекты выполнения колоноскопии с целью скрининга полипов и колоректального рака

Власов В.В.
Почему гастроэнтерология должна быть доказательной

ORIGINAL PAPERS

3 *Gromova N.I.*
The comparative efficacy of antiviral therapy in the patients presenting with chronic hepatitis C using pegylated interferon alpha-2a and alpha-2b in combination with ribavirin

SCIENTIFIC REVIEWS

9 *Kashin S.V., Zeewald Sh., Ivanikov I.O., Nikonov E.L.*
Barrett's oesophagus: the modern approaches to diagnostics, therapy, and reduction of the risk of cancer development

22 *Danese S., Colombel J.-F., Reinisch W., Rutgeerts P.J.*
Review article: infliximab for Crohn's disease treatment — shifting therapeutic strategies after 10 years of clinical experience

CLINICAL MANUALS AND RECOMMENDATIONS

35 *Malfertheiner P., Megraud F., O'Morain C.A., Atherton J., Axon A.T.R., Bazzoli F., Gensini G.F., Gisbert J.P., Graham D.Y., Rokkas T., El-Omar E.M., Kuipers E.J.*
Management of *Helicobacter pylori* infection (the Maastricht—IV/ Florence Consensus Report)

DISCUSSION

63 *Morozov I.A., Nikonov E.L., Miller D.A.*
The structural and functional rationale for the optimal conditions of eradication of *Helicobacter pylori* infection in the stomach

GUIDELINES FOR PRACTITIONERS

68 *Farber A.V., Nikonov E.L.*
The algorithm of therapeutic and diagnostic measures for gastroesophageal reflux disease

77 *Korneeva O.N., Drapkina O.M.*
Prophylaxis of gastrointestinal complications associated with the intake of antiplatelet agents

CASE REPORT

83 *Belyatko E.A., Shifrin O.S.*
The use of infliximab for the treatment of non-specific ulcerative colitis: the cases of postpartum clinical course of the disease

TRAINING WORKSHOP

89 *Kashin S.V.*
Current aspects of practical colonoscopy for the screening of polyps and colorectal cancer

101 *Vlasov V.V.*
Why evidence-based gastroenterology

Почему гастроэнтерология должна быть доказательной

Д.м.н., проф. В.В. ВЛАСОВ

Общество специалистов доказательной медицины, Первый Московский государственный медицинский университет им. И.М. Сеченова

Why evidence-based gastroenterology

V.V. VLASOV

Society of Evidence-Based Medicine Specialists; I.M. Sechenov First Moscow State Medical University

Медицина, основанная на доказательных данных (Evidence-Based Medicine, или доказательная медицина, ДМ), — современная форма научно обоснованной медицинской практики. Противоположностью ДМ может быть все, что угодно, от шаманства до «лечения больного, а не болезни» (т.е. произвольная индивидуализация без научных оснований). Настоящий очерк подготовлен для краткого ознакомления с подходами ДМ в интересах читателей нового журнала «Доказательная гастроэнтерология».

Отличие ДМ состоит в том, что результаты научных исследований признаются «неравными», имеющими разную доказательную силу для врача. До середины XX века медицинские вмешательства обосновывали результатами самых разных исследований — от физиологических и анатомических до простых описаний случаев. Постепенно, с 1950-х, складывается дифференцированное отношение к разным научным сведениям. Признается, что одни виды исследований предоставляют слабые основания для вмешательств, а другие — достаточно убедительные. Объяснение развития язвенной болезни эмоциональными переживаниями может выглядеть правдоподобно, но испытания отвергают терапию сном или седативными препаратами. Таким образом, врачу предстоит различать существующие основания для вмешательства и убедиться в доказанности его полезности. Такая критическая оценка научных публикаций составляет основу ДМ. Важно, что ДМ — это не лечение больных на основании результатов рандомизированных контролируемых испытаний (РКИ). Врач, придерживающийся ДМ, использует все имеющиеся научные данные, но отдает предпочтение результатам наиболее доказательных — РКИ.

Необходимо сформулировать актуальный для клиники вопрос таким образом, чтобы на него можно было найти ответ, имеющий практический смысл. Студенты обычно спрашивают, почему, например, упоминание о лимоне вызывает слювацию, а упоминание о перце — нет. Практикующие же врачи обычно задаются вопросами, напрямую выводящими к практическому применению: помо-

гает ли хинин против ночных судорог в ногах? Есть ли преимущества у ингибиторов АПФ перед диуретиками в предотвращении осложнений и смерти при артериальной гипертензии?

Так, от правильно сформулированного вопроса к поиску результатов исследований, к их критической оценке и обобщению, формулированию вывода применительно к конкретной клинической ситуации определяется круг действий врача, практикующего ДМ.

Поиск информации об исследованиях

Для того, чтобы правильно оценить научные сведения об эффективности интересующего нас вмешательства, например, назначения препаратов висмута при диспепсии, нужно критически оценить ВСЮ совокупность имеющихся результатов исследований, которые всегда в большей или меньшей мере различаются, и ознакомление только с некоторыми из них может дать ложное представление об эффективности вмешательства. Соответственно, от врача требуется уметь найти все или почти все сообщения об исследованиях, отвечающих на возникший вопрос.

Применительно к деятельности современного специалиста мы говорим о поиске информации в компьютерных базах данных. Бумажные источники требуют слишком больших затрат времени. Богатства научной медицинской литературы стали по-настоящему полезны для медицинской практики только в последние 20 лет — благодаря появлению Интернета и доступности системы MEDLINE.

MEDLINE — это электронная версия указателя периодических медицинских изданий, подготавливаемого Национальной медицинской библиотекой США (NLM) для бесплатного использования во всем мире. Электронный поиск позволяет за секунды найти то, что потребовало бы нескольких дней работы в библиотеке. Интернет сделал MEDLINE доступной в любом месте, где есть телефон, компьютер или даже современный сотовый телефон.

Для того чтобы в MEDLINE (<http://www.pubmed.com>) найти статьи, соответствующие изучаемому вопросу, и не потеряться в тысячах неподхо-

Таблица 1

Изучаемый метод	Референтный тест		
	здоровые	больные	всего
Здоровые [Negative]	(TN) A	B (FN)	A+B
Больные [Positive]	(FP) C	D (TP)	C+D
Всего	A+C	B+D	N

дящих записей, нужно знать основные правила составления запроса.

1. Название состояния (заболевания) и его синонимы. В указателе ключевых слов MEDLINE — MeSH надо найти рубрику, к которой отнесено это состояние. Все названия и рубрики объединяются оператором OR (или), чтобы найти максимальное число статей, относящихся к изучаемому состоянию.

2. Аспект проблемы: лечение (treatment), диагноз (diagnosis) или прогноз (prognosis).

3. Вмешательство: в случае, если нас интересует конкретное вмешательство, например ультразвук или название лекарства.

4. Метод исследования. От этого зависит доказательность результата. Метод исследования выясняется с помощью рубрик MeSH, описывающих тип публикации, например применительно к испытанию лечебных вмешательств наибольший интерес для врача представляют РКИ (randomized controlled trials). На метод исследования могут указывать и слова в названии или реферате статьи, например «ослепление» (blind). Рубрики MeSH и слова объединяются оператором OR.

Эти четыре компонента запроса в MEDLINE объединяются оператором AND, чтобы получить только описания статей, соответствующих всем четырем критериям. Правильный запрос позволяет найти большую часть нужных статей, содержащихся в MEDLINE, и одновременно исключить множество ненужных статей. Научиться составлять запросы в MEDLINE не очень сложно: существуют средства помощи на самом сайте и литература, в которой процесс детально описан¹. Для срочного поиска на сайте NLM есть упрощенный вариант «клинические запросы» (clinical queries), где достаточно лишь вписать название состояния и отметить интересующий аспект проблемы (лечение, диагноз, прогноз). Методический фильтр, т.е. выбор доказательного типа исследований, будет применен автоматически. Вы можете подстроить его, выбрав высокую чувствительность (будет найдено больше вероятно релевантных, т.е. соответствующих запросу, статей) или высокую специфичность (будет найдено меньше статей, но они будут более точно соответствовать запросу).

¹Гринхальх Т. Доказательная медицина. ГЭОТАР 2005; Саккет Д. и др. Как учить и преподавать доказательную медицину. ГЭОТАР 2010.

Критическая оценка научного сообщения

В реферате, если он написан по современным правилам, т.е. структурирован и отвечает на основные вопросы о содержании статьи, можно найти достаточно информации, чтобы приблизительно оценить статью. Для полной оценки нужно читать раздел «Методы исследования». Критерии оценки качества исследования различны для исследований разных проблем. Если статья привлекла ваше внимание, то оцените ее с использованием основных критериев. Несоответствие им обесценивает статью.

Диагностика

Определить полезность диагностического теста было бы лучше всего, изучив его влияние на исходы лечения. Для этого нужно провести РКИ, в котором группа вмешательства получает лечение на основе новой диагностики, а контрольная группа — на основе старой. Лучшие исходы в группе новой диагностики обосновывают ее преимущества. Однако такие РКИ осуществлять сложно, поэтому лишь в редких случаях удастся найти сведения, позволяющие судить об эффективности диагностики по показателю качества лечения.

Сопоставление с «золотым стандартом»

Обычно изучают диагностическую точность метода, т.е. степень совпадения его результатов с результатами лучшего из существующих способов (нередко неприемлемого из-за его особенностей), например с результатами интраоперационного или патологоанатомического диагноза. Такой самый надежный метод называют «золотым стандартом» (референтным тестом, РТ). Новый метод обычно связан с прижизненной, более приемлемой диагностикой. Задача изучения диагностического теста (ДТ) — оценить, насколько хорошо диагноз, поставленный с помощью нового теста, соответствует «золотому стандарту», т.е. насколько он точен.

Традиционный метод оценки диагностической точности — построение «латинского квадрата» (четырехпольной таблицы). В статье должна быть приведена такая таблица (табл. 1), или сведения должны быть достаточны для того, чтобы ее построить. В противном случае следует отбросить статью: ее чтение не позволит узнать, полезен ли рассматриваемый ДТ.

Сведенные в латинский квадрат результаты позволяют вычислить так называемые операционные характеристики теста:

Чувствительность (Se) = $D/(B+D)$.

Специфичность (Sp) = $A/(A+C)$.

Прогностичность положительного результата (PVP) = $D/(C+D)$.

Прогностичность отрицательного результата (PVN) = $A/(A+B)$.

Операционные характеристики теста

Чувствительность и специфичность называют стабильными характеристиками ДТ. Это означает, что они относительно независимы от преваленса (частоты больных в общей группе пациентов). Если чувствительность составляет 0,8 (80%), то, сколько бы пациентов ни были больны этой болезнью, $\frac{4}{5}$ из больных будут выявлены с помощью ДТ. Если специфичность составляет 0,9, то среди лиц, не имеющих данного заболевания, положительный результат (ложноположительный, FP) будет получен лишь у 10%.

Если тест характеризуется высокой чувствительностью, то по его отрицательному результату можно с уверенностью исключить подозреваемую болезнь. При обследовании таким ДТ группы людей врач не пропустит в ней больных².

Если тест отличается высокой специфичностью, то его положительный результат (у здоровых лиц его почти никогда не бывает!) дает основания включить подозреваемую болезнь в дальнейшую дифференциальную диагностику.

Для ДТ не существует минимально необходимой величины чувствительности или специфичности. Тест, дающий положительный результат у больных чаще, чем у здоровых, может быть полезен. Решение об использовании ДТ зависит от многих факторов, лежащих за пределами оценки собственно диагностической эффективности ДТ. Например, при вероятности занесения опасной инфекционной болезни даже неспецифические проявления, неэффективные с точки зрения обычной диагностики, могут использоваться для выявления и изоляции возможно больных людей.

При выполнении ДТ врача в основном интересует вероятность болезни у лиц с положительным результатом теста и у лиц с отрицательным результатом, т.е. «прогностичность» положительных и отрицательных результатов. Этими показателями нельзя ограничиваться в оценке ДТ, поскольку их величина зависит от преваленса. Например при чувствительности и специфичности 90% (это высокие показатели) тестирование в условиях низкого преваленса (больных только 10%) даст $PVP=50\%$ и $PVN=99\%$, в условиях высокого преваленса (80%) — $PVP=97\%$ и $PVN=69\%$. Таким образом, даже если автор статьи делает упор на высокую надежность

²Нередко по ошибке считают, что чувствительный тест подходит для включения болезни в дальнейшую диагностическую проработку, для выявления больных.

окончательного диагноза (прогностичности положительного результата), надо учитывать чувствительность и специфичность. Многообещающий результат может оказаться простым следствием того, что в изученной группе очень высок преваленс, например вследствие предварительного отбора пациентов.

Если известен преваленс и установлены чувствительность и специфичность ДТ, то можно количественно оценить вероятность болезни при применении теста. На таких вычислениях основаны системы компьютерной диагностики.

Для каждого показателя должны быть вычислены доверительные интервалы (ДИ, confidence interval). ДИ — это интервал, в котором с заданной вероятностью содержится истинная величина. Обычно медицинские исследования проводятся на небольших группах больных. Полученные результаты используются для того, чтобы судить обо всех пациентах с таким заболеванием (о популяции). Оценка, полученная в выборке, отражает ситуацию в популяции, но приблизительно. Чаще всего для описания степени приближения используется 95% ДИ. Он с вероятностью 95% содержит истинную (популяционную) величину, например чувствительности диагностического теста. Чем больше выборка пациентов, тем точнее оценка. При обследовании малочисленной группы оценка чувствительности 70% очень неопределенна, например с 95% вероятностью от 40 до 98%. Естественно, нельзя уверенно судить о диагностической точности теста, если его чувствительность может быть и 98%, и 40%.

Слепая оценка результатов

В исследовании ДТ оценка результатов применения теста должна быть проведена «вслепую» — так, чтобы специалист не знал результатов других анализов у данного пациента, прежде всего, результатов референтного теста (РТ). Результаты РТ должны оцениваться тоже вслепую, независимо от результатов изучаемого теста. В противном случае возникает ошибка типа «смещения к предполагаемому диагнозу». Они возникают обычно произвольно, например вследствие того, что знающий о предположительном диагнозе специалист склонен более подозрительно относиться к оценке результатов по рентгенограмме (ЭКГ).

Вслепую должна быть определена надежность самой оценки результатов теста. Это означает, что результат теста должен быть повторно проанализирован вслепую тем же специалистом или другими специалистами в условиях, когда они не располагают остальными данными о пациенте. Хороший тест должен быть воспроизводимым, т.е. давать близкие результаты при выполнении разными специалистами в разных условиях. Воспроизводимость результатов инструментальных и, тем более, физикальных

исследований не так высока, как представляется. Относительно низка воспроизводимость ультразвуковых, рентгенологических, радиоизотопных, эндоскопических и других исследований, связанных с оценкой «изображения».

Спектр патологии у обследованных лиц

Выборка пациентов для диагностического эксперимента должна быть репрезентативной, т.е. отражать особенности популяции пациентов. Для этого в выборке должны быть адекватно представлены сопутствующие и схожие заболевания и разные формы основного заболевания. Возможны случаи, когда предлагаемый ДТ имеет описанные характеристики лишь применительно к одной, тяжелой, форме заболевания или применительно к поздней его стадии, или только в отсутствии других схожих заболеваний.

Для того чтобы оценить выборку пациентов, надо знать, по какому принципу она составлена, каковы были критерии включения и исключения пациентов, предполагалось ли обследование посетителей обычной поликлиники или пациентов специализированного стационара. От этого зависит не только преваленс болезни в изучаемой группе, но и тяжесть ее течения, и соотношение разных ее форм.

Итак, доказательное исследование точности диагностического теста должно проводиться путем сравнения результатов теста с результатами РТ в эксперименте, в котором каждый пациент обследуется этими двумя методами параллельно, и их результаты оцениваются вслепую. Исследование приносит пользу только в том случае, если спектр пациентов, включенных в него, соответствует условиям конкретной практики.

Клинический контекст диагностики

Само по себе стремление к точности диагноза не является оправданным уже потому, что во многих случаях специфическое лечение отсутствует и проводится лечение неспецифическое. Уточнение диагноза иногда бывает бесполезным, ибо приводит к отсрочке лечения, увеличивая как страдания пациента, так и расходы. Использование теста в клинической практике обычно подразумевает применение его в совокупности с другими тестами. Тем не менее полезность изучаемого теста должна быть показана при его самостоятельном, отдельном применении. Встречаются исследования, в которых при очевидной неэффективности диагностики с помощью изучаемого теста предлагается применять его в качестве «дополнения к другим методам исследования». Это нужно считать ложной рекомендацией.

Исследования течения и прогноза болезни

Врач и пациент хотят знать, что может произойти: в какие сроки возможно выздоровление (или наступление смерти); насколько можно изменить те-

чение болезни; какие возможны осложнения и в какие сроки. Для того чтобы разобраться в течении болезни, ее вероятных исходах, научиться определять зависимость исходов от особенностей симптоматики, нужно ознакомиться с исследованием, в котором наблюдали за развитием болезни. Исследования развития болезни выполняют разными методами, в частности изучают анамнез у больных и таким образом воссоздают процесс развития болезни. Этот способ ведет к серьезным ошибкам, поскольку анамнез трудно поддается уточнению и, кроме того, врач никогда не знает, насколько репрезентативна группа пациентов, которых он наблюдает. Единственным способом получения надежных данных о течении болезни является когортное исследование.

Исходная группа пациентов

В исходную группу (когарту) должны быть включены пациенты в детально описанной и одинаковой фазе болезни. Если фаза болезни не учтена, то нельзя с уверенностью использовать результаты исследования. Например, в ходе исследования прогноза при остром инфаркте миокарда (ОИМ) в качестве исходной группы пациентов могут быть взяты лица, поступившие в приемное отделение с острым коронарным синдромом и отсутствием подъема сегмента ST на ЭКГ. Судить по этой группе об исходах всех форм ОИМ в целом будет неверно. В действительности летальность при ОИМ будет значительно выше, так как часть больных погибает, не успев поступить в палату интенсивной терапии. Другие, с подъемом ST на ЭКГ, имеют существенно иной прогноз. Исследование корректно, но его результаты имеют ограниченное применение. Если же исходная группа пациентов не была четко определена, то статья не представляет интереса.

Методы оценки стадии, тяжести и других особенностей течения заболевания должны быть точно описаны. Желательно использование испытанных методов, и, если применяется нестандартный метод, необходимы сведения, позволяющие оценить его качество.

Источником подбора определяется, насколько результаты исследования можно переносить на условия, в которых работает конкретный врач. При оказании квалифицированной помощи, в особенности при хронических и тяжелых состояниях, пациентов помещают в специализированные центры. С одной стороны, в них концентрируются самые сложные больные, а с другой — там возможны диагностика и лечение в максимальном объеме. Поэтому полученные в таком центре результаты не соответствуют практике семейного врача.

Обычно врачи первичной помощи сталкиваются с болезнями меньшей тяжести по сравнению с состоянием пациентов специализированных центров, поэтому их пациенты менее склонны соглашаться

на травмирующее лечение и хуже выполняют предписания (комплаенс). С другой стороны, у более легких больных эффект лечения менее заметен. Соответственно вмешательства, оправданные в специализированном центре, могут быть неподходящими для общей врачебной практики. Обращаясь к врачу общей практики, пациенты нередко не имеют определенного заболевания (их страдания порой носят даже немедицинский характер). К таким пациентам можно применять вмешательства, только исходя из имеющихся о них сведений.

Итак, практика в специализированных центрах отличается от практики семейного врача. Но из таких центров исходит большая часть научных сообщений. Не означает ли это, что такими сообщениями следует вообще пренебрегать? Ни в коем случае! Точно так же, как не следует пренебрегать сообщениями из зарубежных центров, где диагностика и лечение могут принципиально отличаться от практики в России. Нужно оценивать в первую очередь методический уровень исследования, отмечая существующие различия и делая на них по мере возможности поправку.

Полнота отслеживания

Отслеживание принципиально важно для качества исследования. Результат наблюдения можно считать соответствующим изначально отобранной группе только, если за всей группой велось наблюдение в течение необходимого срока. Отсев более 10% пациентов внушает подозрение, а более 20% — основание для того, чтобы сомневаться в результатах. Возможно, прогноз благоприятен просто потому, что 20% умерших участников потерялись, отсеялись. В статье обязательно должны быть приведены для сравнения полные исходные характеристики выбывших и отслеженных пациентов. Если выбывшие пациенты отличаются от отслеженных или число выбывших велико, искажения результата неизбежны.

Критерии исхода

Лишь документированный смертельный исход является очевидным и надежным результатом. Все остальные исходы требуют формулировки точных критериев, в том числе и любая выдвинутая причина смерти. Одно дело, если она устанавливается в клинике, другое — если по свидетельству о смерти, где причина указывается очень приблизительно. Если речь идет об обострении процесса или, напротив, о наступлении ремиссии, то для этих исходов должны быть сформулированы соответствующие критерии. Исходы должны оцениваться вслепую, т.е. анализирующий состояние пациента врач не должен знать о том, в каком состоянии пациент находился ранее и каковы результаты других инструментальных исследований.

Испытания методов лечения и профилактики

Существует множество разнообразных оснований предполагать, что то или иное медицинское вмешательство эффективно. Единственным способом установить, действительно ли оно помогает, является контролируемое испытание. Контролируемыми называют испытания, в которых одних пациентов лечат испытываемым методом, а других — контрольным методом (методом сравнения). Наиболее совершенным методом сравнительного испытания являются рандомизированные контролируемые испытания (РКИ).

Контролируемые испытания

При изучении сообщения об испытании лечебного или профилактического вмешательства следует определить, насколько его структура позволяет выявить действенность вмешательства и устранить возможное влияние от привходящих факторов. Правильная структура испытания соответствует следующим критериям:

1. Наличие «контрольной группы» (одной или более).
2. Ясные критерии выбора (включения и исключения) пациентов.
3. Включение пациентов в исследование до назначения лечения.
4. Случайный выбор лечения.
5. Правильный метод рандомизации.
6. Отсутствие у пациентов информации о лечении («слепота»).
7. «Слепая» оценка результата лечения.
8. Информация об осложнениях и побочных эффектах лечения.
9. Информация о числе пациентов, выбывших в ходе эксперимента.

Структура исследования

Главный критерий качества исследования — понятная структура: «опытная» группа, получающая изучаемое лечение, и «контрольная», получающая плацебо или «обычное» лечение. Действенность вмешательства оценивается сравнением исходов в группе вмешательства с исходами в группе контроля. В случае, если такого сравнения нет, оценить эффект изучаемого лечения невозможно, и, следовательно, статья не представляет интереса.

Контрольная группа может получать плацебо — вмешательство, похожее на испытываемое, но не содержащее активного компонента. Пример — таблетки из крахмала или разрез кожи без артроскопической операции. В случае, если существует действенное лечение, то лишать пациентов такого лечения нельзя, и в контрольной группе применяют стандартное (обычное) лечение.

Критерии включения пациентов в исследование и критерии исключения должны быть описаны де-

тально. Если они неясны, то врач не знает, можно ли описанных пациентов рассматривать как похожих на тех, с которыми он сам сталкивается, и будет ли вмешательство действовать в его популяции так же, как в отобранной для испытания группе. Критерии включения и исключения должны применяться до рандомизации.

Рандомизацией называют распределение больных в группу вмешательства и контрольную группу случайным образом, например по таблице случайных чисел. Основной целью рандомизации является создание одинаковых групп. Только с помощью рандомизации можно получить группы пациентов, схожие по всем признакам, в том числе и по тем, которые исследователь не измеряет. Одновременно целью рандомизации является сокрытие получаемого лечения, поэтому метод и результат рандомизации не должен быть известен участникам испытания. Таким образом, из рассмотрения нужно исключать статьи об испытаниях, где распределение пациентов по группам было неслучайным или метод рандомизации представляется неудовлетворительным (по первой букве имени, номеру истории болезни и т.п.), или способ распределения не описан вообще.

При слепом методе испытания пациенты не знают, какому именно вмешательству они подвергаются. Желательно, чтобы лечение было неизвестно также врачам (двойное слепое испытание). Это необходимо для того, чтобы действия врачей и оценки состояния больных врачами не зависели от того, какое лечение получает пациент. Должно быть понятно, какие оценки в РКИ (диагноз, оценка исхода) выносились вслепую, а какие — нет.

Анализ результатов исследования

Часть больных могут отказываться от назначенного лечения или по каким-то причинам подвергаться необходимому им, но не запланированному в испытании лечению. Возникает вопрос, к какой группе таких больных относить? Существуют несколько решений этой проблемы.

А. Пациентов, которые не получили лечения, положенного по рандомизации, анализируют в зависимости от фактически полученного лечения. Недостаток этого анализа в том, что переходить на «традиционное лечение» могут преимущественно пациенты с большей тяжестью болезни, и чаще это происходит в ходе применения более инвазивного метода. Соответственно, исходы «традиционного лечения» могут быть ухудшены, а результаты нового — выглядеть лучше, чем они есть на самом деле.

Б. Пациентов, прекративших экспериментальное лечение, и контрольных, изменивших свое лечение, исключают из анализа (анализ в соответствии с выполнением протокола). Это плохой метод, поскольку побочные эффекты могут быть ярче выра-

жены у более тяжелых больных, и их исключение изменит исходы. В группе плацебо таких случаев непереносимости лечения скорее всего будет меньше. Результат лечения новым методом в данном примере окажется завышенным.

В. Анализ в соответствии с назначением по результатам рандомизации: исходы лечения рассматривают у больных в группах, выделенных при рандомизации, независимо от того, какое лечение они фактически получили. Главное достоинство метода в том, что он моделирует реальную практику. Ведь врач выбирает лечение сегодня, не зная точно, как придется вести больного завтра. В качественном исследовании должны приводиться результаты именно такого анализа.

Оставаться в исследовании в течение длительного времени обычно могут не все больные люди по разным причинам. Если вышло много пациентов, то нельзя узнать, относятся ли конечные результаты ко всем пациентам, включенным в исследование. Если выбывшие лица отличаются по своим характеристикам в начале исследования от отслеженных или если утрачено более 20%, то следует ожидать больших искажений в оценке действенности вмешательства.

Вмешательство должно оцениваться по самому существенному исходу, называемому основным или первичным. Например противодиабетическое лечение не может быть оценено вполне по стойкости и величине гипогликемического эффекта. В первую очередь нужно знать, каковы показатели смертности. Самый блестящий гипогликемический эффект перечеркивается даже небольшим увеличением смертности, слепоты или частоты ампутаций. Основной причиной подмены конечных результатов является недостаточность данных для их анализа или благоприятный для заказчика исследования характер изменения суррогатных исходов.

Статистическая и клиническая значимость результатов

Обнаружение статистически значимой разницы в исходах означает преимущество одного из сравниваемых вмешательств. Отсутствие статистически значимого различия может быть также следствием малой численности пациентов в РКИ. Это называется недостаточной статистической мощностью испытания. В хорошо спланированном РКИ необходимое число пациентов, исходя из предполагаемой величины эффекта и заданной вероятности ошибок, должно быть рассчитано заранее.

Статистически значимо то, что действительно существует с высокой вероятностью. Обычно эту величину обозначают как $p < 0,05$, т.е. отсутствие различия имеет низкую вероятность — меньше 5%. Клинически значимая величина своими размерами (например снижением смертности) убеждает врача в

Таблица 2

Лечение	Неблагоприятный исход		Всего
	наблюдался	отсутствовал	
Применялось	A	B	A+B
Плацебо	C	D	C+D

необходимости изменить свою практику в пользу нового образа действий. Если различия статистически незначимы, то надо проверить, достаточна ли численность участников. Если различия статистически значимы, то надо обратить внимание на величину эффекта. При большом числе участников испытания статистическая значимость может присутствовать, а эффект может быть очень малым, клинически незначимым, неблагоприятным для пациента.

Для количественного выражения результата вмешательства, размера эффекта, используют несколько показателей (табл. 2).

Риск при проведении лечения = $A/(A+B)$.

Риск при отсутствии лечения = $C/(C+D)$.

Абсолютное снижение риска (absolute risk reduction, ARR) $ARR = C/(C+D) - A/(A+B)$.

Число пациентов, получающих лечение, на один предотвращенный неблагоприятный исход выражается (ЧПЛП, number needed to treat, NNT) $NNT = 1/ARR$.

Абсолютное повышение риска (absolute risk increase, ARI) $ARI = A/(A+B) - C/(C+D)$ рассчитывается, когда лечение по сравнению с контролем дает большую частоту неблагоприятных исходов, например, при лечении возникают вредные побочные эффекты чаще, чем в контрольной группе (например, кровотечения у больных, принимающих варфарин).

Число пациентов, получающих лечение, на один дополнительный вредный исход выражается (ЧПЛВ, number needed to harm, NNH) $NNH = 1/ARI$.

Относительный риск (RR) $RR = [A/(A+B)]/[C/(C+D)]$.

Различие риска (снижение относительного риска) (RR reduction, RRR) $RRR = 1 - RR$.

Шансы на неблагоприятный исход при лечении = A/B .

Шансы на неблагоприятный исход при отсутствии лечения = C/D .

Отношение шансов (OR) $OR = (A/B)/(C/D)$.

Сопоставляя два вида лечения, сравнивают вероятность неблагоприятных исходов в группах. Подразумевается, что лечение приводит к снижению риска, и разница между этими вероятностями называется «абсолютным снижением риска». Снижение риска лучше всего отражает ту пользу, которую дает метод в условиях испытания и с наименьшими искажениями. С точки зрения перспективы применения метода лечения, удобен показатель

ЧПЛП. Он переводит относительные величины в число больных, которых надо лечить, чтобы предотвратить один неблагоприятный исход. Это самый наглядный показатель. Оба эти показателя отражают величину эффекта только применительно к той частоте исходов, которая имела место в данном РКИ в контрольной группе.

Относительный риск — вторичный относительный показатель, используемый, чтобы сделать эффект вмешательства менее зависимым от того, в каких условиях и при каком риске проводится вмешательство. Важно, что с помощью этого показателя клинически незначительный эффект, например снижение риска ампутации от 0,01 до 0,003, может быть представлен как снижение риска в ТРИ РАЗА.

Величина каждого показателя должна оцениваться не сама по себе, а в связи с ее доверительным интервалом (ДИ). Обычно представляют 95% ДИ. При статистически значимом эффекте ЧПЛП выражается например как 12 (95% ДИ 8—24). Если ДИ для OR или RR включает единицу, это означает, что эффект лечения статистически незначим.

Другие методы оценки эффективности вмешательств

Данные доброкачественных РКИ существуют не для всех применяемых в медицине вмешательств. Немало вмешательств практикуется на основании того, что они «применяются уже давно» и «хорошо себя зарекомендовали». Во многих странах, в том числе в России, лекарства допускаются или в недавнем прошлом допускались к применению без доказательств их эффективности, полученных в РКИ.

Если вы не находите статьи по изучаемой проблеме, описывающей РКИ вмешательства на высоком методическом уровне, нужно убедиться в том, что таких исследований действительно нет. Прежде всего, необходимо проверить, правильно ли осуществлялся поиск информации, надежные ли источники информации использованы, корректно ли составлен запрос. Убедившись в отсутствии РКИ, правильнее всего отказаться от такого неизученного вмешательства. Если по каким-то причинам это неприемлемо, то можно перейти к методически менее совершенным исследованиям: нерандомизированным испытаниям, сравнениям исходов в больницах, применяющих разные методы лечения (географический контроль), сравнениям исходов после введения нового вмешательства с исходами, имевшими место до того (исторический контроль). Таких исследований всегда больше, они менее доказательны, но зато более ярко подтверждают эффективность вмешательства. Необходимо, однако, помнить, что именно такими исследованиями обосновывали в свое время все виды лечения, позднее оказавшиеся неэффективными, — лечение язвенной болезни сном, незрелых новорожденных кислородом и т.п.

Побочные эффекты и осложнения

Описание отдельного случая неблагоприятного явления у пациента, принимающего лекарство, обычно не доказывает связи этого явления с приемом лекарства. Для выявления побочных эффектов исследование должно доказать, что они чаще встречаются при вмешательстве, чем без него. Поскольку побочные эффекты редки (на практике обычно не внедряют вмешательства, дающие частые или опасные побочные эффекты), их трудно изучать путем РКИ, их изучают в исследовании типа сравнения с контролем (ИСК). Например, когда заподозрили, что прием талидомида (Т) беременными женщинами вызывает уродства плода, достаточно было сопоставить частоту употребления Т матерями здоровых новорожденных и матерями новорожденных с уродствами.

Поскольку группы принимающих Т и не принимающих Т не формируются способом рандомизации, они различаются по многим признакам. Соответственно, различия исходов беременности могут быть связаны с разными причинами. Поэтому обычно результаты ИСК считаются поддерживающими гипотезу о связи неблагоприятного эффекта с определенным фактором, только когда частота эффекта при наличии этого фактора (приеме Т) многократно превышает частоту в отсутствие его (у не принимавших Т). Самыми слабыми свидетельствами возможного вреда от медицинского вмешательства являются описания отдельных случаев, более надежными представляются результаты ИСК, еще более убедительны когортные исследования, но самые достоверные результаты дают контролируемые испытания.

Решение о том, следует ли применять данный тип лечения, зависит от соотношения ожидаемых полезных эффектов и осложнений. Например, частота хорошо известных осложнений вследствие приема нестероидных противовоспалительных средств низка, поэтому известные осложнения не препятствуют назначению этих лекарств, когда их применение показано. В случае, если польза не столь велика или вред слишком велик, возникает необходимость тщательно взвесить и соотнести то и другое.

Обзоры

Для врача будет большой удачей обнаружить обзор исследований по вставшей перед ним проблеме. Обзор является синтетическим (обобщающим) исследованием текущего состояния вопроса. Это исследование можно считать удовлетворительным, если оно отвечает минимальным критериям качества систематического обзора (systematic review, СиО).

1. Предмет обзора ясен и имеет четкие границы (каким пациентам производилось вмешательство и каковы его исходы).

2. Метод поиска релевантных (т.е. соответствующих проблеме) источников удовлетворителен.

3. Метод отбора источников понятен: обычно отбираются самые доказательные исследования.

Методы СиО были разработаны в тесной связи с методологией ДМ. Наибольшая коллекция СиО высокого качества входит в Кокрановскую библиотеку, и с рефератами этих обзоров, освещающих результаты лечебных и диагностических вмешательств, можно бесплатно ознакомиться в MEDLINE.

СиО отличаются систематическим поиском всех сообщений об оригинальных исследованиях. В случае привлечения для поиска печатных или компьютерных БД должно быть указано, какие их разделы (рубрики, ключевые слова) использованы и какие дополнительные критерии отбора применялись. Таким образом, можно оценить полноту поиска, как и соответствие метода поиска рассматриваемой проблеме. Чем меньше число исследований, включенных в обзор, и чем ниже качество этих исследований, тем больше вероятность ошибочного вывода обзора (подтверждение или отрицание эффекта).

Обычно не все найденные исследования включают в обзор. Должно быть сформулировано правило, по которому их отбирали. Например это могут быть только исследования на людях, только РКИ, выполненные двойным слепым методом.

Если различия между результатами исследований невелики, то можно объединить их данные в общем статистическом анализе (метаанализе). Полезность метаанализа вытекает из следующего. Когда число пациентов в исследовании невелико, то маловероятно получение статистически значимых эффектов или различий (низкая статистическая мощность исследования). В результате небольшой, но клинически значимый эффект может быть не обнаружен. Поскольку схожие исследования дополняют друг друга и их методология также схожа, в СиО возможно объединение данных, и увеличение числа пациентов позволяет получить более надежные, статистически значимые оценки.

Оценка медицинских технологий

В оценке медицинской технологии (процедуры, метода, программы) условно выделяется три этапа:

— *технологический*: оцениваются технические характеристики — точность и воспроизводимость измерения, токсичность лекарства;

— *клинический*: оценивается диагностическая или терапевтическая действенность в контролируемом эксперименте; оценка проводится в строго определенных условиях и включает клинически

важные исходы (интервал до первого обострения, средняя продолжительность жизни); в дальнейшем в более широких испытаниях и в менее строгих условиях оценивается клиническая эффективность, т.е. эффективность при обычном применении вмешательства;

— *комплексный*: анализ включает возможно наиболее полный комплекс естественных оценок — клинических, экономических и социальных.

В медицинской практике результаты ведения больного редко зависят от одного вмешательства. Врач, пациент и общество одновременно принимают во внимание действенность, опасность и цену соответствующего лечебного вмешательства, и приемлемость для пациента отдельных специфических последствий. Для исследований комплексных ситуаций существует особый подход, который принято называть анализом принятия решений (decision analysis). Этот вид анализа состоит в последовательном рассмотрении вероятных результатов каждого из возможных действий, с тем чтобы выбрать наилучший вариант. С 1980-х годов в некоторых развитых странах созданы государственные агентства для оценки медицинских технологий. Они выполняют комплексные оценки отдельных вмешательств, в основном новых и дорогостоящих, учитывая, как эти вмешательства могут быть использованы в рамках существующей системы здравоохранения, и выносят рекомендации о том, следует ли эти вмешательства внедрять (оплачивать в рамках государственной системы или в рамках программ медицинского страхования). Наиболее известное и продуктивное агентство работает в Великобритании (<http://www.nice.org.uk/>).

Клинические рекомендации

Медицинские профессиональные общества, государственные организации и страховые компании готовят для врачей «рекомендации» и «указания», т.е. вторичные документы, регламентирующие порядок действий в соответствии с современным состоянием медицинской науки и другими обстоятельствами. Такие клинические рекомендации (КР) создаются на основе коллективных усилий и обычно подвергаются строгой процедуре редактирования и проверки. Поэтому американские, финские или английские врачи относятся к КР, выпущенным их медицинскими обществами, серьезно. Эти рекомендации — стандарт практики. В случае возникновения конфликта врач может опираться на них с уверенностью, доказывая, что действовал верно.

КР подготавливают по специальной технологии. Отличия хорошо разработанных КР:

— полное освещение проблемы, например ведения пациентов с травмой позвоночника, больных диабетом, в определенных условиях (поликлиника, стационар, специализированный центр);

— освещение разных аспектов — пользы и вреда, но не с точки зрения баланса, как это имеет место в анализе принятия решений, а с точки зрения формулировки качественной рекомендации в отношении того, как поступать;

— открытое представление ценностной ориентации авторов, в том числе с четким использованием некоторой экономической модели.

Перечисленное делает КР более важным для врача документом, чем отдельные исследования, опубликованные в журналах. Разработка таких КР сложна, поэтому, несмотря на то что сегодня прилагаются большие усилия в этой сфере, для многих серьезных ситуаций КР отсутствуют.

КР представляют читателю-врачу лишь самые важные выводы, не перегружая его исходными материалами или необходимостью критической оценки этих материалов. В этом есть определенный минус: читатель не может быть вполне уверенным в том, что КР действительно соответствуют передовому уровню медицинской науки. Поэтому обращение к КР не освобождает врача от необходимости критически оценивать методы, использованные авторами КР.

Следует задать себе ряд вопросов.

1. Ясно ли сформулированы возможные варианты действий и исходы?
2. Рассматривают ли КР все возможные варианты болезни и вмешательства?
3. Надежен ли использованный процесс поиска, выбора и обобщения данных?
4. Насколько надежен и понятен процесс «взвешивания» различных исходов?
5. Включают ли КР недавние значительные успехи в данной области?
6. Проходили ли КР независимое рецензирование?
7. Есть ли в КР практически важные решения, насколько они обоснованы?
8. Насколько сильно влияет неопределенность в предпочтениях и ценах на выводы?
9. Соответствуют ли найденные КР проблеме, с которой столкнулся врач, и с особенностями его конкретной практики?
10. Не связаны ли составители КР с производителем определенного препарата или протеза, и не могут ли быть КР искажены в угоду этому производителю?

Обсуждаемые в 2012 г. «стандарты специализированной помощи» не являются КР. С наилучшей подборкой КР можно ознакомиться на <http://guidelines.gov/>.

Уровни доказательности сведений

Применительно к каждому вопросу (эффективен ли метод диагностики, эффективно ли профилактическое вмешательство) можно получить дока-

зательства в исследованиях разного дизайна. Выше мы рассмотрели основные дизайны исследований и их элементы, которые делают исследование доказательным. Соответственно, анализируя всю совокупность сведений, имеющих отношение к рассматриваемому вопросу, можно выстроить шкалу доказательности сведений, которыми мы располагаем.

1. СиО нескольких качественных исследований.

2. По крайней мере одно хорошо организованное экспериментальное исследование.

3. Хорошо организованные квазиэкспериментальные исследования (нерандомизированные контролируемые испытания, исследования одной группы до и после лечения, когортные, исследования серий случаев, ИСК подобранных парных случаев).

4. Хорошо организованные не экспериментальные исследования (сравнительные дескриптивные, корреляции).

5. Описания клинических случаев.

При обобщении литературных данных в ходе подготовки рекомендаций должны быть оценены доказательность имеющихся исследований, гетерогенность их результатов, величина полученного эффекта и точность оценки этого эффекта. На окончательном этапе составления КР шкалу доказательности еще упрощают, чтобы обобщить все разнообразие имеющихся сведений.

Сила и убедительность сведений.

А. Сведения типа 1 или согласующиеся результаты исследований типов 2, 3 или 4.

Б. В основном согласующиеся сведения типов 2—4.

В. Противоречивые сведения типов 2—4.

Г. Отсутствие сведений или сведения только типа 5.

В разных изданиях и в разных организациях, разрабатывающих КР, шкала доказательности и обозначение силы и убедительности сведений могут разли-

чаться в деталях. Наиболее продвинутой является шкала доказательности GRADE, все шире применяемая во всем мире (<http://www.gradeworkinggroup.org>).

Варианты действий и их выбор

Если список рассматриваемых в КР вариантов действий не включает всех возможных способов действий, то значение рекомендаций снижается, поскольку упущенное вмешательство может быть лучшей, или наиболее безвредной, или наиболее доступной альтернативой. В медицинской практике сохраняются методы, эффективность которых никогда не была проверена в доказательных испытаниях, и внедряются подобные новые, например лечение различных заболеланий низкоэнергетическими лазерами. В результате, если КР рассматривают достаточно обширный список вмешательств, то в отношении большего числа вмешательств не удастся найти сведений, основанных на доказательствах уровня А. Поэтому список рассмотренных вмешательств никогда не бывает исчерпывающим, но он должен включать все вмешательства, полезность которых проверена в сравнительных испытаниях.

Поскольку большинство РКИ проводится в специализированных центрах, постольку важно, чтобы при подготовке КР для врачей общей практики учитывались особенности их пациентов. Точно так же, если упущен один из возможных исходов или осложнений, оценка ситуации может быть неверной. Например, если в рекомендациях по профилактике рассматривается только смертность от сердечно-сосудистых заболеланий, а изменения в общей смертности не анализируются, то такие рекомендации, очевидно, искажают картину. Ведь для врача, как и для пациента, в первую очередь важны смертность и ожидаемая продолжительность жизни, а не отдельная причина смерти.