

Основные принципы доказательной медицины В ОНКОЛОГИИ

 С.Б. Петерсон*, А.И. Бенековский*, В.Н. Чехонадский**

* Кафедра онкологии РГМУ

** ГУ Российский онкологический научный центр
им. Н.Н. Блохина РАМН, Москва

Доказательная медицина (ДМ) — фундаментальная медицинская наука, основным направлением которой является прогнозирование для конкретного пациента течения его болезни и последствий проводимого (планируемого) лечения на основании изучения клинических данных аналогичных больных с использованием строгих научных методов, обеспечивающих точность прогноза. В России эту науку называют чаще всего именно “доказательной медициной”, в то время как в англоязычных странах более распространено другое название — “клиническая эпидемиология”. Цель ДМ — разработка и применение таких методов клинического наблюдения, которые дают возможность избежать влияния систематических и случайных ошибок и получить максимально достоверный прогноз.

Количественный метод оценки лечения был предложен менее 200 лет назад, когда французский врач Пьер Луи подсчитал число случаев смерти и выздоровления при лихорадке в зависимости от применения кровопускания. Его осуждали за то, что он позволил себе на основе безжизненных цифр сеять сомнения в целительной силе пиявок, но в итоге его данные послужили причиной для пересмотра существовавших тогда методов лечения.

Среди отечественных основоположников ДМ необходимо отметить академика Н.Н. Блохина (1912–1993), уделявшего огромное внимание доказательности всех клинических результатов и выступавшего

непримиримым противником любых ненаучных методов врачевания, а также профессора И.И. Шатрова (1902–1977), который в своих работах излагает основные принципы ДМ и определяет рациональные основы сотрудничества врачей с математиками.

Традиционное клиническое мышление заключается в том, что если мы знаем механизм заболевания, то можем предсказать течение болезни и выбрать правильное лечение (на основе данных биохимии, анатомии, физиологии и других фундаментальных наук). Но механизмы развития заболеваний известны нам лишь отчасти, а на исход болезни влияют и многие другие факторы (генетические, физические, психологические, социальные и т.п.), поэтому такие прогнозы следует рассматривать только как клинические гипотезы, которые должны выдержать проверку в клинических испытаниях. Личный опыт врача тоже важен для принятия клинических решений. Однако ни один врач не обладает достаточным практическим опытом, чтобы распознать все трудноуловимые, длительно протекающие, взаимодействующие процессы, которые имеют место при большинстве хронических заболеваний.

Хотя личный опыт врача и знание механизмов развития заболевания, безусловно, важны, необходимо принимать во внимание следующие **основные положения ДМ**:

1) в большинстве случаев диагноз, прогноз и результаты лечения для конкретного больного однозначно не определены и по-

тому должны быть выражены через вероятности;

2) эти вероятности для конкретного больного лучше всего оцениваются на основе предыдущего опыта, накопленного в отношении групп аналогичных больных;

3) поскольку клинические наблюдения проводятся на свободных в своем поведении больных, и делают эти наблюдения врачи с разной квалификацией и собственным мнением, результаты могут быть подвержены систематическим ошибкам, ведущим к неверным заключениям;

4) любые наблюдения, и клинические в том числе, подвержены влиянию случайности;

5) чтобы избежать неверных выводов, врачи должны полагаться на результаты исследований, основанных на строгих научных принципах и выполненных с использованием способов минимизации систематических и случайных ошибок;

6) сведение к минимуму систематических ошибок достигается правильной структурой (организацией) исследования, адекватной его задачам;

7) сведение к минимуму случайных ошибок достигается корректным статистическим анализом данных.

Взаимоотношения врача и пациента в рамках клинических исследований (КИ) регламентирует Хельсинкская декларация, принятая 18-й ассамблеей Всемирной медицинской ассоциации в 1964 г. и затем неоднократно пересмотренная (последняя редакция принята на 52-й ассамблее в Эдинбурге в 2000 г.).

В настоящее время считается общепризнанным, что неправильно проведенное (как при планировании, так и при анализе данных) КИ является неэтичным, поскольку пациенты при этом подвергаются неоправданному риску, неэффективно используются финансовые ресурсы и время исследователей, после публикации неверных результатов дальнейшие исследования направляются в неправильное русло, а их

применение в медицинской практике может нанести вред больным. Поэтому КИ нельзя проводить без участия специалиста по прикладной медицинской статистике уже на этапе планирования.

При планировании КИ необходимо заранее четко установить, прогноз какого исхода лечения мы будем стараться определить. Наиболее значимые **клинические исходы**:

- смерть;
- заболевание — набор симптомов, физических и лабораторных данных, отклоняющихся от нормы;
- дискомфорт — субъективно тягостные симптомы (боль, тошнота, одышка, зуд, шум в ушах и т.п.);
- инвалидизация — неспособность к обычной деятельности на работе или дома;
- неудовлетворенность — эмоциональная реакция на болезнь и проводимое лечение (тоска, гнев и др.).

Основные понятия ДМ

Популяции и выборки

Популяция — большая группа людей, проживающих в определенном географическом регионе или обладающих некоторым признаком (например, старше 65 лет). Она может состоять из пациентов, госпитализированных в определенную клинику, или пациентов с определенным заболеванием (что чаще имеет место в КИ). Таким образом, можно говорить об общей популяции, госпитальной популяции или популяции пациентов с конкретным заболеванием.

Выборка — это часть популяции, полученная путем отбора. КИ обычно выполняются на выборках, и оценивать характеристики популяции приходится путем оценки этих характеристик в выборке.

Виды ошибок

Систематическая ошибка — это систематическое (неслучайное, одностороннее) отклонение результатов от истинных значений. Наблюдение за больными (при

лечении и в процессе КИ) особенно подвержено систематическим ошибкам. Больные ведут произвольно: отказываются участвовать в исследовании; выбывают из исследования; меняют метод лечения и т.п. К тому же сам исследователь склонен верить в успех нового метода.

Основные категории систематических ошибок в КИ:

- ошибка, обусловленная отбором, возникает, когда сравниваемые группы пациентов различаются не только по изучаемому признаку, но и по другим факторам, влияющим на исход;
- ошибка, обусловленная измерением, возникает, когда в сравниваемых группах больных используются разные методы измерения;
- ошибка, обусловленная вмешивающимися факторами, возникает, когда один фактор связан с другим, и эффект одного искажает эффект другого.

В одном КИ могут присутствовать систематические ошибки всех трех категорий.

Случайная ошибка — отклонение результата отдельного наблюдения в выборке от истинного значения в популяции, обусловленное исключительно случайностью.

Систематическая ошибка вызывает отклонение оценки от истины либо в одну, либо в другую сторону. Случайная ошибка вызывает отклонение оценки от истины с равной вероятностью и в одну, и в другую сторону. Систематическую ошибку можно предотвратить, а случайную ошибку предотвратить нельзя.

Достоверность и обобщаемость

Достоверность определяется тем, в какой мере полученные результаты справедливы в отношении данной выборки. Это внутренняя характеристика выборки.

Обобщаемость — внешняя характеристика выборки, определяется тем, в какой мере результаты данного КИ применимы к другим группам больных.

Типы данных, норма и отклонение

Данные, полученные в результате клинических наблюдений, бывают трех типов: качественные, порядковые и количественные.

Качественные данные — это величины, которые нельзя расположить в естественном порядке (например, пол, врожденные нарушения метаболизма и др.). Они могут описывать важные дискретные события: смерть, хирургическое вмешательство и т.п. Качественные данные, которые могут быть отнесены только к двум категориям, называются дихотомическими (или бинарными).

Порядковые данные — величины, которые могут быть расположены в естественном порядке или ранжированы (от малого до большого, от хорошего до плохого), но размер интервала между такими категориями не может быть выражен количественно. Примером служит стадия онкологического заболевания, степень сердечной недостаточности и т.п.

Количественные данные — это такие величины, которым присущ естественный порядок расположения с равными интервалами между последовательными значениями, независимо от их места на шкале. Количественные данные могут быть непрерывными и дискретными. Непрерывные могут принимать любое значение на непрерывной шкале, например большинство биохимических показателей крови, масса тела, артериальное давление и т.п. Дискретные данные всегда выражаются целыми числами, например число беременностей, рождений, аборт и т.п.

Данные всех этих типов могут быть охарактеризованы достоверностью, воспроизводимостью и вариациями измерений.

Достоверность измерений (точность оценки) показывает, в какой степени полученные данные соответствуют истинным значениям, т.е. действительно отражают оцениваемое явление.

Таблица 1. Способы представления центральной тенденции и дисперсии

Способ представления	Определение	Преимущества	Недостатки
Центральная тенденция			
Среднее	Сумма полученных значений, деленная на число наблюдений	Удобна для математической обработки	Подвержена влиянию крайних значений
Медиана	Точка, по обе стороны которой находится равное число наблюдений	Не подвержена влиянию крайних значений	Не слишком удобна для математической обработки
Мода	Наиболее часто встречающееся значение	Проста для понимания	Иногда мода отсутствует, а иногда их бывает несколько
Характеристики рассеяния			
Диапазон значений	Рассеяние от минимального до максимального значений в распределении	Охватывает все значения	Сильно подвержен влиянию крайних значений
Стандартное отклонение	Абсолютная величина среднего арифметического отклонения индивидуальных значений от среднего значения	Удобно для математической обработки	Не всегда пригодно для описания негауссовых распределений
Процентили, децили, квартили	Доля наблюдений, находящихся в определенной части распределения	Характеризуют “необычность” значения независимо от типа распределения	Неудобны для статистических расчетов

Воспроизводимость — вероятность того, что при повторных измерениях некоего устойчивого явления, сделанных разными людьми, на разных приборах, в разное время и в разных местах, будет получен один и тот же результат.

Вариации — это разница в результатах измерений, связанная с процессом измерений, биологическими изменениями, происходящими в организме в течение времени, и биологическими различиями между разными индивидуумами.

Частотные распределения для клинических показателей имеют разную форму и характеризуются центральной тенденцией и рассеянием (дисперсией), которые могут быть представлены различными способами (табл. 1).

Значения лабораторных показателей для здоровых и больных часто перекрываются. Вследствие этого, а также из-за относительно низкой распространенности пато-

логии в общей популяции часто бывает невозможно четко разграничить две группы пациентов по результатам однократного тестирования. Точка разделения, на которой заканчивается норма и начинается патология, выбирается произвольно и обычно связана с одним из трех критериев патологии — это состояние:

- 1) необычное с точки зрения статистики;
- 2) проявляется в виде болезни;
- 3) поддается лечению.

Диагноз

Для оценки эффективности диагностического теста строят так называемую четырехпольную таблицу и сравнивают результаты теста при наличии и отсутствии заболевания. Для простоты изложения мы заполним такую таблицу гипотетическими цифрами (табл. 2). Предположим, что были обследованы 149 пациентов; каждому из них тест проводился однократно. Всего

тест был положительным в 62 случаях, а отрицательным — в 87. Затем была проведена верификация диагноза с помощью некоего “золотого стандарта”, позволяющего с абсолютной точностью определить наличие или отсутствие болезни, при этом болезнь выявлена у 37 человек, отсутствовала у 112. Полученные результаты были занесены в четырехпольную таблицу.

Если у пациента болезнь присутствует, и результат теста на наличие болезни положительный, то такие результаты называют **истинноположительными** (в числовом примере в табл. 2 таких результатов 27).

Если у пациента болезнь отсутствует, и результат теста на наличие болезни отрицательный, то такие результаты называют **истинноотрицательными** (в примере таких результатов 77).

Если у пациента болезнь присутствует, а результат теста отрицателен, то такие результаты называют **ложноотрицательными** (в примере таких результатов 10).

Если у пациента болезнь отсутствует, а результат теста положительный, то такие результаты называют **ложноположительными** (в примере таких результатов 35).

В современной научной литературе приняты **7 основных характеристик диагностического теста**.

1. Чувствительность — показывает, в каком проценте случаев распознается болезнь. В рассматриваемом числовом примере

$$Se = 27 : 37 = 73\%.$$

2. Специфичность характеризует точность отрицательного результата, т.е. насколько отрицательный результат специфичен для отсутствия заболевания. В нашем примере:

$$Sp = 77 : 112 = 69\%.$$

3. Распространенность заболевания:

$$P = 37 : 149 = 25\%.$$

4. Прогностическая ценность отрицательного результата теста:

$$-PV = 77 : (10 + 77) = 88\%.$$

5. Прогностическая ценность положительного результата теста:

$$+PV = 27 : (27 + 35) = 44\%.$$

6. Ложноотрицательные ошибки (ошибки первого рода, α -ошибки) — какой процент отрицательных результатов оказывается ложным:

$$\alpha = 10 : (10 + 77) = 12\%.$$

7. Ложноположительные ошибки (ошибки второго рода, β -ошибки) — какой процент положительных результатов оказывается ложным:

$$\beta = 35 : (27 + 35) = 56\%.$$

Нетрудно заметить, что:

$$-PV + \alpha = 100\%;$$

$$+PV + \beta = 100\%.$$

Частота и вероятность событий

В ходе КИ можно определить только частоту некоторых событий или эффектов, а для построения прогноза необходимо иметь не частоту, а вероятность. Для демонстрации различия между этими понятиями приведем пример из радиационной онкологии.

Появление лучевого повреждения нормальных тканей зависит от большого количества различных, до конца еще не изученных факторов. Поэтому в первом приближении можно считать, что оно носит случайный характер. Если вероятность появления изучаемого эффекта при облучении конкретного больного дозой D составляет некоторую величину η , то по причинам чисто случайным в группе из N больных, облученных этой дозой, не будет диагностировано ни одного лучевого по-

Таблица 2. Четырехпольная таблица для оценки эффективности диагностического теста

Тест	Болезнь	
	присутствует	отсутствует
Положительный	27	35
Отрицательный	10	77

Таблица 3. Количество больных в исследуемой группе без лучевых повреждений, необходимое для клинического обоснования низких вероятностей лучевых повреждений

Доверительная вероятность Р	Количество больных при заданной вероятности		
	5 %	10 %	15 %
0,7	23	11	7
0,9	45	21	14
0,95	58	28	18

вреждения нормальных тканей с вероятностью:

$$P_D = (1 - \eta)^N.$$

Поэтому отсутствие лучевых повреждений в группе из N больных, облученных дозой D, правильно может быть истолковано следующим образом: с доверительной вероятностью $P = 1 - P_D$ вероятность эффекта при данном воздействии составляет

$$\eta \leq 1 - P_D^{1/N} = 1 - (1 - P)^{1/N}.$$

Отсюда следует, что отсутствие лучевых повреждений в группе из 100 больных означает лишь, что с доверительной вероятностью P, равной 0,8; 0,9 или 0,95, вероятность такого эффекта составляет не более чем 1,6; 2,3 и 3% соответственно.

Это же выражение позволяет количественно оценить надежность клинического определения вероятности изучаемых отрицательных эффектов. Для расчета минимального количества больных в группе (N), необходимого, чтобы сказать, что отсутствие в этой группе лучевых повреждений свидетельствует о том, что вероятность эффекта при облучении дозой D не превышает η с доверительной вероятностью P, используется формула

$$N \geq \frac{\ln P_D}{\ln(1 - \eta)} = \frac{\ln(1 - P)}{\ln(1 - \eta)}.$$

Результаты оценок, полученных с использованием этого выражения, представлены в табл. 3. Если новому методу облучения были подвергнуты 58 больных, и ни у

одного из них не было лучевых повреждений, то частота лучевых повреждений в исследуемой группе составляет 0%, а в отношении вероятности лучевых повреждений при таком методе облучения мы можем сказать только, что она не превосходит 5% с доверительной вероятностью 0,95.

Эффект локального излечения опухоли также зависит от большого количества различных причин, и поэтому в первом приближении можно считать, что и он носит случайный характер. Поэтому если вероятность локального излечения у больного в результате облучения дозой D составляет некоторую величину v , то по причинам чисто случайным в группе из N больных, облученной этой дозой, будет определено N локальных излечений (100% частота локальных излечений) с вероятностью

$$P_D = v^N.$$

Поэтому 100% частота локальных излечений в группе из N больных, облученных дозой D, правильно может быть истолкована следующим образом: с доверительной вероятностью $P = 1 - P_D$ вероятность локального излечения при облучении дозой D составляет

$$v \geq P_D^{1/N} = (1 - P)^{1/N}.$$

Таким образом, 100% частота локальных излечений в группе из 100 больных означает лишь то, что с доверительной вероятностью 0,90 или 0,95 вероятность локальных излечений v составляет не менее чем 97,7 или 95% соответственно.

Аналогично приведенному выше примеру с лучевыми повреждениями, можно оценить минимально необходимое количество больных в группе:

$$N \geq \frac{\ln P_D}{\ln v} = \frac{\ln(1 - P)}{\ln v}.$$

Если новому методу облучения были подвергнуты 58 больных, и у всех был достигнут полный эффект, то частота локальных излечений в исследуемой группе со-

ставляет 100%, а в отношении вероятности локальных излечений при использовании такого метода мы можем сказать только, что она составляет не менее 95% с достоверной вероятностью 0,95 (табл. 4).

Риск

Факторы риска связаны с повышением вероятности развития заболевания. Независимо от того, является ли некий фактор риска причиной заболевания, его присутствие позволяет прогнозировать вероятность возникновения болезни.

Воздействие большинства предполагаемых факторов риска нельзя изучать в эксперименте, поэтому обычно приходится проводить наблюдения, отслеживая воздействие факторов и оценивая частоту заболевания. Один из методов такого наблюдения состоит в отборе когорты, некоторые представители которой подвергаются воздействию фактора риска. Когорта — группа лиц, изначально объединенных каким-либо общим признаком (например, здоровые лица или больные с определенной стадией заболевания) и наблюдаемых в течение определенного периода времени, чтобы проследить, что с ними произойдет в дальнейшем. После периода наблюдения сопоставляют частоту развития заболевания у лиц, подвергавшихся и не подвергавшихся воздействию этого фактора. Хотя когортные исследования риска предпочтительны с научной точки зрения, этот подход не всегда можно применить на практике, поскольку он требует значительных затрат времени, сил и средств.

Прогноз

Прогноз — это описание течения заболевания с момента его начала. По сравнению с факторами риска прогностические факторы имеют относительно высокую распространенность и могут оцениваться даже на основании индивидуального клинического опыта.

Прогноз лучше всего описывается вероятностью возникновения исхода в любой момент развития заболевания. В принципе, это может быть сделано при наблюдении когорты больных до тех пор, пока изучаемый исход не наступит у всех, у кого он может произойти. Однако поскольку такой подход неэффективен, используется другой метод, называемый анализом дожития (анализом времени наступления события). При обработке этих данных следует использовать специальные методы, уменьшающие влияние систематических ошибок.

Все основные исходы лечения, анализируемые в клинической онкологии, подпадают под определение “данные типа времени жизни” (survival data). В статистическом смысле это понятие описывает пребывание объектов в интересующем исследователя состоянии в определенный момент времени. При анализе таких данных особый интерес представляет группа объектов (пациентов), для каждого из которых определено точное событие, часто называемое “отказом”. “Отказ” происходит после некоторого интервала времени для каждого объекта только один раз.

Данные о времени жизни n объектов обычно представлены набором соответствующих интервалов времени, рассматриваемых как случайные величины. Для описания распределения таких данных используют различные статистические функции.

Функция распределения (функция надежности) случайной величины, называе-

Таблица 4. Количество больных в исследуемой группе с полным эффектом, необходимое для клинического обоснования высоких вероятностей локальных излечений

Доверительная вероятность Р	Количество больных при заданной вероятности		
	80 %	90 %	95 %
0,8	7	15	31
0,9	10	22	45
0,95	13	28	58

мая также кумулятивной функцией риска отказа, отражает вероятность того, что отказ произойдет через время, меньшее, чем t :

$$F(t) = P(T < t).$$

Вероятность противоположного события, т.е. того, что отказ не произойдет за время, меньшее, чем t , называется функцией дожития:

$$S(t) = 1 - F(t) = P(T \geq t).$$

Функция плотности вероятности описывает кривую распределения по срокам жизни:

$$f(t) = dF(t)/dt.$$

Интенсивность смерти (функция риска) характеризует риск отказа в момент t :

$$h(t) = f(t)/F(t).$$

Для выявления факторов, влияющих на вероятность того или иного исхода лечения, строится регрессионная модель, основанная на типе распределения данных времени жизни. Если основная задача исследования состоит в изучении качественного влияния воздействующих факторов на время наступления исхода, то выбор модели не имеет решающего значения. В случаях, когда задача связана с относительно “тонкими” вопросами зависимости времени жизни от воздействующих факторов, требуется выбор адекватной модели.

**Окончание статьи читайте
в следующем номере журнала**



Начинается подписка на 2005 год на научно-практический журнал “Атмосфера. Нервные болезни”

Подписку можно оформить
в любом отделении связи России и СНГ.
Журнал выходит 4 раза в год. Стоимость подписки на полгода
по каталогу агентства “Роспечать” – 44 руб.,
на один номер – 22 руб.
Подписной индекс 81610.



Научно-популярный журнал “Нервы” — это журнал для тех, кто болеет, и не только.

Основной задачей издания является информирование пациентов и среднего медицинского персонала о современных подходах к профилактике, диагностике и лечению заболеваний нервной системы. Читательская аудитория журнала “Нервы” — это в первую очередь пациенты, те люди, у кого имеются проблемы со здоровьем, в частности со стороны нервной системы, а также их родственники, которым необходима качественная информация для принятия решений по оказанию помощи своим близким. Журнал также будет интересен здоровым людям, заботящимся о своем здоровье и интересующимся достижениями современной медицины.