

КАК ОРГАНИЗОВАТЬ СОБСТВЕННЫЕ ИССЛЕДОВАНИЯ И НЕ ЗАПУТАТЬСЯ В ИНТЕРПРЕТАЦИИ ЧУЖИХ (часть 1)

Н.В. Жуков
ФГУФНКЦ
ДГОИ Росздрава

Все мы в своей практической деятельности в большей или меньшей степени опираемся на данные чужих исследований, опубликованные в журналах или представленные на научных конференциях. Редкий врач-онколог в настоящее время использует лишь рекомендации «старших товарищей» и циркуляры вышестоящих инстанций. В нашей работе всегда есть место творчеству, пускай и основанному на анализе результатов чужих исследований. На что же мы обращаем внимание при анализе представленных результатов? На $p < 0,0001$? На количество больных в исследовании? На частоту ремиссий? На что же действительно нужно обращать внимание, чтобы не пасть жертвой чужих заблуждений или чужого тщеславия? Как не навредить больным, приняв за истину результат, который хотя и опубликован в авторитетном зарубежном журнале, однако является явно предварительным (а автор забыл или не захотел об этом написать)?

Более того, на определенном этапе своей деятельности многие врачи приходят к выводу, что результатов исследований, проведенных другими, для них уже недостаточно. И начинают проводить свои.

Давайте вспомним, как же мы проводили свое собственное исследование, например, для диссертации или статьи в журнале. Начало исследованию, скорее всего, положил визит руководителя отделения (клинки) на международную конференцию (возможные варианты: ваш визит на конференцию, прочтение вами или вашим руководителем статьи в международном журнале). На этой конференции (в этом журнале) было представлено интересное для вас (вашего руководителя) исследование, результаты которого хотелось бы проверить (подтвердить, воспроизвести) на собственном опыте. Хотя, возможно, что инициировать исследование побудил и неразрешенный вопрос из клинической практики, ранее не освещавшийся в доступной литературе. Кроме того, поводом для инициации исследования достаточно часто является неожиданно полученный противоопухолевый эффект от терапии, которую назначили «от отчаяния». Следом за этим

обычно генерируется план исследования, представляющий собой намерение пролечить некоторое количество больных и посмотреть, что из этого получится. Возможно даже, что в начале исследования по образу и подобию международных исследований составляются критерии включения, исключения, план лечения (однако не столь жесткие или не слишком обязательные для дальнейшего исполнения — у нас не Запад, аудиторов нет). Контрольная группа, рандомизация — зачем? Если удастся получить достаточный, с нашей точки зрения, эффект, можно подобрать группу сравнения из тех, кого мы лечили ранее. Сколько включим больных? Ну.. Человек 20 (варианты 15, 30). Почему? Круглая цифра выглядит привычно (в прошлом исследовании тоже было 15, 20, 30), примерно столько больных проходит в нашем центре за месяц (полгода, год). Кроме того, весьма вероятно ограничение по количеству препарата (хоть это бы дали). Информированное согласие? Да это же практически стандартный метод — о нем писали в двух зарубежных статьях (вариант: этим больным все равно нечего терять, а мы им предлагаем лечение). Какая это фаза исследования? Раз препарат применялся раньше, то уж явно не первая. Рандомизации нет, значит не III. Тогда — вторая. Каковы цели исследования? Давайте возьмем пошире (частота ремиссий, переносимость, общая выживаемость, безрецидивная выживаемость, токсичность, влияние молекулярных маркеров и т.д.). В конце концов, все, что не будет иметь $p < 0,05$, потом можно и отбросить (а если $p < 0,05$ вообще не получится, то можно и добавить анализируемые показатели). И вот исследование началось. Однако лаборатория внезапно отказывается делать запланированный объем анализов (исследований). Оказалось, что мы заранее не согласовали с исследовательскими службами дополнительный объем исследований. Ничего, вместо компьютерной томографии вполне подойдет УЗИ, а измерять уровень креатинина так часто, как планировалось, совсем необязательно (мочится-то больной хорошо). Большинство больных, на которых мы так рассчитывали (у нас же проходит очень много больных ра-

ком молочной железы, кишки, желудка) вдруг оказываются нецениваемыми с точки зрения международных критериев (имеют только неизмеряемые очаги). Ничего, оценим так: опухоль сократилась, значит — частичная ремиссия (вариант — удалось хирургически убрать небольшую часть опухоли — стабилизация). Кончился препарат для комбинации с исследуемым — ничего, можно попробовать другую комбинацию. Вариантов масса.

Наконец, набрано требуемое количество больных (или закончился исследуемый препарат). Дальше идет анализ полученных результатов с использованием доступных (известных) статистических данных. Тест на нормальное распределение при использовании метода Стьюдента? Медиана длительности периода наблюдения 9 мес при анализе 5-летней выживаемости (один-то больной наблюдается уже 5,5 года)? Пустяки! Если получено $p < 0,05$, то все хорошо — лечение эффективно и рекомендовано для практического применения (варианты — требует дальнейшего изучения). Можно публиковать. Не получено магического $p < 0,05$, тоже не беда. Возможны варианты: ну есть же четкая тенденция (посмотрите, в нашей группе ремиссий 64,5%, а в историческом контроле всего 33%) — лечение явно эффективно. Доберем еще больных, и все получится. А вот эти больные, не ответившие на наше чудесное новое лечение, явно какие-то не такие (слишком больные, слишком залеченные и т.д.) — давайте исключим их из анализа, и ничего, что критерии исключения появились уже после завершения исследования. Еще один вариант: нет, это лечение неэффективно (ведь не удалось получить статистически значимой разницы на группе из 18 больных). Наиболее популярный вариант в диссертациях: давайте проведем «глубокий анализ» (разобьем на подгруппы, посмотрим, а влияло ли это, это и это). Такой подход даже приветствуется, разумеется, если в одном из многочисленных дополнительных анализов удастся получить заветное $p < 0,05$.

Исследование опубликовано, что дальше? А дальше мы обычно ждем новой международной конференции.

Со статистической достоверностью $p < 0,0000001$ уверен, что многим из вас эта картина знакома (автор имеет достаточно большую выборку собственных и чужих наблюдений на эту тему). Более того, и сам автор был активным соучастником многих таких «исследований», а также безоговорочно верил в чужие исследования, показавшие различия в частоте ремиссий в 15% с $p < 0,005$. Принимал участие и верил, пока не заинтересовался, что же по этому поводу думают люди, которые эти исследования создают по правилам и решают вопрос об их публикации в своих журналах. С обзором этих мыслей я и хочу вас познакомить. Возможно, это при-

несет вам удачу в виде возможности провести исследование, которое будет принято для публикации в международном журнале и значительно повысит ваш авторитет. Надеюсь также, что данная статья поможет вам более критично относиться к зарубежным публикациям и брать за основу для собственных исследований (или за руководство к клинической практике) только доказательные, хорошо организованные исследования.

Клинические исследования: какими они бывают

Проспективное или ретроспективное?

Клиническое исследование представляет собой эксперимент, направленный на определение эффективности и безопасности метода лечения, лекарства, метода обследования и т.д. Как и любой эксперимент, оно должно быть направлено на получение истинных данных, а не на подтверждение желаемого результата. Кроме того, эксперимент (в отличие от простого наблюдения) предусматривает проспективное планирование (заранее четко оговоренные правила проведения, наличие контроля). Только в обсервационных и ретроспективных исследованиях «экспериментатор» является пассивным наблюдателем. В таких исследованиях назначение лечения, стадирование, наблюдение за пациентом не подконтрольны исследователю, и возможность оценки этих данных ограничена (даже по хорошо организованным регистрам). По современным представлениям, статистические результаты, полученные в ретроспективных исследованиях, являются достаточно слабым аргументом для заключения о связи наблюдаемого результата с проводимым лечением (т.е. для вывода, что лечение эффективно или неэффективно). В рутинной практике (являющейся основой для составления баз данных) выбор лечения обычно производится на основании субъективной оценки прогноза пациента конкретным врачом, а спектр диагностических мероприятий по выявлению эффективности лечения может значительно варьировать от пациента к пациенту. Если в результате ретроспективного исследования (грубо говоря, при анализе архивных данных) вы вдруг выявили, что больные, получавшие при метастатическом раке молочной железы комбинацию доксорубицин+паклитаксел, имели худшую по сравнению со схемой CMF выживаемость, не торопитесь делать выводы. Сначала вспомните, почему одним больным вы назначали таксаны с антрациклинами, а другим — CMF. Вспомнили? Действительно, их не распределяли рандомизированно. Оказывается, доксорубицин+паклитаксел в основном получали молодые больные с висцеральным поражением, быстрым прогрессированием болезни, рецепторотрицательными, Her-2-позитивными опухолями. Так что же повлияло на различие в исходах лечения — сама те-

рапия или различия между группами, послужившие поводом для выбора лечения? В ретроспективных исследованиях зачастую именно факторы отбора, а не различия в проведенном лечении отвечают за различие в результатах. В связи с этим в международной онкологической практике наблюдательные исследования обычно применимы лишь в области эпидемиологии опухолей. Ретроспективные исследования также могут быть использованы в качестве отправной точки для проведения исследования проспективного (генерация идеи).

К сожалению, в России большинство клинических исследований в области онкологии являются именно ретроспективными. Во многом это оправдывается скудностью финансовой базы для проведения адекватных проспективных исследований. Для большинства молодых (и не очень) специалистов «рытье архива» является единственной возможностью написания диссертации или статьи, так как «старшие товарищи», отвечающие за распределение финансов и препаратов, не дают возможности провести собственное проспективное исследование. Однако даже если исследование проводится в проспективном формате, за его составление обычно отвечают все те же «старшие товарищи», считающие, что им лучше известно, как при минимуме затрат получить «достоверный» результат.

Об адекватном с точки зрения международных рекомендаций планировании собственных исследований (составление протокола и т.д.) мы напишем в следующем номере, если эта тема окажется интересной. В этой же статье будет представлен обзор видов исследований, проводимых в онкологии.

Исследования I фазы

Исследования данного типа являются наиболее ответственными и «ресурсоемкими», так как представляют собой первый опыт использования препарата у людей. Исследование предполагает сверхтщательное наблюдение за больным (для выявления любых побочных эффектов), готовность клиники бороться с необычными побочными эффектами лечения и т.д. Большинство истинных исследований I фазы зарубежные фирмы-производители предпочитают проводить в своих странах (в России проходят лишь единичные исследования I фазы зарубежных препаратов). Однако и наши исследователи создают свои препараты, поэтому у вас есть реальный шанс поучаствовать в создании или «обкатке» протокола I фазы.

Упрощенно, задачей большинства исследований I фазы является определение дозы нового препарата, которая будет применяться в дальнейших исследованиях более высокого уровня, и определение видов токсичности, которые встречаются при использовании данного препарата у людей. Эффективность лечения никогда не является первичной

целью исследования I фазы. В связи с этим данные исследования не предусматривают включения больных с каким-либо определенным видом опухоли (раз цель — определение переносимости и побочных эффектов, то подразумевается, что больные с любым видом опухоли переносят лечение одинаково). Отбор в исследование ведется совсем по другим параметрам. Учитывая, что эффективность и токсичность данного вида терапии являются лишь предположительными (на основании исследований на животных), с этической точки зрения в исследовании I фазы должны включаться только больные с распространенной опухолью, не имеющие шансов на достижение клинически значимого противоопухолевого эффекта при использовании стандартной терапии, доступной в настоящее время. Лишь единичные препараты, появившиеся на заре химиотерапии (антрациклины, цисплатин) показали адекватную эффективность уже на этапе исследований I фазы. Однако в отличие от симптоматического лечения, которое является единственной альтернативой для таких больных, шанс на достижение противоопухолевого эффекта существует при исследовании любого нового препарата, что оправдывает проведение исследований I фазы.

Существует и второе условие, без соблюдения которого проведение исследований I фазы стало бы бессмысленным. Несмотря на крайне вероятное отсутствие терапевтических перспектив, больные должны иметь сохранную функцию основных органов и систем. Такая «сохранность» пациента необходима для того, чтобы обеспечить адекватную переносимость лечения. В случае если в исследование включаются больные, имеющие серьезные нарушения функции органов, выбранная для последующих испытаний доза, скорее всего, будет неадекватно низкой. Использование такой дозы в исследованиях II фазы может послужить поводом для признания неэффективным препарата, который при применении в адекватных дозах мог бы оказаться клинически успешным.

Существует несколько типов исследований I фазы в онкологии. Наиболее частым является исследование цитотоксических препаратов. Такие исследования обычно предусматривают начало с очень малой дозы препарата, от которой не ожидается развития серьезной токсичности ни у одного пациента. Стартовая доза обычно составляет 1/10 от летальной дозы (рассчитанной в мг/м²) наиболее чувствительного к данному цитостатику животного, использовавшегося в доклинических исследованиях [1]. По ходу исследования доза увеличивается от одной когорты пациентов к другой по заранее запланированной ступенчатой методике. Для каждого дозового уровня обычно используются когорты пациентов от 3 до 10 человек. Эскалация дозы произво-

дится только после того, как пройдет определенное время, необходимое для реализации потенциальных острых побочных эффектов в когорте, получавшей предыдущую дозу. Эскалация до следующего дозового уровня происходит, если во всей исследуемой когорте не отмечено дозолимитирующей токсичности (ДЛТ). Под ДЛТ обычно понимают:

- негематологическую токсичность III степени¹ и выше (за исключением алопеции);
- развитие нейтропенической инфекции (нейтрофилы $< 0,5 \times 10^9/\text{л}$ и температура тела $> 38,3^\circ\text{C}$ или клинически документированная инфекция);
- нейтропения IV степени ($< 0,5 \times 10^9/\text{л}$), длящаяся 7 дней;
- тромбоцитопения III степени и выше \pm геморрагические осложнения.

Если ДЛТ отмечается у 33% больных или менее (1 или 2 больных в когорте из 3 или 6 человек соответственно), то производится набор дополнительных больных (обычно трех) на данном дозовом уровне. Если дополнительных случаев ДЛТ не отмечается, то начинается исследование на следующем уровне, при появлении дополнительных случаев ДЛТ эскалация доз останавливается. Исследование останавливают и в том случае, если частота ДЛТ в когорте превышает 33%. Эскалация дозы на последующих курсах лечения для одного и того же пациента не производится (т.е. для следующего дозового уровня набирается новая когорта пациентов).

Для последующих исследований II фазы рекомендуется последняя доза, при которой частота ДЛТ была $< 33\%$. Эскалация доз обычно проводится по модифицированной шкале Fibonacci [2]. Второй дозовый уровень превышает первый в 2 раза, третий превышает второй на 67%, четвертый на 50% больше третьего, пятый на 40% больше четвертого, а все последующие уровни превышают предыдущие на 33%. К сожалению, единственным преимуществом такого подхода является его высокая безопасность. Однако эта безопасность достигается ценой ряда значительных недостатков:

- часто большое количество пациентов получают слишком низкие (субтерапевтические) дозы исследуемого препарата;
- необходимо длительное время для набора пациентов и окончания исследования;
- исследования, основанные на таком дизайне, дают слишком мало информации о различиях от пациента к пациенту и кумулятивной токсичности.

В связи с этим все более широко стали применяться новые дизайны исследований I фазы. Один из них, получивший название дизайна ускоренного

определения дозы (accelerated titration designs), предусматривает эскалацию дозы исследуемого препарата от курса к курсу у одного и того же пациента. На каждый дозовый уровень включается только один больной (до достижения токсичности II степени и выше) [3]. Такой подход позволяет значительно сократить число пациентов, необходимых для исследования, а также уменьшить число больных (и курсов лечения), для которых используются субтерапевтические дозы препарата. Недостатком такого подхода является некоторое повышение риска достижения токсичности III—IV степени [3].

В связи с развитием целевой терапии все более широко стали проводиться и так называемые исследования IV фазы. Кроме стандартных задач классического исследования I фазы они дополнительно изучают взаимосвязь между дозой препарата и развитием основного биологического эффекта (например, эффективность блокады рецепторов или связи антител с эндотелиальным фактором роста сосудов). Такие исследования обычно страдают двумя изъянами. Во-первых, когорты из 3—6 пациентов может оказаться явно недостаточно для соотношения дозы препарата и биологического эффекта (принимая во внимание различия между больными и неточность измерения биологического эффекта). Во-вторых, далеко не всегда достижение предполагаемого биологического эффекта реализуется в дальнейшем в противоопухолевое воздействие. К сожалению, такие исследования зачастую дают недостаточно дополнительной информации (по сравнению с классическими исследованиями I фазы) о дозе препарата, которая должна применяться в дальнейшем. Однако полученная информация иногда все же позволяет судить о потенциальной эффективности препарата. Если при достижении ДЛТ удается добиться лишь минимального желательного биологического эффекта (блокада 5% рецепторов и т.д.), то ожидать клинического эффекта в последующих исследованиях, скорее всего, бесперспективно. Другой вариант — длительная блокада 100% целевых рецепторов может свидетельствовать об отсутствии необходимости эскалации дозы несмотря на то, что ДЛТ не достигнута.

Отдельным видом являются сравнительные исследования I фазы. С ранее описанными видами исследований I фазы их объединяет лишь основная цель (отработка дозы). Пример задачи такого исследования: выяснить, в какой последовательности должно осуществляться введение цисплатина и паклитаксела в комбинации для того, чтобы оба препарата могли использоваться в максимальной дозе. В связи с небольшим объемом выборки, характерной для I фазы исследования, максимальная переносимая доза в таких исследованиях обычно определяется крайне неточно. В сочетании с нерандо-

¹Здесь и далее — степень по критериям NCI CTC v.3.

мизированным характером исследований I фазы это означает, что такие исследования могут дать полезную информацию лишь в случае, если между методиками ожидается большое различие.

Исследования II фазы

Селекция пациентов

В отличие от исследований I фазы, целью которых является отработка дозы, исследования II фазы направлены на определение непосредственной противоопухолевой эффективности при конкретных типах злокачественных новообразований. В связи с этим в данные исследования включают уже пациентов с определенным гистологическим типом опухоли (рак молочной железы, кишки и т.д.). Основной задачей исследований II фазы является получение информации о том, при каком гистологическом типе опухоли препарат будет потенциально эффективен (другими словами, имеет ли смысл проводить в дальнейшем большое исследование III фазы для подтверждения эффективности препарата X при опухоли Y).

Мерой эффективности препарата в исследованиях II фазы чаще всего является непосредственная эффективность (частота ремиссий). В связи с этим результаты исследования II фазы сами по себе практически никогда не способны дать достаточно данных для однозначного признания препарата эффективным (и, соответственно, не должны быть поводом для начала использования предложенного препарата или комбинации в широкой клинической практике). Проводя исследования II фазы или анализируя их результаты, необходимо осознавать, что непосредственный противоопухолевый ответ не является прямым измерением (синонимом) выигрыша пациента в результате проводимого лечения. Достижение целевой частоты ремиссий является лишь поводом для продолжения исследования препарата. Лечение, в результате которого достигается ремиссия, не является обязательно выигрышным для пациента. Уменьшение асимптоматичного метастаза в печени на 50% абсолютно безразлично для пациента, если не приводит к продлению жизни или уменьшению симптомов болезни. Доказать же выигрыш в выживаемости на материале исследований II фазы практически невозможно. Так любимый многими анализ, который показывает, что ответившие на лечение больные живут дольше, чем не ответившие, порочен, если является основанием для вывода, что лечение продлевает жизнь [4, 5]. Во-первых, достижение ремиссии, по определению, предсказывает большую химиочувствительность опухоли и часто сочетается с большей продолжительностью жизни конкретного пациента вне зависимости от вида лечения. Во-вторых, больные, ответившие на лечение, могут иметь другие (дополнительные) благоприятные прогностические фак-

торы. Наконец, лечение само по себе (за счет токсичности или выработки химиорезистентности) может уменьшать выживаемость больных, не ответивших на него, не изменяя продолжительности жизни больных, на лечение ответивших. Для демонстрации выигрыша в выживаемости необходимо доказать, что вся группа больных, получавшая исследуемый препарат (и ответившие, и не ответившие на лечение), живет дольше, чем адекватная контрольная группа, получавшая лечение сравнения. Учитывая, что большинство исследований II фазы не имеет внутренней контрольной группы, доказать такой выигрыш в выживаемости достаточно проблематично.

Для клинических исследований II фазы требуется набор пациентов, которые с большой степенью вероятности ответят на проводимое лечение сокращением опухоли (непосредственный противоопухолевый эффект — ремиссия), однако не имеют реального шанса на значительное продление жизни (и, тем более, шанса на излечение) при проведении стандартной на данный момент терапии. Для получения адекватных результатов исследования (выявления перспективных для дальнейшего исследования препаратов) пациенты должны иметь максимально сохранное общее состояние и минимум предшествующей химиотерапии. Соблюдение запланированной дозовой интенсивности (введение полных доз препарата и соблюдение интервала между курсами лечения) у больных, ранее получивших множественные курсы химиотерапии, обычно крайне затруднено. Кроме того, в ходе предшествующего лечения опухолевые клетки обычно вырабатывают «перекрестную» резистентность ко многим цитостатикам (в том числе возможна выработка резистентности и к исследуемому препарату). В связи с вышеперечисленным отсутствие противоопухолевой активности (или значительная токсичность) у ранее активно предлеченных больных может не свидетельствовать об отсутствии перспектив препарата у больных с менее отягощенным «химиотерапевтическим анамнезом». Данное наблюдение хорошо подтверждается результатами исследований II фазы по применению этопозида у больных с мелкоклеточным раком легкого [6, 7].

Оптимальное время включения больных в исследования II фазы во многом зависит от химиочувствительности опухоли, при которой планируется исследовать препарат. Если препарат испытывается у больных с малочувствительными к химиотерапии опухолями (меланомы, рак почки, саркомы мягких тканей и т.д.), в исследования II фазы рекомендуется включать больных, ранее не получавших химиотерапии. Если перед включением в исследование II фазы пациенты с малочувствительными опухолями уже получали цитостатическое лечение,

то испытываемый препарат (впрочем, как и любой уже зарегистрированный цитостатик) скорее всего будет иметь минимальную противоопухолевую активность или не будет иметь ее вовсе. При более чувствительных опухолях (рак молочной железы, мелкоклеточный рак легкого, неходжкинские лимфомы) желателен включение во II фазу исследования больных, получивших не более одной линии лечения по поводу распространенного заболевания. Несмотря на то что существует искушение еще полечить больных уже зарегистрированными препаратами перед тем, как проводить экспериментальное лечение, именно вышеописанный подход является оптимальным. Раннее включение в исследования II фазы позволяет спасти пациентов от использования исследуемого препарата на этапе болезни, при котором риск развития токсичности от его применения намного больше, чем шанс на достижение клинически значимого противоопухолевого эффекта. С другой стороны, раннее включение в исследования II фазы не препятствует дальнейшему использованию зарегистрированных цитостатиков после прогрессирования заболевания на фоне лечения исследуемым препаратом.

Существует значительная путаница в отношении исследований II фазы, в которых новый препарат используется в монотерапии или в комбинации с ранее использовавшимися схемами лечения. Формально и те и другие называются исследованиями II фазы, так как ограничены пациентами с определенным диагнозом и не имеют внутренней контрольной группы, однако при организации и анализе они имеют ряд существенных различий.

Исследования одного препарата

Целью большинства исследований, предусматривающих монотерапию новым препаратом, является просто ответ на вопрос: имеет ли препарат потенциальную активность в отношении данного типа опухоли? При использовании монотерапии новым препаратом для ответа на этот вопрос обычно достаточно определения частоты ремиссий. Если этот показатель выше определенного уровня — препарат имеет перспективы, если ниже, то не имеет.

В соответствии со своими задачами исследование II фазы должно включить достаточное количество пациентов для того, что бы «отсечь» неэффективные препараты, но при этом не «пропустить» новый перспективный метод лечения. Таким образом, с одной стороны, если исследуемое лечение окажется неэффективным, то его воздействию должно подвергнуться как можно меньше больных. С другой — слишком раннее прекращение исследования может не позволить выявить перспективный препарат. Исходя из этого, одним из наиболее популярных дизайнов исследований II фазы является двухступенчатый [8].

Перед началом исследования определяются две величины — уровень отсутствия интереса к исследуемому препарату (p_0) и целевой уровень активности (p_1). Под p_0 понимают частоту ремиссий, при которой дальнейшее исследование препарата нецелесообразно (уровень отсутствия интереса), под целевым уровнем p_1 — частоту ремиссий, при которой препарат представляет значительный интерес для дальнейшего исследования. Что это значит? Например, мы будем испытывать препарат X при раке поджелудочной железы. Мы знаем, что опухоль крайне химиорезистентна, поэтому отсутствие интереса к препарату может быть только в случае, если частота ремиссий не достигает даже минимального уровня (например, 5%). Это и будет p_0 . Каков же уровень, при котором препарат может быть интересен? Разумеется, хотелось бы 50%. Однако шанс на достижение такого эффекта при раке поджелудочной железы ничтожен, и если мы поставим такую планку эффективности, то будем отвергать все исследуемые препараты. Обычно ожидаемая разница в непосредственной эффективности при испытании нового препарата не превышает 15—20%. Если взять более узкий диапазон, то понадобится большее количество больных для ее выявления, если более широкий — см. выше. Таким образом, целевой уровень эффективности p_1 составит $p_0 +$ ожидаемая разница, т.е. 20 или 25%. Какой диапазон различий выбрать — это ваше дело, однако необходимо помнить — чем меньше разница, тем больше нужно больных для того, чтобы отвергнуть неэффективный препарат. Если речь идет о более чувствительной опухоли, также не рекомендуется занижать уровень p_0 . Например, если поставить p_0 для рака молочной железы в 5%, то, скорее всего, вы примете для дальнейшего исследования любой препарат (включая святую воду), однако каково вам будет проводить III фазу исследований со святой водой?

Но вернемся к раку поджелудочной железы. Предположим, вы выбрали $p_0 = 5\%$, а $p_1 = 20\%$. Дальше необходимо определить приемлемую вероятность альфа- и бета-ошибки. Эти термины означают следующее:

— альфа-ошибка — вероятность принять для дальнейшего исследования (признать потенциально эффективным) лечение с частотой ремиссии ниже p_0 ;

— бета-ошибка — вероятность не принять для дальнейшего исследования (признать неэффективным и остановить дальнейшие исследования) лечение с частотой ремиссий выше p_1 .

Обычно величина вероятности альфа- и бета-ошибки варьирует от 0,1 до 0,05 (чем меньше вероятность ошибки, тем выше шанс сделать правильный выбор, однако тем больше больных нужно включить в исследование).

Дизайн двухступенчатого исследование II фазы при $p_1 - p_0 = 15\%$, вероятности альфа-ошибки 0,05 и бета-ошибки 0,1

| $p_0, \%$ | $p_1, \%$ | Препарат неэффективен, если | |
|-----------|-----------|-----------------------------|------------|
| | | $\leq r_1/n_1$ | $\leq r/n$ |
| 5 | 20 | 1/21 | 4/41 |
| 10 | 25 | 2/21 | 10/66 |
| 20 | 35 | 8/37 | 22/83 |
| 30 | 45 | 13/40 | 40/110 |
| 40 | 55 | 19/45 | 49/104 |
| 50 | 65 | 22/42 | 60/105 |
| 60 | 75 | 21/34 | 64/95 |
| 70 | 85 | 18/25 | 61/79 |
| 80 | 95 | 16/19 | 37/42 |

Тео́рия и пра́ктика клинических исследований

Итак, мы определились, что $p_0=5\%$, $p_1=20\%$, вероятность альфа-ошибки — 0,05, бета-ошибки — 0,1. И после этого, оказывается, ничего сложного уже нет. Уже составлены специальные таблицы, которые позволяют определить количество пациентов, которых нужно включить в исследование, чтобы с заданной вероятностью ошибки выявить лекарства, имеющие перспективы для дальнейшего исследования, и отсеять лекарства, таких перспектив не имеющие. К сожалению, в связи с недостатком объема журнала мы не будем приводить эти таблицы целиком (при желании вы можете получить их из первоисточника [8] или отправить заявку на наш сайт, и таблицы будут опубликованы в продолжении этой статьи в следующем номере). Приведем лишь таблицу определения числа больных, необходимых для различия $p_1 - p_0$ в 15% с альфа-ошибкой 0,05 и бета-ошибкой 0,1 (см. таблицу). Учитывая, что наше исследование является двухступенчатым, в таблице приведены два соотношения. Под n_1 подразумевается максимальное число больных, которое может быть включено в первую ступень исследования, под r_1 — число больных, ответивших на лечение ремиссией. Что это значит на практике? Мы начали лечить больных раком поджелудочной железы препаратом X. Если до набора 21 больного (n_1) хо-

тя бы 2 пациента достигли ремиссии (r_1), то препарат признается перспективным и переходит на II ступень исследования. Причем неважно, какие по счету больные достигнут ремиссии — первые двое, второй и семнадцатый, главное, чтобы 2-я ремиссия была достигнута до включения 21 больного. Как только достигается 2-я ремиссия, препарат переходит на II ступень исследования. Если же после набора 21-го больного не достигнуто двух ремиссий, то препарат признается не заслуживающим внимания и исследование прекращается. На II ступени алгоритм такой же, только для признания препарата перспективным (уже для III фазы исследования) необходимо достижение более r ремиссий до включения n больных. Как только требуемое число ремиссий достигнуто, исследование прекращается и препарат рекомендуется для дальнейшего изучения в III фазе. Если после набора 41 больного отмечено менее 5 ремиссий ($r/n \leq 4/41$), то препарат не имеет перспектив. Все очень просто и позволяет получить осознанные с точки зрения статистики и здравого смысла результаты, однако почти не находит применения в отечественной практике.

Существуют и другие дизайны исследований, однако на них мы останавливаться не будем.

Окончание в следующем номере.

ЛИТЕРАТУРА

- Leventhal B.G., Wittes R.E. Research methods in clinical oncology. N Y, Raven Press; 1988.
- Schneiderman M.A. Mouse to man: statistical problems in bringing a drug to clinical trial. In: Proceedings of the fifth Berkeley symposium on mathematical statistical probability. Berkeley, University of California; 1967. p. 855.
- Simon R., Freidlin B., Rubinstein L. et al. Accelerated titration designs for phase I clinical trials in oncology. J Natl Cancer Inst 1997;89:1138.
- Anderson J.R., Cain K.C., Gelber R.D. Analysis of survival by tumor response. J Clin Oncol 1983;1:710.
- Simon R., Makuch R.W. A nonparametric graphical representation of the relationship between survival and the occurrence of an event: application to responder versus non-responder bias. Stat Med 1984;3:1.
- Ettinger D.S. Evaluation of new drugs in untreated patients with small-cell lung cancer: its time has come. J Clin Oncol 1990;8:374.
- Wittes R.E., Marsoni S., Simon R. et al. The phase II trial. Cancer Treat Rep 1985;69:1235.
- Simon R. Optimal two-stage designs for phase II clinical trials. Control Clin Trials 1989;10:1.