

## Использование косвенных критериев оценки в клинических испытаниях: не ошибаемся ли мы?

Источник: ж/л **Международный журнал медицинской практики** 1997г.

---

Т.Р. Флеминг, Д.Л. ДеМетс;

*Translated, with permission of the American College of Physicians, from: Fleming T.R., DeMets D.L. Surrogate end points in clinical trials: are we being misled? Ann Intern Med 1996;125:605-13.*

Клинические испытания 3 фазы, в которых изучается влияние новых видов вмешательств на жизненно важные для больных клинические исходы (смерть, потеря зрения и т.п.), часто требуют проведения длительной работы с участием большого количества людей. Еще недавно большой интерес вызывало использование "косвенных" критериев оценки, таких как уменьшение размеров опухоли или изменение уровня холестерина, артериального давления, количества CD4-лимфоцитов (лимфоцитов-хелперов) и других лабораторных показателей; предполагалось, что это снизит стоимость и продолжительность клинических испытаний. Теоретически, для того чтобы косвенный критерий оценки действительно заменял собой истинный клинический исход, результат воздействия вмешательства на этот критерий должен быть таким же, как и на истинный клинический исход. Практически же это требование часто нарушается. Среди различных объяснений данного явления упоминается возможность того, что патологический процесс может влиять на клинический исход, развиваясь несколькими патогенетическими путями, часть из которых не связана с косвенным критерием; воздействие вмешательства на эти пути отличается от его влияния на косвенный критерий. Еще более вероятно то, что вмешательство может влиять на клинический исход с помощью механизмов, полностью неизвестных и непредсказуемых. На примере различных заболеваний мы проиллюстрируем, как косвенные критерии оценки привели к ошибкам относительно реального воздействия лечения на здоровье больных.

Косвенные критерии оценки могут быть полезны в клинических испытаниях 2 фазы для выявления того, является ли новый вид вмешательства биологически активным, а также для принятия принципиального решения о достаточной перспективности нового вида вмешательства, оправдывающего проведение крупного окончательного исследования с использованием клинически значимых исходов. В окончательных клинических испытаниях 3 фазы, за исключением редких случаев, когда обоснованность применения косвенного критерия оценки неопровержимо доказана, в качестве основного критерия оценки должен использоваться истинный клинический исход.

Клинические испытания - стандартный научный метод для оценки нового биологического вещества, препарата, устройства или процедуры, предназначенных для профилактики или лечения заболеваний у человека. Испытание 3 фазы планируется так, чтобы можно было выявить клинические преимущества нового вещества и возможные побочные эффекты; оно должно быть окончательной проверкой пригодности данного препарата [1-3]. В испытаниях 3 фазы основным критерием оценки должно быть клиническое проявление, которое имеет существенное значение для больного и которого он хотел бы избежать. В качестве примеров можно назвать смерть, потерю зрения, клинические проявления синдрома приобретенного иммунодефицита (СПИД), необходимость применения искусственной вентиляции легких и другие явления, существенно снижающие качество жизни. Исследования, в которых изучаются такого рода клинические исходы, требуют значительных материальных и временных затрат. В результате в последнее время уделяется большое внимание альтернативным способам решения проблемы или косвенным критериям оценки. Это делается для уменьшения стоимости исследований 3 фазы и уменьшения сроков их проведения [4-17]. В соответствии с определением R.J. Temple [13], "... косвенный критерий оценки в клиническом испытании - лабораторный показатель или симптом, который заменяет клинически значимый исход, прямо характеризующий самочувствие больного, его функциональное состояние и

*выживание. Изменения косвенного критерия оценки, вызываемые лечением, должны отражать изменения клинически значимого критерия (исхода)".*

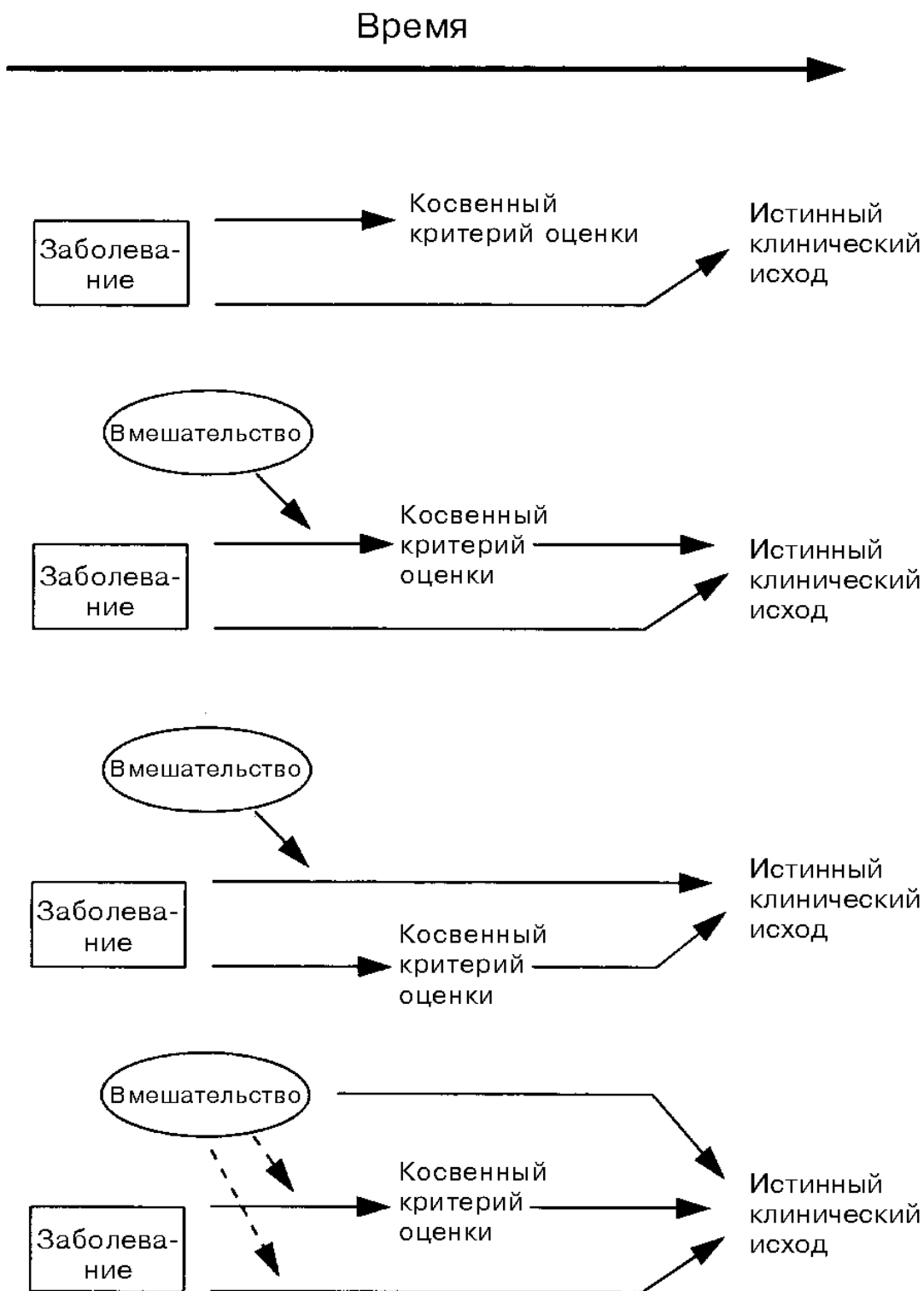
В качестве примеров косвенных критериев оценки можно назвать увеличение количества клеток CD4 или снижение титра содержания вируса в исследованиях, посвященных лечению ВИЧ-инфекции или СПИДа; подавление желудочковых аритмий, снижение уровня холестерина или артериального давления в кардиологических исследованиях; уменьшение размеров опухоли в исследованиях, посвященных лечению рака. Косвенные критерии оценки крайне редко, а то и вовсе не отражают важные клинические исходы в клинических испытаниях 3 фазы. Ниже мы приводим обзор тех требований, которым должны удовлетворять косвенные критерии оценки, чтобы они могли быть использованы для замены клинически значимого исхода.

### **Требования к косвенному критерию оценки**

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

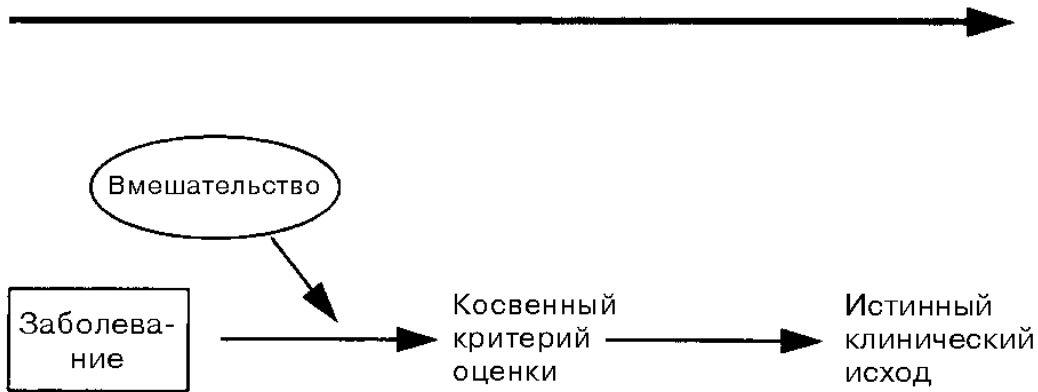
R.L. Prentice [11] разработал принципы определения достоверности косвенного критерия оценки в клинических испытаниях 3 фазы. В соответствии с этими принципами, косвенный критерий должен коррелировать с истинным клиническим исходом и полностью отражать влияние терапии на клинический исход. Первое требование обычно легко выполнить, а второе - нет. Например, в нескольких недавних исследованиях, посвященных ВИЧ-инфекции и СПИДу [14-24], показано, что второе требование не выполняется, если в качестве косвенного критерия оценки прогрессирования симптоматики СПИДа или смертности от него используется количество клеток CD4.

Ряд факторов (рис. 1) может объяснить неудачи при использовании косвенных критериев оценки. Косвенный критерий может коррелировать с прогрессированием заболевания (рис. 1А), но не быть вовлеченным в тот патофизиологический процесс, который приводит к определенному клиническому исходу. Некоторые патогенетические пути могут иметь причинную связь с клиническим исходом и не иметь такой связи с косвенным критерием. Из всех патогенетических звеньев заболевания, влияющих на истинный клинический исход, вмешательство может влиять только на путь, связанный с косвенным критерием (рис. 1Б), или на путь или пути, не связанные с ним (рис. 1В). Очень важно, что вмешательство может также влиять на истинный клинический исход посредством неизвестных механизмов, не связанных с патогенезом заболевания (рис. 1Г). Эффекты, обусловленные известными механизмами вмешательства, могут в значительной степени сводиться к нулю непредусмотренными, неожиданными или нераспознанными механизмами [25].



На рис. 2 представлена схема, отражающая ситуацию, при которой использование косвенного критерия наиболее обосновано. Характерно то, что косвенный критерий связан с одним из патогенетических путей заболевания и воздействие вмешательства на истинный клинический исход опосредовано его воздействием на косвенный критерий. Даже при такой идеальной ситуации использование косвенного критерия оценки может привести к неверным выводам. Воздействие вмешательства на истинный клинический исход может быть недооценено, если имеются существенные помехи при определении его воздействия на косвенный критерий. С другой стороны, воздействие на истинный исход может быть переоценено, если воздействие на косвенный критерий статистически значимо, но по выраженности и по продолжительности недостаточно, чтобы повлиять на истинный клинический исход. Эта переоценка может произойти, например, при проводящемся в настоящее время определении активности ингибиторов протеаз у больных с ВИЧ-инфекцией, когда влияние на косвенный критерий (уровень вирусной РНК в крови) статистически значимо, но недостаточно продолжительно по времени.

## Время



Анализ недавних испытаний с использованием косвенных критериев действует отвращающе, поскольку выявлено множество случаев, когда биологические показатели коррелировали с клиническими исходами, но не позволяли прогнозировать воздействие лечения на клинический исход. Далее рассматриваются примеры неудачного использования косвенных критериев. Можно только высказывать предположения о причинах данных неудач, поскольку даже при ретроспективном анализе наши представления о патогенетических путях и о механизмах воздействия вмешательства являются неполными. В таблице эти предположения сопоставляются с возможными причинами, представленными на рис. 1.

Заболевание и вмешательство	Критерии оценки		Причины, представленные на рисунке 1*			
	косвенные	клинические	А	Б	В	Г
<b>Кардиологические заболевания</b>						
<b>Аритмия</b>						
Энкаинид, флекаинид	Желудочковые аритмии	Выживание		+		++
Хинидин, лидокаин	Мерцание предсердий	Выживание		+		++
<b>Застойная сердечная недостаточность</b>						
Милринон, флосеквиан	Сердечный выброс; фракция изгнания	Выживание		+		++
<b>Повышение уровня липидов</b>						
Фибраты; гормоны; диета; ловастатин	Уровень холестерина	Выживание		+		++
<b>Повышение артериального давления</b>						
Блокаторы кальциевых каналов	Артериальное давление	Инфаркт миокарда; выживание		+		++
<b>Рак</b>						
<b>Профилактика</b>						
Финастерид	Результаты биопсии предстательной железы	Симптомы рака; выживание	++#			
<b>Стадия клинических проявлений</b>						
Фторурацил + лейковорин	Уменьшение опухоли	Выживание		+		++
<b>Другие заболевания</b>						
<b>ВИЧ-инфекция или СПИД</b>						
Антиретровирусные препараты	Число CD4; титр вируса	Симптомы СПИДа; выживание		+	+	+
<b>Остеопороз</b>						
Фторид натрия	Минеральная плотность костной ткани	Переломы костей		+		+
<b>Прогрессирующий гранулематоз</b>						
Интерферон-γ	Гибель бактерий; выработка перекиси	Тяжелая инфекция				++

### Косвенные критерии оценки в кардиологии

#### *Подавление желудочковых аритмий и поддержание синусового ритма*

Классический пример ненадежности косвенных критериев оценки - использование такого показателя как снижение частоты желудочковой экстрасистолии вместо снижения сердечно-сосудистой смертности. Наличие желудочковой аритмии почти в четыре раза повышает риск смерти от сердечно-сосудистых осложнений, в частности, риск внезапной смерти [26, 27]. Было высказано

предположение, что подавление желудочковых аритмий после инфаркта миокарда уменьшит уровень смертности. Показано, что три новых препарата (энкаинид, флекаинид, морицизин) эффективны для подавления аритмии; Администрация США по контролю за продовольствием и лекарственными препаратами (FDA) разрешила применять их у больных с тяжелыми или опасными для жизни нарушениями сердечного ритма. До проведения исследований по изучению влияния аритмий на уровень смертности в США более 200 000 человек в год принимали эти препараты. В исследовании CAST (Cardiac Arrhythmia Suppression Trial) [26-28] оценивалось, как эти три препарата влияют на выживаемость у больных с инфарктом миокарда при наличии как минимум 10 желудочковых экстрасистол в минуту. Первые результаты CAST были ошеломляющими. Испытание энкаинида и флекаинида было прекращено на ранних этапах после того, как у больных, получавших эти препараты, было зарегистрировано 33 случая внезапной смерти, в то время как в группах, получавших плацебо, таких случаев было всего 9. Всего умерли 56 больных, получавших энкаинид и флекаинид, и 22 больных, получавших плацебо. После получения окончательных данных оказалось, что число случаев внезапной смерти равнялось соответственно 43 и 16, а общее число летальных случаев - соответственно 63 и 26. Последующие результаты CAST свидетельствовали об увеличении уровня смертности и у больных, получавших морицизин [28].

Два других примера имеют отношение к профилактике возникновения аритмий. Для поддержания синусового ритма после кардиоверсии у больных с мерцанием предсердий использовался хинидин [29]. Мета-анализ результатов 6 испытаний показал, что прием хинидина способствовал сохранению синусового ритма в течение 1 года у 50% больных по сравнению с 25% больных, не получавших его. Однако уровень смертности на фоне приема хинидина увеличился с 0,8 до 2,9%. Предотвращение рецидивов мерцательной аритмии - ценное качество препарата, но оно явно проигрывает на фоне повышения уровня смертности. Такой же недостаток был выявлен у лидокаина; мета-анализ показал, что уменьшение риска возникновения желудочковой тахикардии на одну треть сопровождается повышением риска смерти в такой же степени [30, 31].

### ***Переносимость физической нагрузки при застойной сердечной недостаточности***

У больных с застойной сердечной недостаточностью снижен сердечный выброс, имеется характерная одышка, снижена толерантность к физической нагрузке и высок риск смертельного исхода. Годовая смертность при тяжелой застойной сердечной недостаточности составляет от 20 до 40%. Снижение физической активности - результат уменьшения сердечного выброса и, в меньшей степени, повышения давления в легочных сосудах. При этой патологии в качестве косвенных критериев оценки эффективности новых препаратов используют величину сердечного выброса и фракции изгнания, а в качестве промежуточных критериев - толерантность к физической нагрузке и нарастание симптоматики. Некоторые виды лечения, влияющие на эти критерии, повышают выживаемость [32-35], другие либо неэффективны, либо даже снижают ее.

Диуретики и дигоксин способствуют уменьшению симптомов застойной сердечной недостаточности. В печати до сих пор нет данных о влиянии этих препаратов на выживаемость, хотя группа, созданная для исследования эффективности препаратов наперстянки [36], недавно обнародовала свои результаты, которые свидетельствуют об отсутствии улучшения выживаемости (Американская коллегия кардиологов, март 1996 г. Неопубликованные данные).

Одним из первых инотропных препаратов, предложенных для лечения застойной сердечной недостаточности, был милринон. Выполненные исследования показали, что он повышал сердечный выброс и толерантность к физической нагрузке. Милринон является инотропным средством (как и дигоксин), увеличивающим силу сокращения сердца. Поскольку FDA сочла, что такого рода препараты могут обладать побочными эффектами при длительном приеме (как было в случае использования инотропных  $\beta$ -агонистов), было проведено рандомизированное двойное слепое плацебо-контролируемое исследование для определения влияния милринона на уровень смертности. В этом исследовании, известном под названием PROMISE (Prospective Milrinone Survival Evaluation), было продемонстрировано увеличение общей смертности у больных, получавших милринон, по сравнению с получавшими плацебо (29 и 23% соотв.;  $p=0,04$ ) [37].

Другой препарат, флозеквинан, созданный для лечения застойной сердечной недостаточности, является вазодилататором, уменьшающим рабочую нагрузку сердца. Препарат был условно разрешен FDA, поскольку он мог увеличивать толерантность к физической нагрузке в тех случаях, когда другие препараты (в том числе, диуретики и ингибиторы ангиотензин-превращающего фермента) были неэффективны или больные не могли переносить лечение по полной схеме. Это условное разрешение было действительно вплоть до завершения испытаний, которые должны были выявить влияние флозеквинана на общую смертность, в частности, исследования PROFILE (Prospective Flosequinan Longevity Evaluation). В нем [38] были получены убедительные доказательства того, что флозеквинан увеличивает общую смертность (относительный риск составил 1,43), что заставило производителя изъять препарат из продажи.

Хотя повышение сердечного выброса, фракции изгнания и толерантности к физической нагрузке коррелирует с более длительным выживанием при застойной сердечной недостаточности, улучшение этих показателей не является надежным прогностическим критерием влияния лечения на уровень смертности. Конечно, повышение толерантности к физической нагрузке само по себе является важным клиническим исходом и могло бы рассматриваться как желательный эффект, несмотря на возможное отрицательное влияние препарата на смертность. Трудно утверждать, однако, что врач или больной смогут принять верное решение, не зная о влиянии лечения как на симптоматику, так и на смертность. Влияние препарата на уровень смертности можно оценить только при длительных исследованиях.

### ***Снижение уровня липидов***

Хотя давно известно, что концентрация липидов в крови (особенно общего холестерина, его фракций и триглицеридов) позволяет достоверно прогнозировать уровень смертности от сердечно-сосудистых заболеваний, некоторые исследователи оспаривают существование взаимосвязи между уменьшением содержания липидов и снижением уровня общей смертности [39-41]. С семидесятых годов для снижения уровня холестерина применяются такие препараты как клофибрат и ниацин. Однако, как показало крупное семилетнее исследование CDP (Coronary Drug Project) [42], ни один из них не уменьшил общую смертность.

Проведены мета-анализы многих крупных клинических испытаний для того, чтобы выявить влияние различных гипохолестеринемических препаратов на специфическую (связанную с повышенным уровнем холестерина) и общую смертность [43-47]. Хотя эти работы и отличаются по виду применявшихся препаратов и по включенным в них исследованиям, их выводы в общем совпадают с выводами D.J. Gordon [47], который осуществил недавно самый обширный мета-анализ. Он рассмотрел результаты 50 рандомизированных контролируемых исследований по оценке эффективности гипохолестеринемических вмешательств, включавших в себя диету, прием фибратов, гормонов, смол и ловастатина. Среднее снижение уровня холестерина в этих исследованиях составило 10%. Достигнут и желаемый благоприятный эффект лечения - снижение смертности от ИБС в среднем на 9%. К сожалению оказалось, что такая комплексная гипохолестеринемическая терапия неожиданно увеличила смертность от причин, не связанных с ИБС, на 24%. В этих 50 исследованиях использование одних только гипохолестеринемических лекарственных препаратов фактически повысило общую смертность на 1%. Таким образом, отрицательное влияние на смертность, не связанную с ИБС, полностью нивелировало ожидаемое положительное действие терапии.

Недавнее клиническое испытание, проведенное в Скандинавии, показало, что у больных со стенокардией или инфарктом миокарда в анамнезе симвастатин снижает уровень холестерина на 25%, а общую смертность - на 30% [48]. Это первое крупное испытание, в котором продемонстрировано положительное влияние гиполипидемического препарата на общую смертность. Независимо от того, связано ли достигнутое с помощью симвастатина снижение смертности с уменьшением содержания холестерина в крови, оценка эффективности лечения только на основании снижения уровня холестерина явно недостаточна. Без учета таких клинических

критериев оценки как общая смертность фибраты и гормоны могли бы широко применяться благодаря их гипохолестеринемическому действию.

### ***Снижение артериального давления***

Эпидемиологические исследования свидетельствуют о том, что артериальная гипертензия является еще одним фактором риска смерти от сердечно-сосудистых заболеваний [49]; снижение артериального давления на каждый миллиметр ртутного столба приводит к снижению смертности на 5% и частоты инсультов на 10%. В одном из ранних крупных исследований по лечению гипертензии (Hypertension Detection and Follow-up Program) [50] выявлено снижение общей смертности при мягкой гипертензии на фоне ступенчатого лечения, начинающегося с диуретиков. Результаты недавнего исследования SHEP (Systolic Hypertension in the Elderly) [51] подтвердили положительное воздействие гипотензивной терапии в виде низких доз диуретиков на выживаемость, частоту несмертельного инфаркта миокарда и инсульта.

Представляют интерес современные подходы к лечению гипертензии [52]. Во многих странах препарат может получить одобрение на основе его косвенной эффективности, т. е. способности снижать артериальное давление. С начала восьмидесятых годов два новых класса гипотензивных препаратов (ингибиторы ангиотензин-превращающего фермента и антагонисты кальция) заняли около 50% всего арсенала гипотензивных средств из-за их положительного влияния на этот косвенный критерий оценки и меньшего количества побочных явлений по сравнению с диуретиками и б-блокаторами. Эти препараты более чем в 30 раз дороже диуретиков. Однако ни в одном из проведенных рандомизированных испытаний не выяснилось, уменьшают ли препараты какого-либо из этих двух классов сердечно-сосудистую смертность или заболеваемость при первичном профилактическом лечении больных с артериальной гипертензией. Недавно проведенное популяционное исследование типа "случай-контроль" позволило высказать предположение о том, что применение антагонистов кальция может повышать риск инфаркта миокарда при артериальной гипертензии [53]. Проводились рандомизированные испытания этих препаратов при застойной сердечной недостаточности и инфаркте миокарда. Мета-анализ влияния антагонистов кальция на выживаемость в этих испытаниях выявил возможное неблагоприятное воздействие данных лекарственных средств [54]. Высказано предположение [52], что использование этих препаратов, клинически полностью не исследованных, для лечения артериальной гипертензии может стоить более 2,5 млрд \$ в год только в США.

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

### **Изучение рака**

#### ***Исследования, посвященные профилактике рака***

Исследования по профилактике рака предпринимаются для того, чтобы выявить такой образ жизни пациента или такие виды вмешательств, которые могли бы способствовать снижению риска возникновения рака у лиц из групп высокого риска. Поскольку таким лицам, в принципе здоровым, профилактическое лечение может проводиться в течение многих лет, представляется чрезвычайно важным определить долговременный риск и положительные стороны подобного лечения.

Важным примером тому является проводимое в настоящее время клиническое испытание финастерида (Proscar, Merck & Co., Inc., West Point, Pennsylvania), назначаемого для химиопрофилактики рака предстательной железы. Если бы в качестве критериев оценки использовались клинические исходы (исчезновение симптомов болезни или снижение смертности), то для получения достоверных выводов о профилактическом действии финастерида необходимо

было бы обследовать более 50 000 мужчин. В данном исследовании косвенным критерием оценки служит наличие раковых клеток в биоптатах предстательной железы после 7-летнего наблюдения или (при наличии клинических показаний) раньше. Использование этого косвенного критерия позволило уменьшить число включенных в исследование в три раза. Наибольшее беспокойство вызывает то, что финастерид понижает концентрацию в плазме крови такого широко используемого маркера рака как простат-специфический антиген. Это обстоятельство может таким образом повлиять на решение о проведении биопсии, что, по предварительным расчетам, 40% участников в течение 7 лет она так и не будет сделана. Финастерид может уменьшить объем предстательной железы и изменить строение ее ткани; в итоге количество ложноположительных результатов в группе вмешательства и в контрольной группе будет разным. Кроме того, следует отметить, что доклиническая стадия рака предстательной железы встречается примерно у 30% мужчин старше 50 лет [15], но только у 9% он разовьется до клинической стадии и менее чем у 3% исход заболевания будет смертельным. Таким образом, финастерид может уменьшить частоту положительных результатов биопсии и не оказывать влияния на смертность или течение рака в клинической стадии. Если данное испытание не будет проведено на большем материале или в течение более длительного времени, уменьшение частоты рака, подтвержденное только результатами биопсии, приведет к широкому использованию финастерида при том, что его единственным ощутимым эффектом будет нарушение либидо и развитие импотенции [15].

### ***Исследования, посвященные лечению рака***

Ответная реакция опухоли на терапию часто использовалась в качестве косвенного критерия оценки в исследованиях по лечению поздних стадий рака, особенно при раке молочной железы, толстого кишечника и при малодифференцированном раке легкого. Категориями этого косвенного критерия являются полная ответная реакция (опухоль не определяется при обследовании), частичная ответная реакция (уменьшение объема опухоли на 50% или более), отсутствие изменений или прогрессирование. К сожалению, ответная реакция опухоли не является надежной заменой такому исходу как выживание [3, 55]. Во многих исследованиях показано, что влияние лечения на этот косвенный критерий не сопровождается изменениями уровня смертности.

Использование косвенных критериев оценки привело недавно к ошибочным результатам при лечении поздних стадий рака толстого кишечника. При назначении 5-фторурацила в сочетании с лейковорином было выявлено статистически достоверное суммарное повышение общей и частичной ответной реакции (на 23%) по сравнению с использованием одного 5-фторурацила (на 11%). Несмотря на такое различие ответной реакции опухоли на лечение, уровень выживания практически не отличался (относительный риск 0,97). Эти результаты получены при мета-анализе почти 1400 случаев [56].

Факторы, обусловившие неудачные результаты применения косвенного критерия оценки (суммы полной и частичной ответной реакции), - меньшее количество полных ответных реакций по сравнению с частичными, небольшое количество долговременных ответных реакций, высокая вероятность отрицательного влияния на выживаемость непредвиденных механизмов действия этих токсичных препаратов.

### **Другие заболевания**

#### ***ВИЧ-инфекция и СПИД***

Наиболее интенсивно использование косвенных критериев оценки обсуждалось при планировании и анализе результатов клинических исследований, посвященных ВИЧ-инфекции и СПИДу. В своем обзоре T.R. Fleming [15] суммировал данные самых крупных исследований по СПИДу, в которых изучалось влияние блокаторов синтеза вирусной ДНК - аналогов нуклеозидов на косвенные критерии оценки и клинические исходы. Результаты посвященной этому вопросу конференции 1993 года [57] свидетельствуют о том, что влияние лечения на наиболее популярный критерий оценки -

количество CD4-лимфоцитов - не позволяет точно прогнозировать влияние лечения на клинические исходы, т. е. на прогрессирование СПИДа или время наступления смерти.

В данном обзоре, включившем в себя 16 крупных исследований по СПИДу, благоприятные изменения косвенного критерия оценки (количества клеток CD4-лимфоцитов) были статистически значимы в 7 из 8 исследований, в которых показано положительное влияние лечения на клинический исход (прогрессирование СПИДа или смерть). К сожалению, они оказались статистически значимыми и в тех 6 из 8 исследований, где не отмечено положительного влияния терапии на течение СПИДа или на смертельный исход. Повышение числа CD4-лимфоцитов было статистически достоверным только в 2 из 4 исследований, в которых лечение оказывало достоверное положительное влияние как на выживаемость, так и на этот критерий. Число CD4-лимфоцитов статистически достоверно повышалось в 6 из 7 исследований, в которых лечение не улучшало выживаемость. В трех дополнительных исследованиях, в том числе в исследовании Concorde [18], продемонстрирована обратная зависимость между уровнем выживания и увеличением количества CD4-лимфоцитов.

В рандомизированное исследование Concorde [18] было включено 1749 ВИЧ-инфицированных бессимптомных пациентов, которым назначалось либо немедленное, либо отсроченное (при появлении клинической симптоматики) лечение зидовудином. За 3-летний период наблюдения снижение количества CD4-лимфоцитов было замедлено в группе немедленной терапии, при этом средняя разница между группами составила 30-35 клеток/мм<sup>3</sup>. Кроме того, у больных в группе отсроченной терапии 50% снижение CD4-лимфоцитов происходило раньше. Однако клинические исходы не отражали этих различий в косвенном критерии оценки. Время прогрессирования ВИЧ-инфекции до развития клинических проявлений СПИДа, клиническая картина СПИДа и время наступления смерти практически не отличались у 175 больных группы немедленной терапии и у 171 больного группы отсроченной терапии. Что касается смертельного исхода, то его частота была значительно ниже при отсроченной терапии (95 и 76 смертельных исходов). Таким образом, выдвинутое ранее предположение о необходимости назначения зидовудина при бессимптомном течении ВИЧ-инфекции не подтвердилось при длительном клиническом наблюдении.

### ***Остеопороз у женщин в менопаузе***

У женщин в менопаузе снижается масса костной ткани и развивается остеопороз, что неизбежно повышает риск переломов бедренных и других костей [58, 59]. Для увеличения массы костной ткани и уменьшения частоты переломов достаточно часто назначаются эстрогены или препараты кальция. Широко используется фторид натрия, стимулирующий образование костной ткани, хотя его применение с этой целью и не одобрено FDA. B.L. Riggs et al. [58] провели плацебо-контролируемое рандомизированное испытание, в которое были включены 202 женщины с остеопорозом и переломами позвонков. Длительность наблюдения равнялась 4 годам. Лечение повышало минеральную плотность костной ткани в поясничном отделе позвоночника на 35% ( $p < 0,001$ ). Однако в группе, получавшей препарат фтора, чаще, чем в группе плацебо, отмечались новые переломы позвонков (163 и 136 переломов соотв.) и переломы иной локализации (72 и 24 перелома соотв.;  $p = 0,01$ ). B.L. Riggs et al. сделали заключение о том, что лечение препаратом фтора повышает минеральную плотность, но кости при этом становятся более ломкими.

### ***Хронический прогрессирующий гранулематоз***

Доверие к косвенным критериям оценки повышает также вероятность ложноотрицательных заключений, что может повлечь за собой отказ от эффективных методов лечения. Это показано в недавнем исследовании, посвященном прогрессирующему гранулематозу у детей [60]. У детей с этим заболеванием нарушена функция иммунной системы: макрофаги поглощают микроорганизмы, но из-за отсутствия выработки перекисных соединений кислорода не убивают их. В результате развиваются тяжелые, рецидивирующие и часто опасные для жизни инфекционные осложнения. Считалось, что интерферон- $\gamma$  обладает терапевтической активностью, способствуя увеличению выработки перекиси водорода и гибели бактерий. В первоначальном плане рандомизированного

плацебо-контролируемого исследования предполагалось, что больные контрольной группы будут получать плацебо в течение такого короткого интервала времени, что возможно будет определять только влияние интерферона- $\gamma$  на косвенные критерии оценки (выработку перекиси водорода и способность макрофагов убивать бактерии). Еще до начала этого исследования план был изменен таким образом, чтобы иметь возможность изучить более длительное воздействие терапии на истинный клинический исход, т. е. на частоту возникновения тяжелых инфекций. Результаты этого исследования [60], проведенного международной группой по изучению хронического прогрессирующего гранулематоза, показали, что интерферон- $\gamma$  статистически достоверно снижал на 70% частоту тяжелых инфекций. Удивительно, что при этом не выявили какого-либо влияния терапии на косвенные критерии оценки. Если бы проводилось более короткое исследование, в котором изучали бы воздействие интерферона- $\gamma$  только на продукцию перекиси водорода и бактерицидную способность макрофагов, то оценка результатов лечения была бы быстрой, но неверной. Это не позволило бы оценить реальную эффективность лечения.

## Заключение

Воздействие на косвенные критерии оценки часто не позволяет предсказать истинный клинический эффект вмешательства. Чаще всего это объясняется тем, что вмешательство имеет непредвиденные механизмы действия, не связанные с патогенезом заболевания. Эти механизмы, часто неожиданные и нераспознанные, могут оказывать на истинный клинический исход влияние, не согласующееся с тем, которого можно было бы ожидать только на основе изучения косвенных критериев оценки.

К сожалению, невозможность прогнозировать истинные клинические исходы при использовании косвенных критериев оценки не является такой уж редкой проблемой. В таблице в качестве примеров представлены различные заболевания, методы их профилактики и лечения, в том числе, данные подробного мета-анализа результатов клинических испытаний. Есть и другие примеры неудачного использования косвенных критериев оценки - испытания вакцин, когда в качестве косвенного критерия используется наличие нейтрализующих антител или клеточный иммунный ответ [61]; испытания тромболитиков, когда критерием оценки является восстановление кровотока по сосуду [62-68]; скрининговые программы по выявлению рака, когда критерием оценки является стадия клинических проявлений процесса [69, 70]; изучение эффективности витаминных добавок при лечении пигментного ретинита, при котором в качестве косвенного критерия используется изменение показателей электроретинограммы [71, 72]; оценка эффективности кислородной терапии при хронических обструктивных заболеваниях легких с использованием физиологических показателей газообмена в качестве критерия оценки [73]; использование результатов зондирования зубов как критерия эффективности лечения [74, 75]; хирургические исследования, где критерием эффективности операции служит восстановление кровотока [76].

Тщательное подтверждение достоверности косвенного критерия оценки производится редко. В окончательных клинических испытаниях по изучению влияния вмешательства на клинические исходы частота ложноположительных и ложноотрицательных результатов должна быть низкой, обычно от 2,5 до 10%. Следовательно, чтобы достоверно замещать собой клинический исход, косвенный критерий оценки должен обеспечивать высокий уровень точности при прогнозировании влияния вмешательства на истинный клинический исход. Прогнозы с точностью 50% (например, на основании числа CD4-лимфоцитов при ВИЧ-инфекции) - то же самое, что прогнозы с помощью подбрасывания монетки. Методы определения достоверности косвенных критериев оценки обсуждаются в работах D.Y. Lin et al. [77], L.S. Freedman et al. [78] и V. DeGruttola et al. [79]. Статистические методы определения достоверности обычно требуют проведения мета-анализа, поскольку необходимые размеры выборки значительно больше, чем при типичном исследовании 3 фазы. Для подтверждения обоснованности использования косвенных критериев требуется также глубокое понимание причинно-следственных путей развития заболевания, ожидаемых и непредвиденных механизмов действия вмешательства. Однако практически эти требования трудновыполнимы.

Косвенные критерии следует использовать там, где они лучше себя зарекомендовали - при оценке новых методов лечения, когда их биологическая активность изучается в предварительных испытаниях 2 фазы. Результаты таких испытаний в свою очередь помогут решить, следует ли проводить более крупные и длительные клинические испытания. Информация о косвенных критериях оценки в окончательных испытаниях 3 фазы важна для понимания механизмов действия вмешательства, однако главной целью исследований должно быть получение прямых доказательств безопасности вмешательства и его влияния на истинные клинические исходы.

---