

Оценка популяционной эффективности медицинских технологий

В. В. Власов¹, О. Ю. Реброва²

¹ Первый Московский государственный медицинский университет (МГМУ) им. И. М. Сеченова Министерства здравоохранения и социального развития Российской Федерации, Москва, Россия

² Российский национальный исследовательский медицинский университет (РНИМУ) им. Н. И. Пирогова Министерства здравоохранения и социального развития Российской Федерации, Москва Россия

Рассмотрены показатели, используемые для оценки вклада факторов риска в заболеваемость в популяции по результатам различных по дизайну исследований. Обсуждаются вопросы оценки эффективности таких популяционных технологий как изменение образа жизни (первичная профилактика) и скрининговые технологии (вторичная профилактика).

КЛЮЧЕВЫЕ СЛОВА: добавочный риск, относительный риск, дополнительное число случаев, дополнительный популяционный риск, скрининг, изменение образа жизни, систематические ошибки.

ВВЕДЕНИЕ

Медицинские технологии (МТ) – это методы диагностики, лечения, профилактики и реабилитации, включая лекарства и изделия медицинского назначения, вакцины, медицинские процедуры (в том числе хирургические), а также любые другие мероприятия по охране и поддержанию здоровья человека. МТ могут применяться как на уровне индивидуума (в клинике, индивидуальная профилактика), так и на уровне популяции (воздействие на группы людей, например, в виде фторирования питьевой воды). Применяемые на уровне популяции МТ являются почти исключительно профилактическими.

МТ, применяемые в клинике, довольно тщательно изучаются, поскольку многие национальные лицензирующие органы систем здравоохранения требуют доказательств их эффективности. Популяционные же технологии, прежде всего направленные на профилактику, изучаются гораздо меньше, хотя их стоимость весьма высока.

Профилактику классифицируют на две категории:

- Первичная – мероприятия, направленные на минимизацию влияния модифицируемых факторов риска (ФР) на заболеваемость.
 - Примеры: прием фолиевой кислоты фертильными женщинами, меры по сокращению доступности табака.
- Вторичная – скрининг с целью выявления у кажущихся здоровыми людей тех или иных заболеваний, контроль над которыми может быть более успешным в случае их обнаружения на ранней (доклинической) стадии.

- Примеры: скрининг на цервикальный рак, рак прямой кишки.

Первичная профилактика может быть организована на разных уровнях:

- на национальном уровне (например, йодирование соли; это не обязательно государственная программа);
- в системе здравоохранения (например, вакцинация);
- на отдельных предприятиях и производствах (например, применение защитных масок);
- применительно к отдельному человеку (например, советы врача).

Несмотря на то, что мероприятия первичной профилактики, организуемые вне системы здравоохранения, не принято называть МТ, в любом случае речь идет о влиянии на показатели здоровья – заболеваемость, фертильность и др. – и, следовательно, в широком смысле их тоже можно относить к МТ. Вторичная профилактика почти всегда организуется в рамках системы здравоохранения. Иногда выделяют другие виды профилактики – третичную, примордиальную и т.д., но эти варианты выделяются на основе интерпретации момента воздействия и не влияют на способы оценки эффективности технологии.

В настоящей статье мы рассмотрим способы оценки эффективности МТ, применяемые на уровне популяций, т.е. оценку популяционной эффективности (в отличие от клинической эффективности). Популяционная эффективность может быть оценена в относительно простых исследованиях типа клинического испытания (КИ). В правильно проведенном КИ, т.е. исследовании, обладающем внутренней валидностью, измеряется эффект применения МТ

у лиц, соответствующих критериям включения и исключения в КИ и условиям оказания медицинской помощи. Экстраполяция данных КИ на людей, отличающихся от включенных в испытание, составляет трудноразрешимую пока проблему. Поэтому в оценках популяционной эффективности МТ сохраняется неопределенность, зависящая от того, насколько правильно оценка, полученная в валидном исследовании, использована применительно к более широкому массам людей.

ОЦЕНКА ПОПУЛЯЦИОННОЙ РОЛИ ФАКТОРА РИСКА

Фактором риска (ФР) заболевания называют признак, связанный с вероятностью развития заболевания в будущем. Среди ФР выделяют немодифицируемые и модифицируемые. С последними связана надежда на то, что устранение такого ФР должно приводить к снижению вероятности заболевания. Иногда специально подчеркивается, что истинный ФР должен соответствовать критерию снижения риска после его устранения. К сожалению, большинство из найденных ФР, например, сотни ФР ишемической болезни сердца, никогда не проверялись на предмет результативности их модификации.

Первичная профилактика заболеваний – достаточно распространенная практика, включающая самые разные воздействия как социального, так и медицинского характера. Среди них можно назвать следующие:

- пропаганда здорового образа жизни,
- вакцинация,
- хлорирование воды,
- применение автомобильных ремней безопасности,
- средства защиты на предприятиях (резиновые перчатки, маски и др.),
- медикаментозные воздействия (аспирин для профилактики острых сосудистых заболеваний, фолиевая кислота для профилактики дефектов развития эмбриона и др.).

Первичной профилактике традиционно уделяется больше внимания, чем вторичной, поскольку привлекательная идея первичной профилактики дает надежду на искоренение заболеваний, что выражается в утверждении «предотвратить легче, чем лечить» (англ. эквивалент «ounce of prevention worth more than a round of treatment»). В такой общей форме правильность этого положения никогда не была доказана. Более того, можно утверждать, что в такой общей форме это утверждение неверно. Тем не менее в некоторых странах, в том числе в России, предпочтение мерам профилактики перед мерами лечения даже внесено в законы о здравоохранении.

Влияние медикаментозных первичных профилактических воздействий неплохо изучено, в отношении некоторых лекарств существует даже вполне

достаточный корпус исследований. В то же время многие рекомендации по коррекции образа жизни не имеют надежных научных обоснований. Главным образом это связано с тем, что степень влияния самого ФР на заболеваемость плохо, ненадежно оценена. В такой ситуации вряд ли может быть известно о популяционной эффективности мер, направленных на его коррекцию. Например, в течение десятилетий предпринимаются попытки оценить вред, наносимый повышенным потреблением пищевой соли, и определить, насколько может быть полезным сокращение потребления соли, но уверенных оценок так пока и не получено. Вместе с тем, повсюду в мире предпринимаются попытки потребление соли сокращать.

В соответствии с концепцией управления риском последовательность действия должна быть следующей:

1) оценка риска – сравнение вредного влияния фактора окружающей среды (ФР) на общественное здоровье с вредом от других воздействий или социальных факторов, а также с пользой, которая связана с этим фактором и сравниваемыми факторами;

2) контроль над экспозицией – мероприятия, направленные на снижение экспозиции фактору риска ниже допустимого максимального предела (т.е. собственное проведение первичной профилактики);

3) мониторинг риска – измерение нового уровня риска (частоты неблагоприятных исходов, которые вследствие контроля над ФР должны стать реже, и новых неблагоприятных эффектов, вызванных применением МТ) после принятых мер по контролю над экспозицией.

ОЦЕНКА ВЛИЯНИЯ ФАКТОРА РИСКА НА ЗАБОЛЕВАЕМОСТЬ ПО РЕЗУЛЬТАТАМ КОГОРТНЫХ ИССЛЕДОВАНИЙ

ФР как признак (экспозиция), приводящий к негативным явлениям, в исследованиях оценивается по его связи с нежелательным исходом (болезнью или иным). Например, в когортных исследованиях, у людей, подвергающихся экспозиции, чаще возникает нежелательный исход, чем в интактной части когорты. В исследованиях В исследовании сравнения с контролем (ИСК) выявляют, что у людей с нежелательным исходом чаще отмечается наличие ФР в прошлом.

По результатам когортного исследования оценка влияния экспозиции на заболеваемость (либо другой исход) получается непосредственно из данных исследования. Обычно вычисляются приведенные ниже показатели (рис. 1), терминология применительно к этим показателям может сильно различаться.

1. Добавочный риск (attributable risk) – разность абсолютных рисков (долей заболевших, т.е. вероятностей заболеть) среди лиц, подвергшихся воздействию ФР (имевших признак, далее для краткости мы рассматриваем ФР как внешнее воздействие, но описываемое

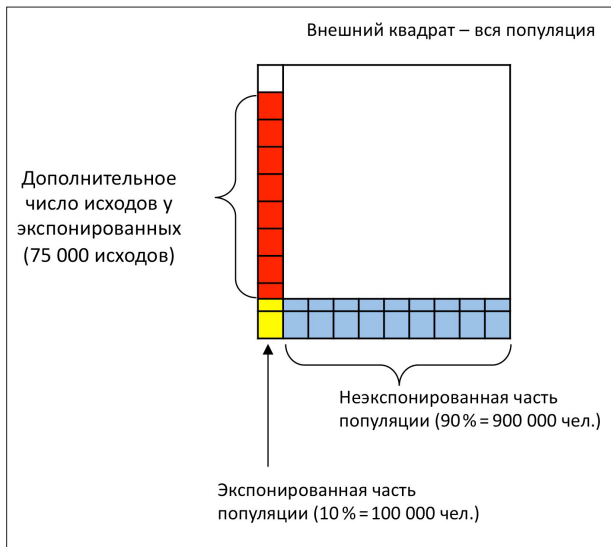


Рис. 1. Соотношение добавочного риска, относительного риска, дополнительного числа случаев и дополнительного популяционного риска.

Примечание. Вероятность неблагоприятного исхода у лиц, принадлежащих к экспонированной части популяции (таких 10 % от всей численности популяции), составляет 90 %, а у неэкспонированных – 15 %. Добавочный риск = $90 - 15 = 75$ %. $OR = 90 / 15 = 6$. На миллион населения экспонировано 100 000 чел., у них возникает всего 90 000 случаев неблагоприятного исхода ($100\,000 \times 0,9$), из них $100\,000 \times 0,75 = 75\,000$ – дополнительное число исходов (на рисунке – красные клеточки). У неэкспонированных 900 000 чел. возникает 135 000 исходов ($900\,000 \times 0,15$) (на рисунке – синие клеточки). Всего на 1 000 000 населения возникает 225 000 исходов (красные + желтые + синие клеточки). Дополнительный популяционный риск (ДПР) = $75\,000 / 225\,000 = 33$ %.

справедливо и для ФР другого типа – особенностей метаболизма человека, особенностей его поведения и т.д.), и лиц, не подвергшихся этому воздействию. Отражает дополнительную вероятность заболевания, и поэтому используется в клинических ситуациях как индивидуальная оценка степени повышения риска. При этом ФР нередко рассматривается как этиологический (причинный) фактор, что является большим упрощением.

Синонимы:

- разница рисков (risk difference);
- избыточный риск (excess risk);
- избыточная частота в популяции (population excess rate).

2. Относительный риск (ОР, relative risk, RR) – отношение абсолютных рисков (долей заболевших, т.е. вероятностей заболеть) среди лиц, подвергшихся воздействию ФР, и лиц, не подвергшихся этому воздействию. Является косвенной мерой влияния ФР на заболевание. Если «добавочный риск» прямо указывает на степень повышения вероятности нежелательного исхода, то ОР указывает только на кратность повыше-

ния. Это означает, что в группе высокого риска лица с ФР могут иметь, например, повышение риска с 15 % до 30 %, а в группе низкого риска – с 3 % до 6 %: в обоих случаях $OR = 2$. Нередко в исследованиях как профилактических, так и лечебных вмешательств указывается привлекательный ОР, но скрывается малый добавочный риск. Такое поведение исследователей надо расценивать как дезинформацию.

3. Дополнительное число случаев нежелательных исходов (attributable number, AN) – число случаев, возникновение которых зависит от экспозиции. Может быть оценено с помощью формулы:

$$AN = Ne \times \text{добавочный риск},$$

где Ne – численность подвергающейся воздействию (экспонированной) части популяции.

4. Дополнительный популяционный риск (ДПР; population attributable risk, PAR) – дополнительная заболеваемость, обусловленная присутствием данного ФР, в долях от популяционной заболеваемости [1].

Синонимы на английском языке:

- population attributable fraction (PAF);
- population attributable risk proportion;
- population attributable risk percent.

В совокупности эти показатели, прежде всего ДПР, позволяют сравнивать относительное значение различных ФР. Опасный, т.е. связанный с высоким риском заболевания, но редкий ФР может быть малосущественным для популяционного здоровья, а относительно слабый, но широко распространенный ФР – весьма важным, поскольку связан с возникновением большого числа новых случаев заболевания. Так, курение – очень распространенный ФР, меры против которого малоэффективны, но ввиду именно высокой распространенности этого ФР даже небольшое сокращение доли курящих в популяции может приводить к большому популяционному эффекту [2].

Велик соблазн сделать вывод, что если ДПР составляет 33 %, то при устранении ФР можно предотвратить 33 % заболеваний. Это весьма легкомысленное предположение: величину положительного эффекта от устранения ФР можно оценить только в профилактическом эксперименте. Обычно в таком эксперименте размер эффекта много ниже, чем предсказывает ДПР (если эффект вообще наблюдается!).

Данные, аналогичные данным когортного исследования, можно получить при анализе результатов клинических испытаний (КИ). Такой анализ все чаще выполняется по окончании КИ. Поскольку в испытание вовлекаются строго определенные группы людей, анализ таких материалов представляет существенную ценность. Как правило, связь фактора риска и заболевания, выявляемая при анализе данных КИ, оказывается более слабой, чем в когортных исследованиях [3].

ОЦЕНКА ВЛИЯНИЯ ФАКТОРА РИСКА НА ЗАБОЛЕВАЕМОСТЬ ПО РЕЗУЛЬТАТАМ ИССЛЕДОВАНИЙ «СЛУЧАЙ-КОНТРОЛЬ»

Если данные о связи ФР и заболевания получены в ИСК, то описанная выше оценка влияния ФР на заболеваемость в популяции и его роли в заболеваемости неприменима. Прежде всего, данные ИСК не позволяют рассчитать вероятность развития заболевания при наличии или отсутствии ФР, равно как и относительный риск (ОР, relative risk, RR). Единственным корректным показателем связи ФР и исхода является отношение шансов (ОШ). Тем не менее в тех случаях, когда изучаются редкие экспозиции и редкие исходы, ОШ приближается к ОР. Критерия «редкости» не существует, но ориентироваться можно на частоты ниже 1 %. Соответственно, возникает возможность приблизительно оценить ДПП как

$$Pe(RR - 1) / (1 + Pe(RR - 1)) \times 100 \%,$$

где Pe – доля экспонированных субъектов среди контролей (незаболевших).

Обратим внимание на то, что доля экспонированных среди заболевших в формуле не присутствует в явном виде, а представлена через RR.

Подавляющее большинство сообщений о выявлении новых ФР заболеваний – это сообщения о результатах ИСК. Эти исследования наиболее подвержены систематическим смещениям (ошибкам). Делать на их основании вывод о причине болезни или оценивать возможности ее профилактики – совершенно бессмысленное упражнение.

ОЦЕНКА ПОПУЛЯЦИОННОЙ ЭФФЕКТИВНОСТИ ВМЕШАТЕЛЬСТВ

Традиционное определение ФР включает в себя указание не только на связь признака (фактора) с развитием заболевания в будущем, но и на снижение частоты такого исхода при устранении ФР. Иными словами, после того как установлен вклад ФР в заболеваемость и предложены меры по его устранению или ослаблению, должны быть проведены исследования эффективности данной технологии профилактики. В идеале такие исследования должны быть проведены как рандомизированные слепые КИ, в которых основная группа получает профилактическое вмешательство, а контрольная – плацебо-вмешательство или не получает вмешательства (ожидает его). В таком исследовании должно быть показано, что в группе вмешательства по сравнению с контрольной произошло статистически значимое снижение заболеваемости. Это полный эквивалент так называемых испытаний на преимущество (superiority trial).

Если КИ лекарственной профилактики проводятся применительно почти ко всем вмешательствам, обещающим положительный эффект, то исследований немедикаментозных профилактических вмешательств немного. В качестве примера можно привести профилактическое изменение пищевого рациона – снижение потребления

соли. Несмотря на то, что пищевая соль в истории человечества всегда была ценным продуктом, издавна высказывалась и до сих пор распространена идея о ее вреде. Большая часть исследований, поддерживающих эту идею, выполнена как ИСК или экологические исследования, реже – как когортные исследования. Измерение потребления соли затруднено, поэтому в большинстве исследований используются суррогатные оценки, вроде национального суммарного потребления соли на душу населения. Тем не менее применительно к снижению потребления соли с целью профилактики хронических заболеваний выполнены ряд КИ. Они обобщены в когортном систематическом обзоре «Снижение пищевого потребления соли для предотвращения сердечно-сосудистых заболеваний» [4]. При анализе 7 исследований, выполненных с участием 6489 лиц (с повышенным и нормальным артериальным давлением) и сроками наблюдения от 7 до 36 месяцев (а в одном случае срок наблюдения после окончания КИ составил 12,7 лет), выяснено, что ни заболеваемость, ни смертность в группе сниженного/запрещенного потребления соли не отличалась от группы обычного потребления соли (рис. 2). При этом ограничение потребления соли приводило в испытании к увеличению смертности у лиц с сердечной недостаточностью (ОР = 2,59 [1,04; 6,44]). К сожалению, данные КИ, указывающие на отсутствие статистически значимого эффекта от снижения потребления соли, и малый размер эффекта даже в КИ с самыми благоприятными для идеи результатами не убеждают сторонников обессоливания пищи. Они, базируясь на вере и отталкиваясь от несистематическим образом обобщенных данных недоказательных исследований, продолжают продвигать в национальные профилактические программы ограничение потребления соли – технологию профилактики, эффективность которой не доказана.

Помимо того, что действенность программы ограничения потребления соли не доказана, такая программа трудно реализуема. Люди малоспособны на длительное употребление не нравящейся им еды. Такая технология обречена на плохую, неполноценную реализацию. Поэтому даже если бы эта программа была действенной, ее популяционная эффективность оставалась бы сомнительной. Внедрение таких вмешательств не только не приводит к ожидаемой пользе, но еще и исчерпывает резервы комплаенса даже у людей, готовых к затратам сил и времени на профилактику болезней.

Второй пример – ситуация с первичной профилактикой сердечно-сосудистых заболеваний посредством комплексных вмешательств. Дело в том, что когортные исследования, так же как и международные сравнительные исследования второй половины XX века, выявили целый ряд ФР сердечно-сосудистых заболеваний. Каждый из ФР давал низкий относительный риск и низкий дополнительный популяционный риск. Лишь выделение очень небольшой группы людей с крайне высокими

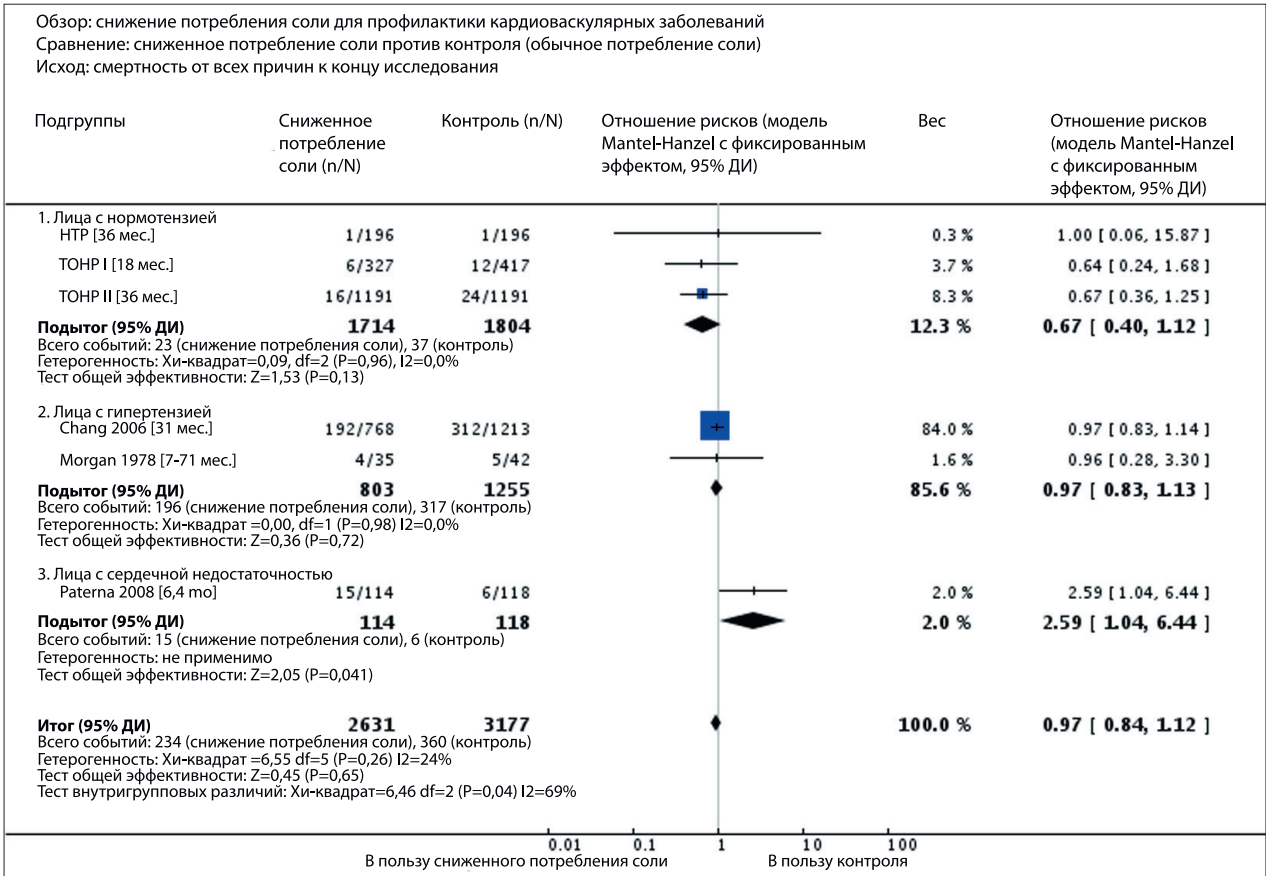


Рис. 2. Отсутствие различий смертности от всех причин при сниженном и обычном потреблении соли (результаты метаанализа; Taylor R.S. et al, [4])

значениями ФР (очень высоким артериальным давлением, очень высокой концентрацией холестерина в плазме) позволяет выделить людей с высоким относительным риском. Однако таким уровнем ФР объясняется очень небольшое относительное число неблагоприятных исходов. Аналогичным образом, выделение людей, подверженных нескольким ФР одновременно, приводит к выделению людей с сильно повышенным риском. Но и в этом случае возникающие у них заболевания составляют лишь небольшую долю от общего числа неблагоприятных исходов (низкий ДПР). Все же модификация одновременно множества ФР казалась очень перспективной, и был проведен ряд КИ для оценки ее возможного эффекта. Эти КИ были обобщены в систематическом обзоре [5]. Обзор ограничивался первичной профилактикой, т.е. теми КИ, в которые включали людей, не имеющих ишемической болезни сердца. В обзоре были рассмотрены результаты 55 таких испытаний продолжительностью от 6 месяцев до 12 лет (медиана продолжительности наблюдения 12 месяцев). Согласно этим результатам, комплексные вмешательства действительно позволили снизить артериальное давление, концентрацию холестерина, частоту и интенсивность курения. Однако, вопреки ожиданиям, комплексные вмешательства почти не влияли на риск развития ишемической болезни сердца и на показатели смертности и заболеваемости.

Возможны различные объяснения неэффективности таких профилактических вмешательств, но факт остается ясным: первичная профилактика, основанная на модификации ФР ишемической болезни сердца, неэффективна. Даже если допустить, что в отдельных случаях эффект существует, он настолько мал, что его не удается обнаружить. При высокой заболеваемости и смертности от ишемической болезни сердца велико желание и пациентов, и медиков, а также правительств сделать что-то для снижения сердечно - сосудистого риска. Поэтому в ряде стран вмешательства такого типа – направленные на модификацию одновременно множества ФР – внедряются в практику. Между тем, это не только приводит к напрасной трате средств здравоохранения, но и дезориентируют людей в их стремлении сохранить здоровье.

ОЦЕНКА ПОПУЛЯЦИОННОЙ ЭФФЕКТИВНОСТИ СКРИНИНГА

Как отмечалось выше, скрининг – это выявление у людей заболеваний и состояний с целью вмешательства, позволяющего предотвратить неблагоприятный исход (возникновение или прогрессирование заболевания, инвалидность, смерть). Таким образом, цель скрининга в более узком смысле – выявление состояний на бессимптомных (малосимптомных) стадиях (т.е. у людей, которые не знают о наличии у них таких

состояний). Обязательное условие эффективности программы скрининга – возможность на этих ранних стадиях более эффективного лечения болезни (или снижения риска), чем на стадиях с клиническими проявлениями, выявляемыми традиционным образом.

Основой скрининга является собственно диагностический метод, который должен быть высокоэффективным, т.е. обладать высокими значениями операционных характеристик (чувствительности и специфичности). Поскольку скрининговый тест применяется заведомо в условиях низкого преваленса, требования к его чувствительности и специфичности заведомо выше, чем в случае клинической диагностики, поскольку иначе прогностическая ценность положительного результата (т.е. вероятность истинно положительного результата среди всех положительных результатов) оказывается крайне низкой. В общем случае скрининговый тест – это не диагностический тест, который применяется в клинической диагностике; к нему предъявляются специальные требования, помимо требования высокой чувствительности и специфичности. Например, при наличии очень дешевого и безопасного теста его можно применять и в случае низкой специфичности, если доступно дообследование людей с положительными результатами.

Вместе с тем, скрининг категорически не сводится к диагностике целевых состояний. Скрининг – это комплексная технология, эффективность и затратная эффективность которой могут быть оценены только в совокупности. Самый эффективный скрининговый тест может оказаться малоэффективным для общества, если не соблюдены другие условия. Перечень таких условий был предложен в 1968 г. Вильсоном и Джаннером [6], а в дальнейшем рекомендован Всемирной организацией здравоохранения.

Несоответствие хотя бы одному из приведенных критериев делает внедрение программы скрининга необоснованным.

Повторим: конечной целью скрининга является не раннее выявление заболеваний, а предотвращение нежелательных исходов – именно это важно для общества. Следовательно, при испытании программы скрининга необходимо показать, что нежелательные для больных и общества исходы будут наблюдаться реже. Например, при скрининге на рак должны улучшаться не только исходы выявленных случаев заболевания, но и достигаться популяционно важный результат – снижение общей смертности.

При испытании программ скрининга (оценке их популяционной эффективности) необходимо проводить масштабные исследования на уровне популяции с обеспечением максимально возможной степени чистоты эксперимента – наличие опытной и контрольной популяций, рандомизации участвующих популяций и т.д. (рис. 3).

Эта схема обычно не может быть реализована применительно к отдельным людям и даже применительно

Критерии Вильсона–Джаннера для оценки программы скрининга

1. Состояние, на которое нацелена программа, должно быть важной проблемой здоровья общества.
2. Развитие болезни (состояния) должно быть хорошо изученным.
3. У этого состояния должна быть выявляемая ранняя стадия.
4. Вмешательство на ранней стадии должно быть более эффективно, чем на поздних стадиях.
5. Должен быть доступен (высокоэффективный – В.В., О.Р.) диагностический тест для выявления ранней стадии болезни.
6. Тест для выявления ранней стадии должен быть приемлемым.
7. Должны быть определены интервалы между повторными обследованиями.
8. Должны быть обеспечены дополнительные диагностические и лечебные вмешательства, потребность в которых возникает в результате скрининга.
9. Физический и психологический вред должны быть меньше, чем польза от программ.
10. Стоимость программы в отношении к ее полезности должна быть не выше, чем у прочих медицинских вмешательств (программ).

но к отдельным медицинским организациям. Если рандомизировать отдельных пациентов, то в рамках одной медицинской организации не удастся реализовать разные способы обследования и ведения пациентов в группах вмешательства и контрольной группе. Медицинские работники станут работать стереотипно, участники опытной и контрольной групп будут обследоваться и лечиться примерно одинаково. В результате возможное различие в исходах между группами останется невыявленным. Поэтому прибегают к рандомизации целых регионов, из которых одни становятся регионами экспериментального вмешательства, а другие – контрольными.

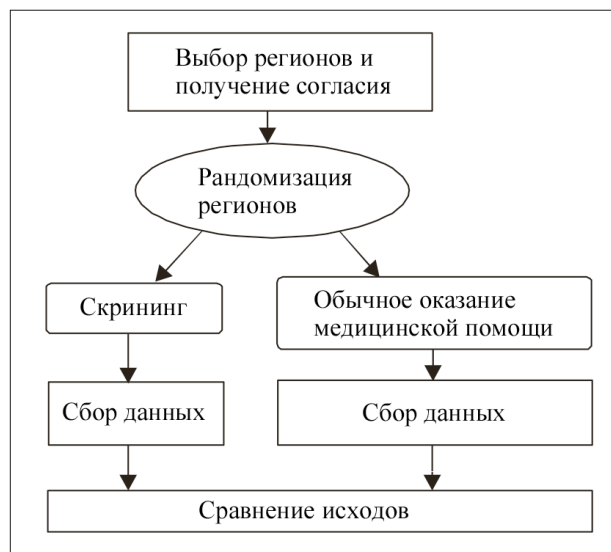


Рис. 3. Схема исследования для оценки эффективности программы скрининга

В испытании должно быть показано, что скрининг действительно приводит к благоприятному изменению важных исходов. Таким образом, основными критериями популяционной эффективности программ скрининга являются снижение специфической смертности и общей смертности. Эти показатели изучаются в сопоставляемых популяциях и сравниваются с использованием тех же количественных мер, что и в случае клинических вмешательств:

- ОР и его 95 % доверительный интервал – отношение абсолютных рисков исхода в сравниваемых популяциях (применяется в проспективных исследованиях)
- ОШ исхода в сравниваемых популяциях и его 95 % доверительный интервал (применяется в ретроспективных исследованиях).

При оценке эффективности программ скрининга нужно помнить о возможных систематических ошибках:

- **Систематическая ошибка, связанная с опережением диагностики** (lead time). Опережением диагностики называют обнаружение заболевания при скрининге, предшествующее диагностике при клиническом проявлении заболевания. Систематическая ошибка может возникать, если раннее вмешательство неэффективно, но увеличение времени жизни с установленным диагнозом воспринимается как увеличение времени дожития. Способ избежать этой ошибки – анализировать возрастные показатели смертности, но не летальности и не времени дожития от момента установления диагноза.
- **Систематическая ошибка, связанная с быстрым прогрессирующим заболеванием** (length time bias). Скрининг, как и любое периодическое обследование, эффективнее выявляет медленно прогрессирующие случаи, более благоприятные с точки зрения прогноза заболеваний. Быстро прогрессирующие случаи клинически проявляются (а значит, становятся объектом клинической диагностики) в промежутках между обследованиями и в силу своего быстрого течения отличаются большей летальностью. Соответственно, летальность среди случаев, выявляемых при скрининге, ниже, чем летальность в обычной клинической практике, однако это не связано с эффективностью скрининга, а лишь с особенностями выявляемых случаев. Данная ошибка устраняется эффективной рандомизацией при сравнении схем «обычная медицинская помощь + скрининг» и «обычная медицинской помощи».
- **Систематическая ошибка в связи с низким комплаенсом** (compliance bias). Люди, добровольно включающиеся в профилактические программы, обычно точнее соблюдают рекомендации врача, проявляя большую готовность к сотрудничеству. Это может обуславливать лучший прогноз в сравнении с прогнозом для тех, кто не стремится участвовать в программе скрининга. Данная ошибка также нивелируется эффективной рандомизацией.

Подробнее о систематических ошибках можно прочитать в [6].

Для скрининга обычно избираются дешевые тесты, но вследствие больших объемов обследования и большого числа ложноположительных результатов, создающих необходимость в дообследовании, а также резкого увеличения числа лечебных или профилактических вмешательств, стоимость программ скрининга очень велика. Она оказывается особенно большой в расчете на один выявленный случай болезни и на один предотвращенный неблагоприятный исход. Например, если тест дешев (стоит 100 руб.), то при выявлении одного “подозрительного” результата на 1000 обследованных только первичное тестирование будет стоить 100 тыс. руб. на один случай. При цене дообследования “подозрительного случая” 10 тыс. руб. и выявлении болезни у 1 из 50 “подозрительных” цена дообследования оказывается равной 500 тыс. руб. на один случай болезни. В итоге затраты составят $50 \times 100\,000 + 500\,000 = 5,5$ млн руб. на один выявленный случай болезни. Даже при использовании таких символически низких расценок расходы выглядят катастрофически высокими. Между тем, они даже не включают стоимости лечения и прямых и косвенных затрат пациентов. Таким образом, без проведения клинико-экономического анализа реализация программ скрининга не может считаться обоснованной.

При оценке программы скрининга важно установить, насколько данная программа служит всему обществу, а не отдельным группам населения, имеющим наилучший доступ к медицинской помощи. Поскольку абсолютное большинство лиц, включившихся в программу скрининга, не больны, но испытывают на себе все побочные эффекты участия в программе, очень трудно сохранить в программе всех, кто в нее включился, на протяжении всего периода исследований. Если обследование сопряжено с неприятными процедурами (как, например, маммография или сигмоскопия), то в течение нескольких лет число участников может существенно сократиться. Если скрининг сопровождается большим числом ложноположительных результатов, то участники подвергаются травматичным обследованиям (например, биопсии) и испытывают стресс при дообследовании. Побочные эффекты могут быть ответственны за то, что не будет обнаружено ожидаемое снижение общей смертности – оно окажется скрыто из-за повышения смертности от побочных эффектов обследования и лечения (болезнь оказывается не хуже вмешательства).

Такие всесторонние исследования эффектов скрининга проводятся редко, поскольку дороги и должны включать очень большое число участников. Даже если такие исследования проведены, ситуация остается непростой из-за проблем, возникающих при экстраполяции данных, полученных в КИ, на популяцию. Примером могут служить результаты маммографи-

Примеры эффективных и неэффективных скринингов для профилактики хронических неинфекционных заболеваний

Наименование профилактического мероприятия	Пояснение	Источник
Выявление низкой физической активности и соответствующая рекомендация по оптимизации физической активности	Ежедневная физическая активность, вовлекающая все группы мышц. Минимальная двигательная активность взрослого – 30 минут ходьбы в среднем темпе в день	Глобальная стратегия по питанию, физической активности и здоровью. WHA57.17. ВОЗ, 2004
Выявление стеноза сонной артерии	Не проводить у лиц без симптомов недостаточности мозгового кровообращения	U.S. Preventive Services Task Force (USPSTF) 2007
Скрининг на рак простаты	Лицам старше 75 лет не рекомендуется. До 75 лет соотношение пользы и вреда неясно	USPSTF 2011
Скрининг на рак яичников	Не рекомендуется, в том числе по генетическим тестам (на мутацию BRCA). Рекомендуется только при наличии этих мутаций в семейном анамнезе.	USPSTF 2005
Фенилкетонурия	Тест следует проводить всем новорожденным	USPSTF 2008
Выявление рака легких	Недостаточно оснований для рекомендации скрининга любым из имеющихся методов (неясен баланс пользы и вреда)	USPSTF 2004
Определение уровня общего холестерина крови	1 раз в 1–2 года; для мужчин – с 35 лет, для женщин – с 45 лет	Курс на оздоровление. Европейская стратегия профилактики и борьбы с неинфекционными заболеваниями. ВОЗ, 2006
Скрининг на рак молочной железы у женщин	При отягощенном семейном анамнезе – с 35 лет, иначе с 50 и до 74 лет специальным рентгеновским маммографом 1 раз в 2 года. Женщины должны быть информированы о потенциальных пользе и вреде скрининга. Обучение самообследованию и профилактический врачебный осмотр не рекомендуются	Борьба против рака. EB114/3. ВОЗ, 2004; Профилактика рака и борьба с ним. WHA 58.22. ВОЗ, 2005; USPSTF 2009
Скрининг на рак шейки матки (ПАП–тест)	Всем сексуально активным женщинам, при отрицательных результатах – до 65 лет.	Борьба против рака. EB114/3. ВОЗ, 2004; Профилактика рака и борьба с ним. WHA 58.22. ВОЗ, 2005; USPSTF 2009
Скрининг на рак прямой кишки	Лицам 50–75 лет 1 раз в 3–5 лет сигмоскопия или ежегодное определение скрытой крови в кале.	Борьба против рака. EB114/3. ВОЗ, 2004; Профилактика рака и борьба с ним. WHA 58.22. ВОЗ, 2005; USPSTF 2008
Скрининг на депрессию взрослых и детей	Проводить при наличии доступа к специализированной помощи для уточнения диагноза и психотерапии	USPSTF 2009
Обучение укладыванию спать ребенка в возрасте до 1 года	Детям до 1 года не рекомендуется спать на спине из-за возможной внезапной смерти во сне	Европейская стратегия «Здоровье и развитие детей и подростков». Инструмент действий. – ВОЗ, 2005
Оценка физического развития (антропометрия)	В детском возрасте – для выявления тяжелых нарушений питания	Стратегические направления улучшения здоровья и развития детей и подростков. – ВОЗ. WHO/FCH/CAN/02.21
Выявление сколиоза у детей	Проведение скрининга бессимптомных детей не рекомендуется	USPSTF 2004
Выявление остеопороза	Рекомендуется женщинам старше 65 лет, а при повышенном риске – старше 60 лет	USPSTF 2002
Скрининг на деменцию	Соотношение пользы и вреда от скрининга пожилых людей остается неясным	USPSTF 2003

ческого скрининга на рак молочной железы (РМЖ). Несмотря на то, что сомнения в эффективности такого скрининга и благоприятном соотношении полезных и вредных эффектов существовали всегда, опубликование в 2000 г. систематического обзора выявленных его эффектов [7] привело к многолетним дискуссиям и изменению оценок применительно к перспективе отдельной женщины, участвующей в скрининге, и к перспективе системы здравоохранения.

В Кокрановском обзоре [8] анализировались 8 клинических исследований скрининга на РМЖ, включавших 600 000 женщин. Три исследования с адекватной

рандомизацией не выявили снижения смертности от РМЖ за 13 лет наблюдения (ОР = 0,90 [0,79; 1,02]), а также общей смертности (от всех причин) за 13 лет наблюдения (ОР = 0,99 [0,95; 1,03]). В то же время число люмпэктоми и мастэктомий было выше в группах скрининга (ОР = 1,31 [1,22; 1,42]). Скрининг ведет к 30 %-ной гипердиагностике и, следовательно, к излишним вмешательствам. Это означает, что на 2000 женщин, проходящих скрининг в течение 10 лет, одной женщине будет продлена жизнь, в то время как дополнительно 10 женщин будут подвергнуты лечебному вмешательству, а 200 женщин будут испытывать

психологический стресс в течение многих месяцев вследствие ложноположительных результатов скрининга. Причиной низкой эффективности программы является недостаточная специфичность скринингового метода – маммографии, дающей значительное число ложноположительных результатов.

К настоящему времени Кокрановским сотрудничеством выполнено 36 систематических обзоров скрининговых технологий (с ключевым словом «screening» в заголовке обзора) и еще 22 находятся в процессе подготовки. Некоторые рекомендации канадских медицинских организаций переведены на русский язык (<http://familymedicine.ru/content/category/4/19/32>), и органам управления здравоохранением можно рекомендовать при принятии решений о финансировании скрининговых программ (в рамках национальной программы профилактики) руководствоваться научно-обоснованной информацией. В табл. приведены примеры эффективных и неэффективных скринингов со ссылками на соответствующие действующие международные и национальные документы.

При рассмотрении вопроса о включении программы скрининга в пакет профилактических мер системы здравоохранения должны учитываться не только надежды на раннее выявление заболеваний, но всеобъемлющий баланс пользы и вреда, возможных при реализации скрининга. Анализ затрат должен включать не только затраты на скрининг и дообследование, но и на лечение дополнительно выявленных случаев заболеваний. Например, в России при финансировании дополнительных диспансеризаций не было выделено дополнительного финансирования на дообследование и лечение дополнительно выявленных больных. В условиях ограниченных возможностей системы здравоохранения по оказанию онкологической помощи дополнительно выявленные «случаи» рака (простаты, яичников и т.д.) вступают в конкуренцию за ограниченные ресурсы, чем ухудшается

положение больных с клиническими формами заболевания и резко снижается удовлетворенность населения качеством медицинской помощи.

ЛИТЕРАТУРА

1. Флетчер Р., Флетчер С., Вагнер Э. Клиническая эпидемиология. Основы доказательной медицины. М., Медиа Сфера. 1998; 134-139.
2. Smith S. C., Benjamin E. J., Bonow R. O., Braun L. T., Creager M. A., Franklin B.A. *et al.* AHA/ACCF Secondary Prevention and Risk Reduction Therapy for Patients With Coronary and Other Atherosclerotic Vascular Disease: 2011 Update: A Guideline From the American Heart Association and American College of Cardiology Foundation. *Circulation*; 2011 Nov 29; 124(22): 2458-73.
3. Tzoulaki I., Siontis K., Ioannidis J. P., Ioannidis J. P. A. Prognostic effect size of cardiovascular biomarkers in datasets from observational studies versus randomised trials: meta-epidemiology study. *Brit. Med.J.* 2011; 343: d6829.
4. Taylor R. S., Ashton K. E., Moxham T., Hooper L., Ebrahim S. Reduced dietary salt for the prevention of cardiovascular disease. *Cochrane.Database.Syst.Rev.* 2011; CD009217.
5. Ebrahim S., Taylor F., Ward K., Beswick A., Burke M., Davey S. G. Multiple risk factor interventions for primary prevention of coronary heart disease. *Cochrane.Database.Syst.Rev.* 2011; CD001561.
6. Wilson J. M. G., Junger G. Principles and practice of screening for disease. Geneva: World Health Organization, 1968.
7. Gotzsche P. C., Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet.* 2000; 355: 129-134.
8. Gotzsche P. C., Nielsen M. Screening for breast cancer with mammography. *Cochrane.Database.Syst.Rev.* 2011; CD001877.

Сведения об авторах:

Власов Василий Викторович

Первый Московский государственный медицинский университет им. И.М. Сеченова, Москва, Россия, профессор, д-р мед. наук

Реброва Ольга Юрьевна

профессор кафедры медицинской кибернетики и информатики ГБОУ ВПО РНИМУ им. Н.И. Пирогова, Москва, Россия, д-р мед. наук

Адрес для переписки:

Москва 117997, ул. Островитянова, д. 1
 Телефон: +7(495)545-0927
 E-mail: o.yu.rebrova@gmail.com

RESEARCH. ANALYSIS. EXPERTISE

Methodology

Criteria for Population Efficacy of Medical Technologies

V.V. Vlassov¹, O.Yu. Rebrova²

¹ I. M. Sechenov First Moscow State Medical University (MSMU), 119991, Moscow, Trubetskaya St. 8, bild. 2, Russia

² The Russian National Research Medical University named after N. I. Pirogov, 117997, Moscow, Ostrovityanova St., 1, Russia

The authors describe the indices used to evaluate the impact of risk factors on morbidity based on the data derived from studies of different designs. The problems of the assessment of the effectiveness of population technologies are discussed, including lifestyle changes (primary prevention) and screening (secondary prevention).

KEYWORDS: attributable risk, relative risk, attributable number, population attributable risk, screening, life style, systematic errors.