

Принципы надлежащей практики для моделирования аналитических решений в исследованиях здравоохранения – отчет рабочей группы Международного общества фармакоэкономических исследований (ISPOR) по надлежащей исследовательской практике при использовании моделирования

Рабочая группа под руководством Милтона С. Вайнштейна

Milton C. Weinstein

PhD, Center for Risk Analysis, Harvard School of Public Health, Boston, MA, USA, and Innovus Research, Inc., Medford, MA, USA

Bernie O'Brien

PhD, Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada

John Hornberger

MD, MS, Acumen, LLC, and Stanford University School of Medicine, Stanford, CA, USA

Joseph Jackson

PhD, Pharmaceutical Research Institute, Bristol-Myers Squibb, Princeton, NJ, USA

Magnus Johannesson

PhD, Centre for Health Economics, Stockholm School of Economics, Stockholm, Sweden

Chris McCabe

MSc, Trent Institute for Health Services Research, University of Sheffield, Sheffield, UK

Bryan R. Luce

PhD, MEDTAP International, Bethesda, MD, USA

Цель: математическое моделирование широко используется при проведении экономической оценки в фармакологии и других отраслях здравоохранения. При использовании таких моделей в правительстве и частном секторе пользователи должны иметь возможность оценить качество той или иной модели исходя из принципов научного подхода к надлежащей практике. В данном отчете представлено согласованное мнение рабочей группы, созданной для разработки руководств по проведению исследований с моделированием и сообщению их результатов. **Методы:** рабочая группа была создана по инициативе и с согласия совета директоров *Международного общества фармакоэкономических исследований (ISPOR)*. В рабочую группу вошли опытные разработчики и пользователи математических моделей, работающие в академических условиях и на производстве, представляющие несколько стран Северной Америки и Европы. Встречи рабочей группы проводились трижды, кроме того, постоянно проходила интенсивная переписка и обмен наработанными материалами по электронной почте. На встречах обсуждались рабочие материалы с комментариями основной группы независимых рецензентов и членов ISPOR. **Результаты:** критерии оценки качества разрабатываемых моделей лежат в трех плоскостях: структура модели; данные, используемые в качестве исходных при моделировании; и утверждение модели. Несколько крупных тем охватывают все три области. Модели и результаты их оценки должны содействовать принятию решения, они не являются утверждениями научного характера; поэтому нецелесообразно требовать проспективного утверждения модели перед ее использованием. Тем не менее, исходные предположения в модели, касающиеся причинной структуры и оценки параметров, должны постоянно изучаться, и, соответственно, должна

пересматриваться модель. Структурные предположения и оценки параметров должны излагаться четко и однозначно; а для пользователей возможность определения условной взаимосвязи между исходными и полученными данными должна предоставляться при проведении анализа чувствительности. **Выводы:** Заключение, созданные на основании моделирования, являются важными источниками информации при принятии решений в здравоохранении. В обязанности разработчиков моделей входит проведение исследований на основании моделирования в соответствии с наилучшими применимыми стандартами качества, и сообщение полученных результатов с надлежащим освещением предположений и предостережением о том, что полученные результаты сделаны на основании предположений и исходных данных, на которых строилась данная модель.

Ключевые слова: Математическое моделирование, фармакология, здравоохранение, фармакоэкономика, модели

Введение

Математическое моделирование широко используется при проведении экономической оценки в фармакологии и других отраслях здравоохранения. Целью моделирования является структурирование данных о клинических и экономических исходах и изложение их в форме, которая поможет при принятии информированных решений в клинической практике и при распределении ресурсов. В моделях синтезируются сведения о влиянии на здоровье и экономических затратах, полученные из различных источников, включая результаты клинических исследований, наблюдательных исследований, сведения из баз данных о страховых выплатах, реестров отдельных случаев, статистических данных и исследований предпочтений. Модель является логической математической основой, которая позволяет объединить факты и величины и связать имеющиеся данные с последствиями, которые представляют интерес при принятии решений в здравоохранении. При принятии решений о распределении ресурсов конечным результатом в модели часто является оценка затрат на количество приобретенных лет жизни с поправкой на ее качество (QALY) или другая информация об эффективности расходования средств.

Хотя результаты рандомизированных клинических исследований остаются основой для эффективного проведения анализа, сами по себе они могут быть недостаточно информативными, если конечные точки не переведены в величины, которые могут быть оценены пациентами, медицинскими работниками, страховщиками и общественностью. Например, предположим, что в рандомизированном клиническом исследовании было доказано, что лечение снижает риск развития редкого осложнения на 50%. Далее, предположим, что в другом исследовании было доказано, что различные варианты лечения снижают риск развития различных более распространенных осложнений на 10%. Последнее вмешательство может быть более эффективным и менее затратным, чем первое, однако простого сравнения результатов двух исследований недостаточно. В то же время, при принятии решения может быть полезным проведение моделирования. Результаты сравнения двух тактик лечения будут зависеть от результатов синтеза сведений о частоте развития осложнений в целевой популяции, относительном снижении риска развития этих осложнений при проведении лечения, выживаемости и качестве жизни с осложнением и без него, стоимости вмешательства и медицинской помощи, необходимых для диагностики и лечения данного осложнения.

Ценность модели определяется не только полученными результатами, но и способностью выявлять логические связи между исходными показателями (данными и предположениями) и конечными показателями в форме оцениваемых последствий и затрат. Именно поэтому модель не должна быть «черным ящиком» для конечного пользователя, она должна быть максимально прозрачной, а логика выведения конечного результата должна улавливаться пользователем на интуитивном уровне. Также по этой причине результаты изучения модели не должны быть представлены в виде точечных оценок или безоговорочных утверждений об эффективности или стоимости. Вместо этого, конечные данные, полученные в модели, должны быть представлены как условные, изменяющиеся от исходных показателей и делаемых предположений; также всегда должен проводиться анализ чувствительности для изучения влияния альтернативных данных и предположений на конечный результат.

Целью написания данной статьи является представление согласованной позиции рабочей группы ISPOR по надлежащей исследовательской практике в области исследований с применением моделей. Как и сами модели, представленная позиция отражает мнение рабочей группы на данный момент и может изменяться по мере развития новых технологий в области математического моделирования, совершенствования компьютерных технологий и методов проведения анализа, появления новых сведений в области здравоохранения и окружающей среды, таких как развитие геномной и микробной устойчивости к лекарственным препаратам и др.

Создание рабочей группы

Глава рабочей группы Международного общества фармакоэкономических исследований (ISPOR) по разработке принципов надлежащей исследовательской практики в области исследований с применением моделей Милтон С. Вайнштейн был назначен на должность председателем Комитета медицинских наук ISPOR Брайаном Р. Льюсом. Члены рабочей группы были приглашены для работы главой группы с согласия совета директоров ISPOR. Проводился поиск специалистов, имеющих опыт в создании или использовании фармакоэкономических моделей, имеющих опыт успешной научной работы в данной области, работающих в академических условиях, на производстве или в качестве советников при правительстве; поиск специалистов проводился в разных странах. Также была выделена группа контроля в рабочей группе ISPOR, которая состояла из специалистов, которые впоследствии будут комментировать проведенную работу. Первая встреча рабочей группы ISPOR была проведена на ежегодной научной конференции Международного общества фармакоэкономических исследований для стран Северной Америки, которая прошла в мае 2000 года в Арлингтоне, штат Виржиния. В последующие месяцы члены рабочей группы использовали электронную почту для обмена информацией и идеями. Черновой отчет о проведенной работе был составлен главой рабочей группы и передан ее членам для ознакомления и комментариев. Проверенный черновик был передан группе контроля, и после получения соответствующих комментариев был подготовлен окончательный отчет. Краткое содержание окончательного отчета было представлено на пленарном заседании ежегодной научной конференции Международного общества фармакоэкономических исследований для стран Северной Америки, которая прошла в мае 2001 года в Арлингтоне, штат Виржиния. Комментарии, полученные от аудитории, были включены в новый вариант отчета, размещенный на сайте для публичного обсуждения. Новый отчет был представлен на ежегодной научной конференции Международного общества фармакоэкономических исследований для стран Северной Америки, которая прошла в ноябре 2001 года в Каннах, Франция; а пересмотренный вариант отчета затем был размещен на сайте ISPOR для получения дальнейших комментариев. Данный вариант отчета включает поправки и комментарии из всех перечисленных выше источников.

Определяемые модели

Национальный исследовательский совет в своем отчете об использовании микромоделирования в социальной политике предложил следующее определение математических моделей: «... воспроизводимая, объективная последовательность расчетов, используемая для получения решений некоторого количества задач...» [1]. Нами математическая модель в области здравоохранения была определена как методика проведения анализа, которая учитывает развитие явлений во времени и в различных популяциях, базируется на данных, полученных из первичных и/или вторичных источников, и целью которой является определение влияния изменений оцениваемых последствий на здоровье и затраты.

Как часть нашего рабочего определения, мы предположили, что модели оценки экономической эффективности разрабатываются для оказания помощи в принятии решений. Это означает, что их целью является не получение безоговорочных утверждений о последствиях каких-либо вмешательств, а выявление взаимосвязи между исходным предположением и результатом. К таким исходным предположениям относятся предположения о причинной связи между значениями переменных; количественные параметры, такие как заболеваемость и распространенность болезни, эффективность лечения, выживаемость, оценка состояния здоровья по определенным шкалам, коэффициент использования ресурсов, и затраты на одну оцениваемую единицу; и оценочные суждения, такие как определение природы последствий, которые оцениваются при принятии решений. Надлежащее исследование, основанное на моделировании, делает все эти предположения явными и прозрачными, а результаты сообщает в условной форме в зависимости от всех перечисленных факторов.

Оценка модели

Модели должны использоваться только после тщательной проверки достоверности математических вычислений и их соответствия спецификации модели (внутренняя валидность), подтверждения того, что исходные и конечные переменные соответствуют имеющимся данным (калибровка), и подтверждения целесообразности полученных результатов и возможности их объяснения на интуитивном уровне (внешняя валидность). В тех случаях, когда различные модели одного и того же решения приводят к различным выводам, разработчики моделей должны быть способны объяснить причину этих различий (перекрестная валидизация). Описание модели должно быть достаточно детализированным, чтобы модель можно было воспроизвести математически.

Проверка прогностической валидности – способности модели делать точные предсказания будущих событий – является важной, но не стопроцентно необходимой. Так как будущие события несут в себе информацию, которая не доступна на момент разработки и калибровки модели, сама модель не должна оцениваться отрицательно за неспособность предсказывать эти события. В то же время, хорошая модель должна поддаваться повторной калибровке и изменению спецификации для адаптации к новым сведениям при их получении. Критерием определения того, необходима ли проверка прогностической валидности перед использованием модели и в какой степени она нужна, зависит от получаемых преимуществ в отношении улучшения свойств модели при принятии решений и затрат, связанных с отсрочкой получения информации при приобретении дополнительных данных [2].

Оценка качества моделей

В оставшейся части документа описывается согласованное мнение рабочей группы, касающееся свойств хорошей математической модели, используемой для принятия решений в области здравоохранения. Мы с трудом выбирали из нескольких отличных статей, предлагающих критерии для оценки качества моделей [3-6]. Описываемые свойства классифицированы в зависимости от структуры, данных и валидации.

Структура

1. Модель должна быть структурирована так, чтобы исходные и конечные данные имели отношение к принятию решений и проведению экономической оценки. Как оцениваемые затраты, так и последствия для здоровья должны отражать выбранную перспективу принятия решений. Например, если целью создания модели является помощь в принятии решения о распределении ресурсов среди широкого спектра медицинских вмешательств на социальном уровне, конечные величины в модели должны быть широко применимы, также должны быть включены важные затраты и последствия для здоровья всех членов затрагиваемой популяции. Если используется перспектива более узкая, чем социальная, в отчете должна обсуждаться, хотя бы на качественном уровне, возможность расширения перспективы до социальной.

2. Структура модели должна соответствовать последовательной теории моделируемого состояния здоровья и имеющимся сведениям, касающимся причинной связи между переменными. Это не означает, что все причинные связи должны быть подтверждены, так как это обычно подразумевается при проверке гипотез, когда проверяемый эффект достигает статистической достоверности на общепринятом уровне достоверности (например, $P < 0,05$). Напротив, это означает, что предполагаемые связи не противоречат имеющимся сведениям и соответствуют общепринятым теориям.

3. Если сведения, касающиеся структурных предположений, являются неполными, и не существует общепринятой теории изучаемого заболевания, необходимо указать на существующие ограничения в выбранной структуре модели. По возможности должен быть проведен анализ чувствительности с использованием альтернативных структур модели – например, с использованием альтернативных суррогатных маркеров или промежуточных переменных.

Пункты 4-8 относятся к моделям переходов состояний (или камерным моделям, или моделям Маркова):

4. Состояния здоровья могут быть определены для соответствия основному патологическому процессу, который может быть незамеченным или не поддаваться непосредственному наблюдению, для соответствия наблюдаемому состоянию здоровья, или комбинации этих состояний. Например, скрининговые модели могут определять состояния здоровья на основании основной патологии, клинического состояния или обоих этих показателей. В то же время, следует быть осторожными, чтобы избежать возможных структурных противоречий, когда потребуется изменить как основное заболевание, так и клиническое состояние, как, например, в моделях скрининга рака, когда выявление рака может иметь различный прогноз в зависимости от используемого метода или частоты проведения скрининга. В целом, структурных противоречий можно избежать при моделировании основного заболевания и затем путем калибровки окончательных данных в соответствии с данными о наблюдаемом клиническом состоянии.

5. Если интенсивность переходов или вероятность зависят от событий или состояний, которые могли случиться ранее, эта зависимость, или «память», должна быть отражена в модели. Это может быть сделано путем указания клинического анамнеза или анамнеза лечения в определении состояния здоровья, или путем включения анамнеза в качестве ковариаты в обоснование вероятностей переходов.

6. Состояния не должны опускаться в связи с нехваткой данных. Примером могут быть хронические состояния здоровья, соответствующие редким нежелательным явлениям или последствиям заболевания, которые не наблюдались в клинических исследованиях. В то же время, включение в модель состояний здоровья должно основываться на данных, соответствующих рекомендации 2 (см. выше).

7. Основаниями для включения в модель дополнительных подгрупп состояний здоровья может быть их клиническая важность, влияние на показатель смертности, влияние на качество жизни или предпочтения пациентов, их связь с затратами на ресурсы или любая комбинация вышеназванных причин. Состояния здоровья, которые могут не расцениваться как клинически значимые, могут быть достаточно важными, чтобы включить их в модель отдельно по перечисленным выше причинам. В свою очередь состояния, которые расцениваются как клинически значимые, могут быть включены в модель для усиления ее внешней валидности, даже если они существенно не влияют на результаты оценки модели.

8. Длина цикла в модели должна быть достаточно короткой, чтобы множественные изменения патологического состояния, симптомов, выбора лечения или затрат в рамках одного цикла были маловероятными. Выбор длины цикла должен быть оправданным.

9. Структура модели должна быть максимально простой, и в то же время должна охватывать основные особенности патологического процесса и исследуемого лечения. Нет необходимости моделировать заболевание во всей его сложности и многокомпонентности, если необходимое решение может быть принято на основании оценки модели с комплексной структурой в отношении заболевания или популяционных подгрупп. Если были сделаны упрощения, то это должно быть обусловлено тем, что указанные упрощения вряд ли существенно повлияют на результаты анализа. В некоторых случаях структурный анализ чувствительности, в котором используется менее комплексная модель, может подтвердить, что упрощения существенно не повлияют на результаты моделирования.

10. Возможности и стратегии не должны быть строго ограничены наличием прямых свидетельств, полученных в клинических исследованиях. Также ряд моделируемых возможностей и стратегий не должен быть ограничен общепринятой в настоящее время практикой. Необходимо соблюдение баланса между включением в модель широкого спектра подходящих возможностей и необходимостью поддерживать управляемость модели, объяснимость ее результатов и их доказательность.

11. Хотя структура модели должна отражать основные черты заболевания и методов его лечения, независимо от доступности данных, предполагается, что доступность данных может повлиять на выбор структуры модели. Например, если в клинических исследованиях чаще всего используется определенная ступенчатая система, состояние здоровья может быть хорошо определено по этой системе, даже если другие системы лучше подходят для прогноза исходов или для дифференцировки качества жизни и затрат.

12. Невозможность обеспечить гетерогенность внутри моделируемой популяции может привести к ошибочным результатам. Когда это целесообразно, моделируемые популяции должны быть разделены на группы с различной вероятностью развития изучаемого явления, различным качеством жизни или разными затратами. Это особенно важно, когда частота развития во времени рецидивирующего явления сопоставляется в различных популяциях, различающихся частотой развития этого явления, так как несоблюдение указанного правила может привести к ошибкам при оценке долгосрочных исходов.

13. Временная протяженность модели должна быть достаточной для отражения важных оцениваемых различий между долгосрочными последствиями и затратами на альтернативные варианты и стратегии. Продолжительность жизни пациента является подходящей для многих моделей и практически всегда требуется в моделях, в которых оцениваемые возможности лечения характеризуются различным меняющимся во времени коэффициентом выживаемости. Более короткие временные границы могут использоваться в случаях, когда выживаемость и долгосрочные хронические последствия не различаются между оцениваемыми возможностями лечения или основываются на понимании процесса развития заболевания и эффекта проводимых вмешательств. В любом случае, отсутствие информации о результатах долгосрочного наблюдения не должно служить обоснованием невозможности продлить временные границы модели достаточно далеко для адекватного принятия решения по результатам проводимого анализа.

Данные

Наши рекомендации по выбору исходных данных в модели группируются по трем категориям: идентификация данных, моделирование данных и включение данных в модель.

Идентификация данных:

1. Модель не должна считаться неверной из-за того, что имеющиеся данные не соответствуют идеальным стандартам научных требований. Решение будет принято с моделью или без нее. Отказ от использования модели в связи с недостаточностью данных будет означать, что решение, принятое без наличия данных и модели, лучше, чем решение, принятое без наличия данных, но с использованием модели. При использовании моделирования имеющиеся сведения могут быть проанализированы логически для последующего принятия решения; без использования модели возможность применения логической основы для анализа имеющихся данных будет утрачена.
2. Следует провести систематизированный обзор литературных источников, касающихся ключевых исходных данных в модели. К описанию модели должны прилагаться сведения о том, что такой обзор был проведен, или подтверждение того, что такой обзор провести невозможно по объективным причинам.
3. Диапазон значений (т.е. верхние и нижние границы) должен указываться при проведении оценки базового случая для всех исходных параметров, для которых выполнялся анализ чувствительности. Выбор параметров для проведения анализа чувствительности проводится аналитиком, однако невозможность проведения анализа чувствительности параметра, значение которого является спорным, ставит под сомнение полученные результаты.
4. Спецификация распределения вероятностей для исходных параметров, основанных на неопределенности выборки и/или разбросе параметров между исследованиями, может быть включена в формальный вероятностный анализ чувствительности. В то же время, это не всегда необходимо или экономически обосновано. Для оценки распределения исходных данных методом предпочтения является использование апостериорного распределения, полученного при проведении формального мета-анализа и байесовского анализа, однако на практике иногда удобнее использовать экспертное заключение (см. пункт 7 ниже).
5. Если известные источники данных исключены из оценки параметров, это решение должно быть обосновано.
6. Источники данных и результаты не должны исключаться по отдельности, так как они не достигают общепринятых границ вероятности, определяемых как «статистическая достоверность» (например, при $P > 0,05$). Все сведения, даже если невозможно исключить случайность их появления, могут быть обоснованно включены в модели. Это может быть сделано при условии, что сомнительность используемых сведений явно указана и проверена при проведении анализа чувствительности, а полученные выводы четко описаны как условные и зависящие от используемых исходных сведений.
7. Мнения экспертов могут законно использоваться при оценке параметров в случае, если эти параметры существенно не влияют на конечные результаты, или если эти параметры подвергнуты анализу чувствительности с четким указанием, что полученные результаты зависят от приведенной субъективной оценки. Если была получена экспертная оценка, а результаты чувствительны к процессу получения выводов, то этот процесс должен быть описан детально. Предпочтительнее использование экспертных оценок, полученных при применении формальных методик, таких как дельфийский метод или техника номинальных групп.
8. Необходимо обосновать наличие оправданной возможности получения новых дополнительных данных перед моделированием. «Оправданной» в данном контексте означает, что затраты и отсрочка при получении данных оправданы ожидаемой ценностью новой информации при проведении анализа. Хотя существуют формальные методы оценки ценности информации, достаточно предоставить наводящее соображение насчет того, почему имеющиеся сведения являются оптимальными с точки зрения решения данной задачи. Обычно поставленной задачи можно достигнуть при проведении анализа чувствительности для демонстрации того, что обоснованное изменение данных приведет к получению сходных результатов, или путем обсуждения того, что затраты и задержка при получении дополнительных данных не восполняют утраченные преимущества, связанные с проведением анализа с имеющимися данными.

Моделирование данных:

1. Моделирование данных означает определенные математические шаги, которые предпринимаются для трансформации эмпирических наблюдений в форму, приемлемую для принятия конкретных решений. Примеры включают:
 - Метод сочетания оценки эффективности лечения, полученной в клинических исследованиях, и оценки основных исходов, полученной при проведении эпидемиологических исследований или при анализе данных общественного здравоохранения. Оценка эффективности лечения может основываться на данных, полученных в соответствии с исходно назначенным лечением (intention-to-treat), или данных, полученных в процессе лечения (on-treatment), в зависимости от целей проводимого анализа. Часто используется подход, предполагающий оценку относительного риска

(или отношения рисков) между исследуемыми вариантами лечения, полученного по результатам клинических исследований, и перенос полученных значений на оценку исходной вероятности выживаемости (например, в группе не леченых пациентов, или леченых традиционными методами) или другой конечной точки, используемой в популяционных исследованиях.

- Метод трансформации интервальных вероятностей, полученных из литературных источников или в клинических исследованиях, в текущие показатели и затем в вероятности перехода, соответствующие временному интервалу, используемому в модели.
 - Метод сочетания в модели смертности, связанной с конкретным заболеванием, и смертности по любым причинам. В целом, значения вероятности смерти по любым причинам можно брать из национальных таблиц продолжительности жизни, за исключением случаев появления других утвержденных источников. В целом, нет необходимости делать поправку на то, что смертность по любым причинам включает смертность, специфически связанную с конкретным заболеванием в общей популяции, за исключением случаев, когда данное заболевание является основной причиной смерти в демографической группе, используемой в модели.
 - Метод моделирования выживаемости (например, в виде экспоненциального или гамма распределения, распределения Вейбулла или Гомперца). Выбор функциональной формы для оценки смертности, связанной с заболеванием, должен быть специфичным и обоснованным. В целом, смертность по любым причинам должна моделироваться непараметрически на основании таблиц продолжительности жизни.
 - Моделирование факторов риска или вмешательств как величин, оказывающих аддитивный (суммируемый) или мультипликативный (множительный) эффект на исходную вероятность развития заболевания или смерти от него. Данные, поддерживающие суммируемую или множительную форму анализа, должны быть взяты из исследований, нацеленных на изучение эффектов этих факторов риска или вмешательств в популяции, стратифицированной в зависимости от исходного риска.
 - Метод комбинирования специализированных индексов оценки общего состояния здоровья для создания многомерной функции полезности. Предпочтительно использовать общепринятые инструменты изучения качества жизни со специализированными системами оценки, основанными на методах «принудительного выбора» (метод стандартного риска, временной компромисс).
 - Метод трансформации значений оценки состояния здоровья (таких как значения по балльным шкалам или классификации состояния здоровья) в значения оценки качества жизни.
 - Метод перевода расходов в затраты.
 - Метод поправки на инфляцию или покупательскую способность во времени или в различных странах. Поправка на инфляцию должна производиться с учетом индекса цен на потребительские товары, а именно компонентов, связанных со здравоохранением, или одним или несколькими субкомпонентами, такими как медицинская помощь или оснащение. Выбор между использованием общего индекса цен на потребительские товары и его одним или несколькими компонентами зависит от способа лучшего представления оцениваемых ресурсов – с использованием общей «покупательской корзины» в рамках индекса цен на потребительские товары или покупательской корзины услуг здравоохранения. Ограничением индекса цен на услуги здравоохранения является то, что он отражает не только цены, но и в некоторой степени количество исходных ресурсов, используемых для оказания услуг здравоохранения. Методом выбора для приведения в соответствие показателей в различных странах является использование паритета покупательской способности. В то же время, можно использовать и простую конверсию валют, если исследуемый препарат присутствует на международном фармацевтическом рынке по фиксированной цене.
 - Метод приведения стоимости и эффектов на здоровье к текущим ценам.
2. Допущения, сделанные при моделировании данных, должны быть освещены и подтверждены свидетельствами их общей приемлемости и, предпочтительно, эмпирической валидности. Ключевые этапы разработки модели должны быть документально подтверждены и записаны. Надежность модели может быть улучшена указанием того, как она была сконструирована, например, перед или во время проведения клинических исследований III или IV фазы, и при описании изменений ее структуры и исходных данных в свете появления новых сведений (например, после завершения клинических исследований) и в ответ на обсуждение полученных результатов с экспертами в различных областях.

3. Если альтернативные, но одинаково научно обоснованные, подходы к моделированию данных приводят к получению существенно различающихся результатов, необходимо провести анализ чувствительности для оценки влияния этих альтернативных методов. Например, если модель предсказывает меньшее увеличение ожидаемой продолжительности жизни у пациентов пожилого возраста, и при этом в модели используется мультипликативная спецификация оценки эффекта вмешательства на исходную смертность, следует проверить альтернативную аддитивную модель. Если существуют более убедительные эмпирические сведения в поддержку одной из функциональных форм, эта форма должна использоваться в качестве базового случая, а альтернативные формы должны проверяться при выполнении анализа чувствительности.
4. Методы моделирования данных должны соответствовать общепринятым методам биостатистики и эпидемиологии. Для моделирования мета-анализ является приемлемым и желательным методом, его использование указывает на тщательный подход к определению гетерогенности используемых источников данных. Гетерогенность можно рассматривать путем разделения полученных сведений в зависимости от различных групп первичных исследований, или путем оценки формальных иерархических моделей для комбинирования информации, полученной в гетерогенных исследованиях.

Включение данных в модель:

1. Единицы измерения, временные интервалы и характеристики популяций должны быть единообразными на протяжении всего использования модели.
2. Приемлемо использование вероятностного (метод Монте-Карло, первого порядка) или детерминированного (когортного) имитационного моделирования.
3. В случае использования имитационного моделирования по методу первого порядка или Монте-Карло должны быть предоставлены свидетельства того, что случайная погрешность моделирования (например, стандартное отклонение конечных величин за один цикл анализа) значительно меньше интересующей выраженности эффекта.
4. Все исследования с применением моделирования должны включать интенсивный анализ чувствительности ключевых параметров. Возможно проведение детерминированного (однофакторного или многофакторного) или вероятностного анализа чувствительности.
5. При наличии возможности анализ чувствительности в моделях, использующих имитацию Монте-Карло, должен использовать фиксированное случайное число «зерно» для каждого анализа чувствительности для сведения к минимуму случайной имитационной ошибки.
6. При проведении детерминированного имитационного моделирования анализ чувствительности можно проводить с использованием вероятностной (метод Монте-Карло, первого порядка) методики и специфического распределения вероятности исходных параметров. При указании таких параметров распределения следует уделить особое внимание надлежащему освещению взаимозависимости параметров при их совместном распределении.
7. Если целесообразно и в случае, когда различия выживаемости, скорректированной с учетом качества жизни, между альтернативными методами лечения меньше длины одного периода, следует проводить коррекцию с учетом полупериода для регулирования в модели оценок, зависящих от времени.

Проверка

Наши рекомендации, касающиеся проверки качества моделей, сгруппированы в три категории: внутренняя проверка, межмодельная проверка и внешняя проверка.

Внутренняя проверка:

1. Все модели должны подвергаться тщательной внутренней проверке и исправлению ошибок («дебаггинг»). Также должны быть предоставлены свидетельства того, что такая проверка была выполнена. Процесс включает использование нулевых или экстремальных значений исходных показателей для проверки того, приведет ли это к получению ожидаемых значений конечных показателей. Процесс также может включать проверку программного кода для выявления синтаксических ошибок и тест на тиражирование с использованием эквивалентных исходных значений.

2. При наличии возможности модель должна быть калибрована по данным. Калибровка возможна, когда существуют данные о значениях начальных и конечных величин в моделируемых временных рамках. Данные для калибровки могут быть взяты из национальных статистических отчетов, например, совокупная и специфическая для определенного пола или возраста смертность, уровень госпитализации, количество выполненных процедур или затраты на ресурсы. Калибровочные данные должны быть взяты из независимых источников относительно источников получения информации об исходных величинах в модели. Модель нельзя считать непригодной, если независимый источник калибровочных данных не существует. Однако недостатком модели может являться тот факт, что при наличии независимых данных, пригодных для проверки, в модели не получаются конечные результаты, соответствующие этим данным (или полученное расхождение невозможно объяснить), или исследователь не проверил соответствие калибровочных данных и конечных результатов моделирования.
3. Хотя исходный код в целом остается собственностью разработчика модели, необходимо оставить возможность получения копий модели с адекватным пользовательским интерфейсом для ее экспертной оценки на условиях строгого соблюдения требований безопасности и авторских прав.

Межмодельная проверка:

1. Модели должны разрабатываться независимо друг от друга, что позволит провести межмодельную проверку (конвергентная валидность).
2. Конечные результаты, полученные в модели, существенно отличаются от опубликованных или общедоступных сведений, полученных в других моделях, исследователь должен предпринять попытку объяснить имеющиеся противоречия. Объясняются ли они различиями в структуре моделей или различиями в исходных показателях?
3. Исследователь должен сотрудничать с другими разработчиками моделей для сравнения полученных результатов и объяснения имеющихся различий. (Мы благодарим финансирующие организации, которые поддерживают такой тип сотрудничества, например, программу разработки моделей в онкологии CISNET, которая поддерживается Национальным институтом США по изучению рака).

Внешняя и предиктивная проверки

Модели должны основываться на убедительных доказательствах, доступных на момент их создания. Для таких состояний как ВИЧ и гиперлипидемия, ранние модели показали, что последствия для здоровья зависят от факторов риска (количество CD4 клеток и сывороточный уровень холестерина). Однако данные, полученные впоследствии в некоторых клинических исследованиях, различались от первоначально сделанных оценок, в то время как результаты изучения других моделей подтвердили первоначальные предположения. Анализ результатов клинических исследований привел к созданию моделей второго поколения как для ВИЧ-инфекции, так и для гиперлипидемии, которые в большей степени соответствовали результатам клинических исследований. В случае с ВИЧ-инфекцией это привело к включению в модель оценки эффективного лечения такого показателя как устойчивость вируса к антиретровирусным препаратам, а также уровня ВИЧ РНК как маркера вирулентности заболевания. В случае с гиперлипидемией в модель были включены такие факторы риска как уровень липопротеинов высокой и низкой плотности. Оставшиеся противоречия между непосредственными эмпирическими данными и результатами моделирования объяснить тяжело. Объясняется ли это артефактами в дизайне исследования (например, выбором пациентов, пересечением лечения) или имеющимися биологическими факторами (например, С-реактивный белок и статины, восстановление иммунитета и антиретровирусная терапия) пока неизвестно. Таким образом, модели не только помогают понять суть проблемы на научном уровне, доступном на момент их создания (на момент, пока ограничены данные о долгосрочном наблюдении за результатами лечения), однако они также могут предоставить основу для сравнения и интерпретации информации, полученной в новых исследованиях. Способность модели адаптироваться к новым сведениям и научному пониманию рассматривается как ее сильная, а не слабая сторона.

1. Так как модели разрабатываются для оказания помощи в принятии решений, а их результаты должны рассматриваться как условные и зависящие от исходного предположения, нет необходимости проверять каждую оценку данных или структурное предположение в проспективных исследованиях перед началом использования модели.
2. Решение о получении дополнительных данных для моделирования должно основываться на балансе между ожидаемой ценностью новой информации и стоимостью этой информации.
 - «Ожидаемая ценность информации» относится к концепции теории принятия решений, которая оценивает информацию в зависимости от ее ожидаемого (или среднего) влияния на последующее принятие решений. Например, ожидаемая ценность информации будет равняться нулю в исследовании параметра модели, предыдущий диапазон значений которого не включал порог чувствительности для выбора между вариантами принятия решения. Суждение, касающееся

предыдущей вероятности возможных результатов исследования, является обязательной частью определения «ожидаемой ценности информации».

- «Стоимость информации» включает затраты на ресурсы, необходимые на выполнение эмпирического исследования или опыта, а также затраты, связанные с упущенной пользой в связи с задержкой принятия решения до завершения исследования. Суждение, касающееся предыдущих вероятных эффектов лечения, является неотъемлемой частью оценки стоимости информации.
 - Рекомендации, касающиеся проведения или дизайна исследований, нацеленных на оказание помощи в последующем принятии решений, могут основываться на результатах формального анализа ценности информации или на неформальной интерпретации результатов анализа чувствительности.
3. Модели никогда не должны считаться завершенными или неизменными. Они должны постоянно обновляться, а иногда консервироваться и перемещаться, так как появляются новые сведения, которую могут повлиять на структуру моделей или исходные данные в них. В заключение следует сказать, что модели, которые не соответствуют появляющимся впоследствии сведениям, и которые не пересматривались и не калибровались в соответствии с этими сведениями, должны консервироваться до проведения такой калибровки.

Заключительные комментарии

Так как данные руководства отражают точку зрения рабочей группы на момент их написания, они не должны считаться строгими и окончательными. Данные рекомендации не являются «сводом правил». Различные обстоятельства могут привести к отклонениям от руководств, в зависимости от источников, доступных для исследователя (время, деньги и данные), и от целей моделирования.

На наш взгляд, наиболее важным моментом, на который следует обратить внимание при оценке математических моделей в здравоохранении, является то, что полученные выводы не должны рассматриваться как утверждение об имеющихся фактах или прогноз на будущее. Вместо этого, целью моделирования является синтез сведений и предположений способом, который позволит пользователям понять влияние исходных показателей на конечные последствия и затраты. Конечные результаты в модели всегда зависят от исходных данных в ней; именно поэтому так важно сохранять прозрачность и доступность исходных данных.

Дополнительная литература, посвященная методам моделирования

Целью данного отчета было не предоставление обзора существующих методов моделирования, а выявление тех аспектов, которые рабочая группа рассматривает как надлежащую исследовательскую практику. Мы рекомендуем следующие литературные источники для тех читателей, которые хотят познакомиться с основными методами моделирования. Учебником по основам анализа решений, включая дерево решений и модель Маркова, является книга Hunink et al [7]. Современные методы моделирования при проведении экономической оценки, включая обзор детерминированного и стохастического подходов к моделированию, представлены в работах Kuntz and Weinstein [8]. Обзор методов работы с погрешностями в моделях представлен в работах Briggs [9] и в 11 главе книги Hunink et al [7].

Следующие члены ISPOR предоставили ценные письменные комментарии к черновику данного отчета: Фантипа Саксонг, MS, факультет фармацевтических наук Университета Чулалонгкорн, Бангкок, Тайланд; Мендель Сингер, PhD, Университет Кейс Вестерн Резерв, Кливленд, Огайо, США; и Лесли Вилсон, PhD, MS, Калифорнийский университет, Сан-Франциско, США.

Авторы также благодарят исполнительного директора ISPOR Мэрилин Дикс Смит, PhD, за административную помощь в проведении заседаний рабочей группы.

Список литературы

1. National Research Council. Improving Information for Social Policy Decisions: The Uses of Microsimulation Modeling, Vol. 1, Review and Recommendations. Washington: National Academy Press; 1991.
2. Weinstein MC, Toy EL, Sandberg EA, et al. Modeling for health care and other policy decisions: uses, roles, and validity. *Value Health* 2001;4: 348–61.
3. Sculpher M, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effectiveness models: a suggested framework and example of application. *Pharmacoeconomics* 2000;17:461–77.
4. Hay J, Jackson J, Luce B, et al. Methodological issues in conducting pharmacoeconomic evaluations— modeling studies. *Value Health* 1999;2:78–81.
5. Akehurst R, Anderson P, Brazier J, et al. Decision analytic modeling in the economic evaluation of health technologies. *Pharmacoeconomics* 2000;17: 443–4.
6. Gold MR, Siegel JE, Russell LB, Weinstein MC, editors. Cost-Effectiveness in Health and Medicine. Report of the Panel on Cost-Effectiveness in Health and Medicine. New York: Oxford University Press; 1996.
7. Hunink M, Glasziou P, Siegel J, et al. Decision Making in Health and Medicine: Integrating Evidence and Values. Cambridge: Cambridge University Press; 2001.
8. Kuntz K, Weinstein M. Modelling in economic evaluation. In: Drummond M, McGuire A, editors. *Economic Evaluation in Health Care: Merging Theory with Practice*. Oxford: Oxford University Press; 2001.
9. Briggs A. Handling uncertainty in economic evaluation and presenting the results. In: Drummond M, McGuire A, editors. *Economic Evaluation in Health Care: Merging Theory with Practice*. Oxford: Oxford University Press; 2001.

Principles of Good Practice for Decision Analytic Modeling in Health-Care Evaluation: Report of the ISPOR Task Force on Good Research Practices— Modeling Studies

Milton C. Weinstein

PhD, Center for Risk Analysis, Harvard School of Public Health, Boston, MA, USA, and Innovus Research, Inc., Medford, MA, USA

Bernie O'Brien

PhD, Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada

John Hornberger

MD, MS, Acumen, LLC, and Stanford University School of Medicine, Stanford, CA, USA

Joseph Jackson

PhD, Pharmaceutical Research Institute, Bristol-Myers Squibb, Princeton, NJ, USA

Magnus Johannesson

PhD, Centre for Health Economics, Stockholm School of Economics, Stockholm, Sweden

Chris McCabe

MSc, Trent Institute for Health Services Research, University of Sheffield, Sheffield, UK

Бryan R. Luce

PhD, MEDTAP International, Bethesda, MD, USA

Objectives: Mathematical modeling is used widely in economic evaluations of pharmaceuticals and other healthcare technologies. Users of models in government and the private sector need to be able to evaluate the quality of models according to scientific criteria of good practice. This report describes the consensus of a task force convened to provide modelers with guidelines for conducting and reporting modeling studies. **Methods:** The task force was appointed with the advice and consent of the Board of Directors of ISPOR. Members were experienced developers or users of models, worked in academia and industry, and came from several countries in North America and Europe. The task force met on three occasions, conducted frequent correspondence and exchanges of drafts by electronic mail, and solicited comments on three drafts from a core group of external reviewers and more broadly from the membership of ISPOR. **Results:** Criteria for assessing the quality of models fell into three areas: model structure, data used as inputs to models, and model validation. Several major themes cut across these areas. Models and their results should be represented as aids to decision making, not as statements of scientific fact; therefore, it is inappropriate to demand that models be validated prospectively before use. However, model assumptions regarding causal structure and parameter estimates should be continually assessed against data, and models should be revised accordingly. Structural assumptions and parameter estimates should be reported clearly and explicitly, and opportunities for users to appreciate the conditional relationship between inputs and outputs should be provided through sensitivity analyses. **Conclusions:** Model-based evaluations are a valuable resource for health-care decision makers. It is the responsibility of model developers to conduct modeling studies according to the best practicable standards of quality and to communicate results with adequate disclosure of assumptions and with the caveat that conclusions are conditional upon the assumptions and data on which the model is built.

Key words: mathematic modeling, pharmacology, health care, pharmacoeconomics, models