

Принципы надлежащей практики для применения аналитического моделирования в оценке медицинских технологий: доклад Рабочей группы ISPOR по надлежащей практике исследований — исследования по моделированию

Milton C. Weinstein¹, Bernie O'Brien², John Hornberger³, Joseph Jackson⁴, Magnus Johannesson⁵, Chris McCabe⁶, Bryan R. Luce⁷

¹ — Center for Risk Analysis, Harvard School of Public Health, Boston, MA, USA, and Innovus Research, Inc., Medford, MA, USA

² — Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada

³ — Acumen, LLC, and Stanford University School of Medicine, Stanford, CA, USA

⁴ — Pharmaceutical Research Institute, Bristol-Myers Squibb, Princeton, NJ, USA

⁵ — Centre for Health Economics, Stockholm School of Economics, Stockholm, Sweden

⁶ — Trent Institute for Health Services Research, University of Sheffield, Sheffield, UK

⁷ — MEDTAP International, Bethesda, MD, USA

Переводчики: Павлыш Андрей Владиславович — к.м.н., докторант Первого Санкт-Петербургского государственного медицинского университета им. акад. И.П. Павлова, г. Санкт-Петербург, Российская Федерация

Рецензенты: Колбин Алексей Сергеевич — д.м.н., профессор, Первый Санкт-Петербургский государственный медицинский университет им. акад. И.П. Павлова, г. Санкт-Петербург, Российская Федерация

Резюме. *Цели:* Математическое моделирование широко используется при экономической оценке фармацевтических и других технологий здравоохранения. Пользователи моделей, как в правительстве, так и в частном секторе должны быть в состоянии оценить качество моделей в соответствии с научными критериями надлежащей практики. В настоящем докладе описываются заключения Рабочей группы, созданной для разработки основных принципов проведения и представления исследований с использованием моделирования. *Методы:* Рабочая группа была назначена по совету и с согласия Совета директоров ISPOR. Членами группы выступили опытные разработчики или пользователи моделей из нескольких стран Северной Америки и Европы, работающие в сфере науки и промышленности. Рабочая группа встречалась трижды, обменивалась идеями в процессе многократной переписки по электронной почте, а также консультировалась с основной группой внешних экспертов и с экспертами из числа членов ISPOR. *Результаты:* Критерии оценки качества моделей были распределены в соответствии с тремя составляющими, относящимися к структуре моделей, используемым входным и выходным данным, и валидации моделей. При рассмотрении этих составляющих было затронуто несколько основных тем. В частности, тот факт, что модели и их результаты должны быть представлены в качестве вспомогательной информации, облегчающей процесс принятия решений, но не в качестве утверждений, претендующих на научную истинность.

В связи с этим, нецелесообразно требовать валидации моделей до их использования. Тем не менее, предположения, лежащие в основе моделей и касающиеся структуры причинно-следственных связей и значений параметров, должны постоянно сравниваться с имеющимися данными, и при необходимости в модели должны быть внесены соответствующие коррективы. Предположения о структуре моделей, а также оценки параметров должны быть представлены чётко и ясно. Проведение анализа чувствительности должно обеспечить пользователям моделей возможность оценить условные связи между входными и выходными переменными моделей. *Выводы:* Оценки, основанные на использовании моделей, являются значимым ресурсом для лиц, принимающих решения в области здравоохранения. Разработчики моделей обязаны проводить процесс моделирования в соответствии с лучшими существующими стандартами качества и представлять конечные результаты, должным образом указывая предположения и ограничения моделей и разъясняя то, что полученные выводы обусловлены вышеуказанными предположениями и исходными данными, на которых строились модели.

Ключевые слова: математическое моделирование, оценка технологий здравоохранения, принятие решений, валидация моделей, анализ чувствительности

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

Abstract. *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.

Keywords: mathematical modeling, health technology assessment, decision-making, model validation, sensitivity analysis

Опубликовано: Milton C. Weinstein, Bernie O'Brien, John Hornberger, Joseph Jackson, Magnus Johannesson, Chris McCabe, Bryan R. Luce. Principles of Good Practice for Decision Analytic Modeling in Health-Care Evaluation: Report of the ISPOR Task Force on Good Research Practices — Modeling Studies. // Value Health 2003; 6:9-17.

Введение

Математическое моделирование широко используется в экономических оценках фармацевтических и других технологий здравоохранения. Цель моделирования состоит в структурировании доказательной базы клинических и экономических результатов в той форме, которая может помочь в процессе принятия решений о клинической практике и распределении ресурсов здравоохранения.

Данные о последствиях для здоровья и затратах на лечение собираются из различных источников, в том числе используются данные клинических и эмпирических исследований, базы данных по страховым случаям, реестры историй болезней, статистические данные по здравоохранению и исследования предпочтений. Модель является логической математической основой, которая позволяет объединить качественные и количественные данные и соединить

их с результатами, представляющими интерес для лиц, принимающих решения в области здравоохранения. При принятии решений о распределении ресурсов конечным результатом модели часто является оценка стоимости дополнительного года жизни с поправкой на качество (QALY) или иная мера соотношения затрат и качества.

Несмотря на то, что основными при исследовании действенности остаются данные из рандомизированных клинических исследований (РКИ), взятые отдельно, они могут привести к неправильным выводам, в случае, если конечные критерии оценки не преобразованы в показатели, которые являются значимыми для пациентов, провайдеров, страховых агентов и широкой общественности. Например, предположим, что согласно результатам РКИ, лечение снижает риск редкого осложнения хронического заболевания на 50%. Далее, предположим, что другое исследование показывает, что иная методика лечения

снижает риск другого, более распространённого осложнения, на 10%. Последнее лечение может быть более эффективным и экономически целесообразным, чем первое, но простого сравнения результатов исследований будет недостаточно. Использование модели может помочь при представлении полученных выводов лицам, принимающим решения. Сравнение двух стратегий лечения будет зависеть от синтеза данных о частоте возникновения рассматриваемых осложнений в целевой группе населения, сокращении относительного риска, вызванного предлагаемым лечением, выживаемости и качестве жизни с учётом возникших осложнений и без них, а также расходов на лечение и необходимую медицинскую помощь для диагностики и купирования данных осложнений.

Ценность модели заключается не только в результатах, которые она предоставляет, но и в её способности выявлять логические связи между входными параметрами (т.е. данными и предположениями) и выходами в виде оцененных исходов и затрат. По этой причине модель не должна быть своего рода «чёрным ящиком» для конечного пользователя. Наоборот, она должна быть как можно более прозрачной, для того чтобы логику, лежащую в основе модели и её результатов, можно было бы понять на интуитивном уровне. По этой же причине результаты моделирования никогда не следует представлять в виде точечных оценок или в виде безусловных утверждений о ответственности или затратах. Необходимо указывать на факт того, что полученные результаты действительны только при условии использования допустимых предположений и конкретных входных данных, и они должны включать анализ чувствительности для изучения влияния альтернативных данных и предположений на результаты модели.

Цель этой статьи заключается в представлении общей позиции Рабочей группы ISPOR по надлежащей практике исследований на основе моделирования. Как и сами модели, эта позиция, являясь результатом обсуждений Рабочей группы на настоящий момент, может быть изменена благодаря появлению новых технологий для моделирования и распространению таких принципиально новых технологий здравоохранения и окружающей среды как геномная резистентность или резистентность микробов к лекарственным препаратам.

Предыстория создания Рабочей группы

Председатель Рабочей группы ISPOR по надлежащей практике исследований с использованием моделирования, *Milton C. Weinstein*, был назначен в 2000 году Председателем Совета по науке в здравоохранении в ISPOR, *Bryan R. Luce*. Члены Рабочей группы были приглашены для участия Председателем, по рекомендации и с согласия Совета директоров ISPOR. Мы искали людей, которые бы обладали опытом в

разработке или использовании фармакоэкономических моделей; ведущих научных специалистов в данной области; людей, работающих в академической и промышленной сферах и в качестве советников в государственных учреждениях; и которые представляли бы различные страны и системы здравоохранения. Была также создана референтная группа из членов ISPOR, которые предоставляли комментарии и замечания по работе. Рабочая группа провела своё первое заседание на ежегодном Северо-Американском Научном Конгрессе ISPOR в Арлингтоне, штат Вирджиния, в мае 2000 г. В последующие месяцы для обмена идеями членами Рабочей группы широко использовалась электронная почта. Председателем был подготовлен проект доклада, который был разослан членам Рабочей группы для рассмотрения и внесения возможных изменений и комментариев. Доработанный проект был разослан членам референтной группы, и после получения её замечаний был сформирован новый проект доклада. Резюме данного проекта доклада было представлено на пленарном заседании ежегодного Северо-Американского Научного Конгресса ISPOR в Арлингтоне, штат Вирджиния, в мае 2001 года. Комментарии от слушателей были включены в переработанный проект, который был размещён на веб-сайте ISPOR для общих замечаний. Следующий вариант проекта доклада был представлен на ежегодной европейской научной конференции ISPOR в Каннах, Франция, в ноябре 2001 года, и исправленный проект был опубликован для дальнейших комментариев на веб-сайте ISPOR. Данный доклад отражает все внесённые замечания и комментарии на всех перечисленных этапах.

Определение Модели

Национальный исследовательский совет в своем докладе по вопросам использования микромоделирования в задачах социальной политики предложил следующее определение имитационной модели: «... воспроизводимая, подтверждённая последовательность вычислений, используемая для создания оценок интересующих величин...» [1]. Под моделью оценки технологий здравоохранения мы понимаем аналитическую методологию, которая учитывает события во времени и в различных популяциях, которая основана на данных, полученных из первичных и/или вторичных источников, и цель которой заключается в оценке результатов вмешательства на последствия для здоровья и соответствующие затраты.

В рамках нашего рабочего определения будем считать, что модели экономической эффективности предназначены для помощи в процессе принятия решений. Это означает, что их целью является не делать безоговорочные утверждения о последствиях того или иного медицинского вмешательства, но выявлять связь между предположениями модели и результата-

ми. Эти предположения включают в себя следующие аспекты: структурные предположения о причинно-следственных связях между переменными; количественные параметры, такие как частота выявления новых случаев заболевания и распространенность заболевания; эффективность и действенность стратегии лечения; выживаемость; шкала оценки состояния здоровья; интенсивность использования; затраты; а также оценочные суждения, такие как характер последствий, имеющих значение для лиц, принимающих решения. Качественное исследование, основанное на модели, представляет выводы, обязательно учитывающие вышеперечисленные предположения, уделяя внимание их понятности и прозрачности.

Оценка модели

Модели следует использовать только после тщательного тестирования, чтобы гарантировать, что математические расчёты точны и соответствуют спецификациям модели (внутренняя валидация модели), что входные и выходные переменные моделей согласуются с имеющимися данными (калибровка модели), и что результаты работы моделей имеют смысл и могут быть объяснены на интуитивном уровне (внешняя валидация модели). В случае, если различные модели одного и того же решения приходят к разным выводам, разработчики модели также должны быть способны объяснить источники различий (кросс-валидация модели). Описание модели должно быть достаточно подробным, так что соответствующие математические расчёты могли бы быть воспроизведены.

Тестирование прогностической валидности, т.е. способности модели делать точные предсказания будущих событий, будет дополнительным плюсом, но не является абсолютно необходимым. Поскольку в будущем может появиться информация, которая не доступна на момент создания и калибровки модели, последняя не должна критиковаться за неспособность предсказать будущее. Тем не менее, хорошая модель должна быть восприимчива к повторной калибровке или уточнению для адаптации к новым данным по мере их поступления. Критерием для определения того, следует ли, и если да, то в какой степени, проводить тестирование модели перед её использованием на предмет достоверности предсказаний, зависит от преимуществ с точки зрения улучшения модели для принятия решений, а также от затрат, связанных с задержкой потока информации при получении дополнительных данных [2].

Оценка качества моделей

Остальная часть данного отчёта описывает согласованное мнение членов Рабочей группы в отношении признаков, которые определяют качественную модель принятия решений в сфере здравоохранения.

Мы активно заимствуем информацию из нескольких превосходных работ, которые предлагают критерии для оценки качества моделей [3—6]. Признаки организованы по основным разделам: структура модели, данные и валидация модели.

Структура модели

1. Модель должна быть построена таким образом, чтобы её входные и выходные параметры имели отношение к позиции, с которой проводится экономическая оценка при принятии решения. Как затраты, так и рассматриваемые медицинские исходы, должны отражать выбранную позицию принятия решений. Например, если исследование призвано помочь лицам, принимающим решения, в распределении ресурсов в широком диапазоне медицинских вмешательств на уровне общества, то результаты модели должны быть широко применимы, и в модель должны быть включены все значимые затраты и последствия для всех членов пострадавшей популяции. Если рассматриваемый ракурс уже, чем социальный, то в отчёте следует отразить, по крайней мере, качественно, последствия расширения модели до социальной точки зрения.
2. Структура модели должна соответствовать как теоретическим положениям, отвечающим за описание моделируемого состояния здоровья, так и имеющимся данным о причинно-следственных связях между переменными. Это не означает, что все причинные связи должны быть доказаны, как это обычно понимается при проверке гипотез, когда показывается, что величина эффекта является статистически значимой при принятом уровне значимости (например, $P < 0,05$). Напротив, это значит, что предполагаемые связи не противоречат имеющимся данным и соответствуют общепризнанным теориям.
3. Если данные в отношении предположений о структуре модели являются неполными, а общепризнанная теория, описывающая процесс протекания рассматриваемого заболевания, отсутствует, то в таком случае следует указать ограничения данных, подтверждающих выбранную структуру модели. По возможности, должен быть выполнен анализ чувствительности с использованием альтернативной структуры модели, например, с использованием альтернативных суррогатных точек или промежуточных переменных.

Пункты 4—8 относятся к моделям переходов между состояниями (камерные модели или модели Маркова):

4. Состояния здоровья могут быть определены так, чтобы они отражали соответствующий процесс протекания заболевания, который может быть

- незаметным или ненаблюдаемым, наблюдаемое состояние здоровья, или же сочетание того и другого. Например, в моделях скрининга состояния здоровья могут определяться, основываясь на основной патологии, клиническом статусе, или на том и другом. Тем не менее, во избежание предвзятости и связанных с этим отклонением в структуре модели, следует соблюдать осторожность, когда вмешательства изменяют как основное заболевание, так и клинические проявления, как, например, в моделях скрининга рака, в которых случаи обнаруженного рака могут иметь разные прогнозы в зависимости от способа или частоты скрининга. В целом, погрешности в структуре модели избегаются путём моделирования состояний, описывающих основное заболевание, с последующей калибровкой выходных значений с учётом данных по наблюдаемым клиническим проявлениям заболевания.
5. В случаях, когда скорости переходов между состояниями или вероятности переходов зависят от событий или состояний, имевших место ранее, такого рода зависимость, или «память», должна быть отражена в модели. Это может быть сделано либо путём включения истории болезни и лечения в определение состояний здоровья, либо путём её включения в качестве ковариаты при определении вероятностей перехода между состояниями.
 6. Состояния не должны быть опущены из-за отсутствия данных. Примерами могут служить хронические состояния здоровья, соответствующие редким побочным явлениям, или осложнения болезни, которые не наблюдаются в клинических исследованиях. Тем не менее, включение в модель определенного состояния здоровья должно основываться на доказательствах в соответствии с вышеупомянутой рекомендацией #2.
 7. Причины включения дополнительного дальнейшего деления состояний здоровья могут быть основаны на их клиническом значении, их отношении к смертности, отношении к качеству жизни или предпочтениях пациентов, их отношении к стоимости ресурсов или любой комбинации вышеперечисленных факторов. Состояния болезни, которые не рассматриваются как клинически значимые, возможно также важно включить в модель отдельно именно по этим причинам. С другой стороны, состояния здоровья, которые рассматриваются как имеющие клиническое значение, могут быть включены для усиления внешней валидности, даже если они не оказывают существенного влияния на результаты модели.
 8. Длина цикла в модели должна быть достаточно короткой, так, чтобы были маловероятны различные изменения в патологии, симптомах, принятии решений о лечении или расходах в пределах одного цикла. Выбор длины цикла должен быть обоснован.
 9. Структура модели, описывающая процесс протекания заболевания и медицинские вмешательства, должна быть как можно более простой. Не обязательно моделировать всю сложность болезни, если решение может быть получено при помощи более простой модели, адекватно описывающей состояния болезни или популяционные подгруппы. При этом, если были сделаны упрощения, их использование должно быть оправдано на основании того, что оно не будет оказывать существенного влияния на результаты анализа. Иногда анализ чувствительности структуры модели, который использует менее обобщенную модель, позволяет удостовериться в том, что упрощения не оказывают существенного влияния на результаты.
 10. Варианты и стратегии не должны быть строго ограничены наличием прямых доказательств из клинических исследований. Также, диапазон моделируемых вариантов не должен ограничиваться лишь теми, которые являются в настоящее время принятой клинической практикой. Должен быть баланс между включением широкого круга возможных вариантов и необходимостью сохранять модель управляемой, интерпретируемой и основанной на фактических данных.
 11. В то время как структура модели должна отражать главные черты заболевания и соответствующие медицинские вмешательства независимо от наличия данных, ожидается, что наличие данных может повлиять на выбор в отношении структуры модели. Например, если конкретная система стадирования используется наиболее часто в клинических исследованиях, то состояния здоровья могут также быть определены в соответствии с этой системой, даже если другие системы стадирования работают лучше с точки зрения прогнозирования результатов или с точки зрения дифференциации качества жизни и затрат.
 12. Отсутствие учёта неоднородности в моделируемой популяции может привести к ошибкам в результатах модели. При необходимости, моделируемые популяции пациентов следует разбивать по группам, которые имеют различные вероятности событий, качество жизни и затраты. Это особенно важно, когда частоты повторяющихся событий с течением времени коррелируют в субпопуляциях, которые имеют различные вероятности наступления событий, так как неспособность сделать это может привести к неправильной оценке долгосрочных результатов.
 13. Временной горизонт модели должен быть достаточно длинным, чтобы отразить важные и значимые различия между долгосрочными последствиями и издержками альтернативных вариантов и

стратегий. Горизонт моделирования, равный по продолжительности времени жизни, подходит для многих моделей и почти всегда требуется для моделей, в которых варианты имеют различные изменяющиеся во времени показатели выживаемости. Использование более коротких горизонтов моделирования может быть обосновано, если выживание и долгосрочные хронические последствия не отличаются среди сравниваемых вариантов или основываясь на понимании процесса протекания болезни и эффективности вмешательств. В любом случае, отсутствие долгосрочных данных наблюдения не должно использоваться как оправдание для отказа расширить временной горизонт до необходимого в рамках рассматриваемого анализа по принятию решения.

Данные

Мы рекомендуем группировать входные параметры моделей по трём категориям: идентификация данных, моделирование данных и применение данных.

Идентификация данных

1. Модель не следует критиковать из-за того, что существующие данные недостаточно отвечают идеальным стандартам научной строгости. Решения будут приниматься в любом случае — с моделями или же без них. Отклонение моделей на основании неполных данных будет означать, что решение, принятое при отсутствии как данных, так и модели, лучше, чем решения с моделью, но без данных. При применении модели имеющиеся данные могут быть использованы для логического обоснования решения. При отсутствии модели будет упущена возможность использовать имеющиеся данные в рамках логической структуры исследования.
2. По ключевым входным параметрам моделей должны проводиться систематические обзоры литературы. Модель должна сопровождаться информацией о том, что такие обзоры были проведены, или же необходимо обосновать отсутствие таких обзоров вследствие адекватности и обобщаемости приобретенных данных.
3. Диапазоны изменения (т.е. верхняя и нижняя границы) должны быть указаны для оценок всех входных параметров базового сценария, для которых выполняется анализ чувствительности. Выбор параметров для проведения анализа чувствительности носит оценочный характер и зависит от мнения аналитика, но неисполнение анализа чувствительности для параметра, значение которого может быть оспорено, может поставить под вопрос все сделанные выводы.
4. В вероятностный анализ чувствительности может быть включена информация по описанию вероятностных распределений для входных параметров на основе неопределенности выборки и/или информации об отклонениях значений в разных исследованиях. Это, однако, не всегда является необходимым или экономически-эффективным. С целью оценки распределений входных параметров предпочтительным вариантом методологии является использование апостериорных распределений, полученных из официальных мета-анализов и байесовского анализа, но практические соображения могут привести и к использованию экспертных оценок (см. ниже, пункт #7).
5. Если известные источники данных исключаются из рассмотрения при оценке параметров, такое исключение должно быть обосновано.
6. Источники данных и результаты не должны быть отклонены только на основании того, что они не достигают общепринятых вероятностных порогов, определяющих «статистическую значимость» (например, $P < 0,05$). Все данные, даже недостаточные для исключения случайности как причины, могут быть на законных основаниях включены в модели. Это действительно при том условии, что неопределенность в отношении оценок указана и проверена в анализе чувствительности, и что в выводах четко указана обусловленность результатов используемыми входными значениями параметров.
7. Экспертное мнение является законным методом оценки параметров при условии, что показано, что эти параметры не оказывают существенного влияния на результаты, либо что по этим параметрам проводится анализ чувствительности с четким указанием, что результаты рассматриваются только в условии этих субъективных оценок. Если результаты чувствительны к используемому мнению экспертов, то процесс сбора информации должен быть детально описан. Предпочтительными являются экспертные оценки, полученные более строгими методами, такими как метод Дельфи или метод номинальных групп.
8. Необходимо убедиться, что перед началом моделирования были рассмотрены разумные возможности для получения новых дополнительных данных. «Разумный» в данном контексте означает, что стоимость и задержки, связанные с получением данных, оправданы ожидаемой ценностью новой информации при анализе. Наряду с существующими формальными методами оценки ценности информации допустимо привести эвристические аргументы относительно того, почему текущая доказательная база была оптимальной с точки зрения обоснования текущих решений. Этого часто можно достичь, применяя анализ чувствительности, чтобы показать, что приемлемые диапазоны данных приведут к качественно схожим выводам, или утверждая, что стоимость и задержки в получении данных не стоят упущенной выгоды от использования только имеющихся данных.

Моделирование данных

1. Моделирование данных представляет собой последовательность математических шагов, направленных на то, чтобы преобразовать эмпирические наблюдения в форму, пригодную для моделей принятия решений. Например:
 - a. Метод для объединения оценок эффективности лечения, взятых из клинических исследований, с базовыми оценками результатов основного сценария, взятыми из эпидемиологических или общедоступных медицинских данных. Оценки эффективности могут быть основаны либо на данных по выборкам пациентов, собирающихся проходить лечение, либо на данных по выборкам пациентов проходящих лечение, в зависимости от целей анализа. Часто уместным подходом является получение оценок относительного риска (или отношения шансов) для различных методов лечения, полученных из клинических исследований, и наложение их на базовые вероятностные оценки (например, в отсутствие лечения или при стандартном лечении) выживания или других ожидаемых результатов из популяционных источников.
 - b. Метод для преобразования интервальных оценок вероятностей из литературных источников или из клинических исследований в моментные показатели, а затем в вероятности перехода или вероятности наступления событий, соответствующие временному интервалу, используемому в модели.
 - c. Метод для объединения в модели показателей смертности по причине болезни и общей смертности. В целом, допустимо получать вероятности общей смертности на основе таблиц продолжительности жизни, используемых в государственной статистике, за исключением случаев, когда может быть обоснован иной источник информации. В общем случае, не требуется делать поправку на то, что смертность от всех причин включает в себя смертность по причине конкретной болезни в общей популяции, если только заболевание не представляет собой основную причину смерти в моделируемых демографических группах.
 - d. Метод моделирования выживаемости (например, в виде экспоненциального, гамма распределения, распространения Вейбулла или Гомперца). Выбор функциональной формы для смертности по причине болезни должен быть указан и обоснован. В общем случае, смертность от всех причин следует моделировать непараметрическими методами на основе данных таблицы смертности.
 - e. Моделирование факторов риска или вмешательства как имеющие аддитивный или мультипликативный эффект на исходные вероятности или частоту заболеваемости или смертности. Доказательства, подтверждающие аддитивную либо мультипликативную форму, следует искать в исследованиях, посвященных изучению воздействию фактора риска или вмешательства на популяцию, стратифицированную по исходному риску.
 - f. Метод для объединения полезностей, специфичных для конкретной предметной области, в многомерную функцию полезности. Лучше при этом использовать проверенные шкалы оценки состояния здоровья и качества жизни с заранее определенными системами оценок, основанными на методах вынужденного выбора (метод стандартного риска, метод временного компромисса).
 - g. Метод для преобразования значений состояния здоровья (например, рейтинговых шкал или классификаций состояний здоровья) в весовые показатели качества жизни.
 - h. Метод для преобразования возможных издержек в затраты.
 - i. Метод корректировки учета инфляции или покупательной способности во времени и между странами. Корректировка с учетом инфляции должна быть основана на индексе потребительских цен (ИПЦ), входящих в него разделах, связанных со здравоохранением, или одним или более подразделах, таких как медицинские услуги или оборудование. Выбор между общим ИПЦ и его компонентами или подкомпонентами, связанными со здравоохранением, зависит от того, в каком из случаев цены на ресурсы представлены лучше: в виде общей «рыночной корзины» в ИПЦ или по потребительской корзине рынка здравоохранения. Ограничением в использовании ИПЦ является то, что он отражает не только цены, но и в некоторой степени величины входных ресурсов, используемых для предоставления медицинских услуг. Предпочтительным методом для внесения корректировок в разных странах является использование паритета покупательной способности. Тем не менее, может быть уместно простое конвертирование валюты, при наличии международного рынка поставок по фиксированной цене.
 - j. Метод дисконтирования затрат и последствий для здоровья.
2. Предположения, положенные в основу модели, должны быть описаны и иметь всеобщее признание, а также, желательно, подтверждение их эмпирической валидности. Ключевые шаги, принятые при разработке модели, должны быть тщательно задокументированы и записаны. Степень доверия к модели может быть повышена, если показано, как модель появилась, например,

- до или во время клинических исследований III или IV фазы, и как ее структура и входные данные видоизменялись с появлением новых данных (например, после завершения клинических исследований) в ответ на последующие обсуждения с клиническими, нормативными и стратегическими экспертами.
3. В случае если альтернативные, но не менее оправданные подходы к моделированию данных могут привести к значительно отличающимся результатам, для оценки роли этих альтернатив необходимо выполнять анализ чувствительности. Например, если модель предсказывает незначительный прирост продолжительности жизни в пожилом возрасте, но при этом модель использует мультипликативный вариант эффекта вмешательства на базовый показатель смертности, то следует также проверить и альтернативу в виде аддитивной модели. Если имеются веские эмпирические доказательства в поддержку одной функциональной формы, то данная форма должна быть использована в базовом сценарии, а альтернативная форма (альтернативные формы) должны быть проверены при проведении анализа чувствительности.
 4. Методы моделирования данных должны соответствовать общепринятым методам, используемым в биостатистике и эпидемиологии. Мета-анализ является допустимым и целесообразным подходом для моделирования, при условии, что приняты меры по выявлению гетерогенности среди источников данных. Гетерогенность можно учитывать либо путем разделения оценок, основанных на различном группировании первичных исследований, либо путем оценки формальных иерархических моделей для объединения информации, полученной из гетерогенных исследований.

Включение данных

1. Единицы измерения, временные интервалы и характеристики исследуемой популяции должны быть взаимно согласованы по всей модели.
2. Допустимы либо вероятностное имитационное моделирование (Монте-Карло, первого порядка), либо детерминированное имитационное моделирование (когортное).
3. В случае использования имитационного моделирования Монте-Карло первого порядка должны быть предоставлены доказательства того, что случайная погрешность моделирования (например, стандартное отклонение для выходных значений за один прогон) заметно меньше, по сравнению с величинами представляющих интерес эффектов.
4. Любое моделирование должно включать в себя подробный анализ чувствительности по ключевым параметрам. Допустимым является проведение детерминированного (одностороннего или многостороннего) либо вероятностного анализа чувствительности.
5. Для минимизации случайной погрешности моделирования, при проведении анализов чувствительности в моделях, которые используют имитационное моделирование по методу Монте-Карло, по возможности следует использовать фиксированное начальное распределение случайных чисел в пределах каждого анализа чувствительности.
6. При использовании когортного моделирования анализ чувствительности может быть произведен при помощи вероятностного имитационного моделирования (Монте-Карло, второго порядка), используя заданные вероятностные распределения входных параметров. При определении этих распределений следует проявлять особое внимание, чтобы гарантировать, что взаимозависимость между параметрами надлежащим образом отражена в совместном распределении параметров.
7. Когда это необходимо, и, если различия в выживаемости с поправкой на качество жизни между альтернативными медицинскими технологиями являются менее длины одного цикла, следует использовать полуциклы для корректировки связанных со временем показателей модели.

Валидация модели

Наши рекомендации по валидации моделей сгруппированы в три категории: внутренняя валидация, межмодельная валидация и внешняя валидация.

Внутренняя валидация

1. Модели должны тщательно тестироваться и проверяться на предмет наличия ошибок и последующего их исправления. Необходимо указывать подтверждения того, что такая внутренняя проверка модели была произведена. Внутренняя проверка должна включать в себя проверку правильности работы модели при задании нулевых или крайних значений входных параметров. Также к такой проверке можно отнести проверку программного кода на наличие синтаксических ошибок, а также тестирование репликации с использованием эквивалентных значений ввода.
2. По возможности, модели должны быть откалиброваны по данным. Калибровка возможна, когда существуют данные как по входным, так и по выходным показателям модели на протяжении всего рассматриваемого промежутка времени. Калибровочные данные могут быть взяты из государственной статистики здравоохранения, такие как общая смертность и смертность по возрастно-половым группам, количество госпитализаций, процедур или стоимость ресурсов. Калибровочные данные должны быть взяты из

источников, не зависящих от данных, используемых для оценки входных параметров модели. Модель не следует критиковать, если не существует независимых калибровочных данных. Тем не менее, модель подвергается критике, если независимые данные, пригодные для валидации, существуют и, либо модель не в состоянии произвести результаты, согласующиеся с этими данными (или несоответствия не могут быть объяснены) либо разработчики модели не исследовали соответствие между результатами моделирования и вышеуказанными независимыми данными.

3. В то время как исходный код должен, как правило, оставаться в собственности разработчика модели, обоснованные просьбы о копиях моделей с соответствующим пользовательским интерфейсом должны быть доступны для экспертной оценки в условиях строгой безопасности и защиты прав собственности.

Межмодельная валидация

1. Модели должны быть разработаны независимо друг от друга для возможности сравнения результатов одной модели с результатами другой модели, моделирующей тот же самый процесс (конвергентная валидность).
2. Если получаемые в модели результаты в значительной мере отличаются от опубликованных или общедоступных результатов, полученных на основе других моделей, разработчик должен предпринять серьезные усилия, чтобы объяснить расхождения. Являются ли расхождения следствием различий в структуре моделей или в значениях входных параметров?
3. Разработчики моделей должны сотрудничать с другими разработчиками при сравнении результатов и формулировании причин возможных расхождений. (Мы приветствуем финансирующие организации, которые поддерживают такой тип сотрудничества, например, программу CISNET моделирования рака при поддержке американского Национального института исследований злокачественных новообразований).

Внешняя и прогностическая валидация

Модели должны быть основаны на лучших данных, имеющихся на момент построения модели. В таких областях, как ВИЧ и гиперлипидемия, более ранние модели предполагали, что последствия для здоровья опосредованы факторами риска (количеством клеток CD4, содержанием сывороточного холестерина). Последующие данные некоторых клинических исследований противоречат оценкам исходных моделей, в то время как другие находятся в соответствии с модельными предположениями.

Информация, полученная из клинических исследований, легла в основу второго поколения моделей, описывающих как ВИЧ, так и гиперлипидемию. Оценки, полученные с помощью этих моделей, более точно соответствуют результатам клинических исследований. В случае с моделями для ВИЧ такой результат был достигнут за счет включения резистентности к антиретровирусным препаратам в оценки действенности лечения, а также включения РНК ВИЧ в качестве маркера вирулентности. В случае с гиперлипидемией это было достигнуто путем моделирования липидных фракций ЛПНП и ЛПВП в качестве факторов риска. Оставшиеся расхождения между прямыми эмпирическими данными и результатами моделирования являются необъясненными. Относятся ли они к артефактам дизайна клинических исследований (например, отбор пациентов, перекрестные стратегии лечения), или к основным биологическим факторам (например, С-реактивный белок и статины, иммунологическое восстановление и антиретровирусная терапия) остается на сегодняшний день неизвестным. Таким образом, модели не только используют известные науке данные на момент разработки модели (в то время, когда там еще могут быть ограниченные данные многолетних наблюдений по новым методам лечения), но они также могут служить основой для сравнения и интерпретации информации из новых исследований. Способность моделей к адаптации к новым данным и новому научному пониманию следует рассматривать не как недостаток, а как сильную сторону модельного подхода.

1. Учитывая то, что модели предназначены для помощи в текущем принятии решений, и результаты моделирования являются действительными только в условиях заданных предположений, нет необходимости в проверке каждой оценки и предположения о структуре модели на проспективных исследованиях перед ее использованием.
2. Решение для получения дополнительных данных для добавления в модель должно быть основано на соотношении между ожидаемой ценностью дополнительной информации и стоимостью информации.
 - а. «Ожидаемая ценность информации» относится к одному из основных понятий теории принятия решений, согласно которому информация оценивается с точки зрения ее ожидаемого (или среднего) влияния на последствия решений. Например, ожидаемая ценность информации была бы равна нулю для изучения параметра модели, априорный диапазон значений которого не включает в себя порог для выбора среди вариантов решения. Оценка, связанная с априорными вероятностями возможных результатов исследования является неизбежной частью оценки «ожидаемой ценности информации».

- b. «Стоимость информации» включает в себя стоимость ресурсов для выполнения эмпирического исследования, а также размер ожидаемой упущенной выгоды от затягивания момента принятия решения до завершения исследования. Оценка, связанная с априорными вероятностями эффектов от лечения, является неизбежной частью оценки стоимости информации.
- c. Рекомендации по проведению или дизайну исследований для формирования руководящих принципов в принятии решений могут быть основаны на формальном анализе ценности информации или на неформальной интерпретации выводов анализа чувствительности.
3. Модели никогда не должны считаться полными или неизменными. Они должны систематически обновляться. Иногда от существующих моделей следует отказываться и заменять другими по мере поступления новой информации, которая может быть использована при построении модели или при задании входных значений. Также, следует прекратить пользоваться моделями, которые оказались несовместимыми с обновленными данными и не были скорректированы с учётом последних, до тех пор, пока такая коррекция не будет произведена.

Заключительные комментарии

Хотя данные рекомендации и отражают точку зрения Рабочей группы на момент составления отчёта, они не должны рассматриваться как окончательные и неизменные. Они не являются итоговым «сводом правил». Различные обстоятельства могут привести к отклонениям от описанных рекомендаций, в зависимости от имеющихся в распоряжении разработчика модели ресурсов (времени, денег и данных) и от цели моделирования.

На наш взгляд, самое главное, что следует иметь в виду при анализе модели оценки технологий здравоохранения, это тот факт, что результаты моделирования не должны рассматриваться как утверждения о фактах или как прогнозы будущего. Скорее, цель моделирования заключается в синтезе доказательств и предположений таким образом, чтобы позволить

конечным пользователям получить представление о роли этих входных данных для оценивания последствий и затрат. Результаты моделирования всегда зависят от входных данных, поэтому так важно, чтобы данные на входе модели были настолько прозрачными и доступными, насколько это возможно.

Дальнейшее чтение по методологии моделирования

Цель данного отчёта заключается не в предоставлении некоего обзора методологии моделирования, а скорее в определении тех аспектов методологии, которые Рабочая Группа относит к надлежащей практике проведения исследований. Для читателей, желающих ознакомиться с основами методов моделирования, мы рекомендуем следующие источники. В качестве вводного учебного пособия по теории принятия решений, в том числе деревьям решений и моделям Маркова, смотрите *Hunink* и соавт. [7]. Для ознакомления с современными методами моделирования при проведении экономической экспертизы, включая обзор методов моделирования выживаемости по данным клинических исследований, а также обзор детерминированных и стохастических подходов к моделированию, смотрите Kuntz и Weinstein [8]. Обзор методов обработки неопределенности в моделях представлен в работе *Briggs* [9] и одиннадцатой главе книги *Hunink* и соавт. [7].

Благодарности

Рабочая группа выражает благодарность следующим членам ISPOR за предоставленные ценные письменные замечания по проектам данного отчёта: Phantipa Sakthong, MS, факультет фармацевтических наук, Университет имени Короля Чулалонгкорна, Бангкок, Таиланд; Mendel Singer, PhD, университет Кейс Вестерн Резерв, Кливленд, Огайо, США, и Leslie Wilson, PhD, MS, Калифорнийский университет в Сан-Франциско, Калифорния, США.

Авторы также благодарят Исполнительного Директора ISPOR Marilyn Dix Smith за оказанную организационную поддержку в созыве совещаний Рабочей Группы.

Литература

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.