

**О.Ю. РЕБРОВА,**

д.м.н., Федеральное государственное автономное образовательное учреждение высшего образования «Российский национальный исследовательский медицинский университет имени Н.И. Пирогова» Министерства здравоохранения Российской Федерации, г. Москва, Россия, e-mail: o.yu.rebrova@gmail.com

ЖИЗНЕННЫЙ ЦИКЛ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ ВРАЧЕБНЫХ РЕШЕНИЙ КАК МЕДИЦИНСКИХ ТЕХНОЛОГИЙ

УДК 004.9; 61; 519.688

DOI: 10.37690/1811-0193-2020-1-27-37

Реброва О.Ю. *Жизненный цикл систем поддержки принятия врачебных решений как медицинских технологий* (ФГАОУ ВО РНИМУ им. Н.И. Пирогова, г. Москва, Россия)

Аннотация. Системы поддержки принятия врачебных решений (СППВР) можно классифицировать на информационно-справочные и интеллектуальные, а последние, в свою очередь, на моделирующие и имитирующие рассуждения. Моделирующие системы основаны на формализованных знаниях экспертов, а имитирующие – на моделях, построенных различными методами многомерного анализа данных, в т.ч. методами машинного обучения. СППВР следует рассматривать как медицинские технологии, следовательно после их разработки должны последовать этапы оценки их аналитической (технической) валидности, а затем (независимо от актуальных регуляторных норм, в соответствии с принципами доказательной медицины) – и клиническая валидация, в ходе которой должны быть получены доказательства эффективности таких систем в улучшении исходов пациентов и их безопасности. После получения таких доказательств может быть проведен клинико-экономический анализ с целью обоснования экономической целесообразности использования СППВР.

Ключевые слова: системы поддержки принятия решений, медицина, жизненный цикл, аналитическая валидация, клиническая валидация, клинико-экономический анализ.

UDC 004.9; 61; 519.688

Rebrova O.Yu. *Life cycle of decision support systems as medical technologies* (Pirogov Russian National Research Medical University, Moscow, Russia)

Abstract. Decision support systems (DSS) in medicine can be classified into reference and intellectual, and the latter, in turn, into modeling and imitating human reasoning. Modeling systems are based on formalized expert knowledge, and imitating ones are based on models built by various multidimensional data analysis methods. DSS should be considered as medical technologies, therefore, after their development, assessing of analytical (technical) and clinical validity should follow, regardless of current national regulatory documents. Clinical validation have to be based on principles of evidence based medicine and demonstrate superiority, non-inferiority or equivalence to routine practice. Then a clinical and economic analysis can be carried out in order to justify the economic feasibility of DSS, and later health technology assessment can be performed.

Keywords: decision support system, medicine, life cycle, analytical validation, clinical validation, clinical economic analysis.

Актуальным определением системы поддержки принятия врачебных решений (СППВР) следует считать следующее: «СППВР – это программное обеспечение, позволяющее путем сбора и анализа информации влиять на принятие врачом решения при обследовании пациента, диагностике, назначении лечения с целью снижения ошибок и повышения качества оказываемой медицинской помощи» [1].

Различными исследователями и разработчиками определения СППВР и их типы даются несколько по-разному, однако нам кажется целесообразной следующая их классификация (см. врезку).

Классификация СППВР

1. Информационно-справочные СППВР

Авторы [1] дифференцируют функционал таких систем следующим образом:

- справочно-библиотечные системы;
- электронные медицинские карты;
- интегральный анамнез;
- автоматизация врачебных назначений;
- автоматический контроль требований нормативно-правовых актов;
- автоматический контроль «клинического минимума»;
- автоматическая поддержка клинических руководств;
- автоматический контроль критериев качества медицинской помощи.

2. Интеллектуальные СППВР (системы, имитирующие или моделирующие рассуждения врача, и СППВР)

2.1. Системы, имитирующие рассуждения врача (вычислительные системы):

2.1.1. Модели, построенные с применением методов многомерной математической статистики, нейронных сетей (обычно слабо интерпретируемые, для пользователя-врача выглядят как «черный ящик»).

2.1.2. Модели, построенные с применением методов математической логики (обычно хорошо интерпретируемые). Пример: модели, построенные с использованием деревьев решений, ДСМ-метода, логико-статистического подхода.

2.2. Системы, моделирующие рассуждения врача (экспертные системы; системы, построенные на знаниях). Разрабатываются модели знаний, извлеченных из различных источников (главным образом, текстов и экспертов).

3. Гибридные системы – сочетающие информационно-справочные и интеллектуальные компоненты.

До настоящего времени не получены убедительные доказательства эффективности и безопасности информационно-справочных СППВР, преимущественно разработанных для повышения соответствия врачебной практики клиническим рекомендациям, в улучшении клинических исходов пациентов. Проведены десятки, и возможно даже сотни, рандомизированных контролируемых испытаний (РКИ), однако большинство систематических обзоров (СО) не позволяют сделать вывод об эффективности таких систем [2–6]. Показано некоторое улучшение первичной и вторичной профилактики в результате применения СППВР: по результатам 28 РКИ выявлен сниженный относительный риск новых случаев острых заболеваний (инфаркт миокарда, острое нарушение мозгового кровообращения, пневмония и др.) 0,82, 95% ДИ (0,68; 0,99) [7]. Полезность СППВР на основе встраивания КР для профилактики нескольких состояний показана и в систематическом обзоре [8], однако в него включены неконтролируемые (оценка «до-после») и нерандомизированные исследования, следовательно, полученные доказательства имеют низкий уровень достоверности.

Что касается иСППВР, то таких исследований пока меньше, однако уже выполнены несколько СО. В СО [9] изучены 39 исследований СППВР по ведению острой и хронической боли, основанных на искусственных нейронных сетях (ИНС), статистических методах машинного обучения, деревьях решений, логических методах и др. Хотя получены отдельные отличные результаты (в основном

с использованием ИНС), в целом качество таких систем пока нельзя считать хорошим. СО [10] включает 10 исследований разных дизайнов, оценивавших способы выявления пациентов с риском увеличения интервала QT с применением методов предиктивной аналитики, и какой-либо определенный вывод сделан не был. Тридцать исследований разных дизайнов, включенные в СО [11], обобщены с целью оценки качества диагностики рассеянного склероза с использованием систем на основе нечеткой логики, ИНС, формализации правил и других методов. Наилучшие результаты получены с использованием ИНС – чувствительность составила 97%, точность – 99%. Что касается ИИ в области анализа медицинских изображений, то преимущество алгоритмов, построенных на глубоком машинном обучении, над врачебными оценками отсутствует. Этот вывод сделан по результатам мета-анализа, включившем 82 исследования [12]. Таким образом, сплошных успехов иСППВР тоже пока не наблюдается, а качество имеющихся доказательств низкое.

СППВР для врача фактически играют роль «второго мнения», которое он может учесть или не учесть при выработке своего решения, недаром такие системы также называют рекомендательными. Поскольку в настоящее время многие врачи стараются практиковать в соответствии с концепцией доказательной медицины, СППВР рассматривается ими как любая другая медицинская технология (МТ), потенциально влияющая на состояние больного, и которая, соответственно, должна иметь научное обоснование своей эффективности и безопасности.



Рис. 1. Жизненный цикл медицинских технологий

Конечно, степень влияния СППВР на решение врача требует специального изучения, но в дальнейшем мы будем исходить из предположения ее максимального влияния на решение врача. Это позволит рассматривать СППВР в контексте необходимости максимизации ее эффективности и безопасности.

Рассмотрим жизненный цикл МТ в применении к СППВР, прежде всего интеллектуальным. Типичный жизненный цикл МТ представлен на рис. 1.

I. Разработка и производство

Первая стадия – разработка и производство – с нашей точки зрения, должна включать следующие этапы.

1. Постановка задачи совместно с врачами (в зарубежных документах этому этапу примерно соответствует процесс установки валидных клинических ассоциаций [13]). Необходимо определить, какие именно решения предполагается «поддерживать», и почему это актуально. Инициативное моделирование, которое особенно свойственно «чистым» математикам-неофитам в медицине, получающим в руки «большие данные», без согласования постановки задачи с врачами, как мы полагаем, обречено на неудачу.

Для того чтобы оценить новизну и актуальность решаемой задачи, необходимо выполнить СО – поиск всех имеющихся решений в данной проблемной области. Например, если разрабатывается система диагностики, должны быть изучены все имеющиеся способы диагностики, сопоставлены их операционные характеристики (чувствительность, специфичность, прогностические ценности) с учетом преваленса изучаемых состояний в целевой популяции. Вполне возможно, что уже разработаны и другие ИСППВР, поскольку в настоящее время выполняется все возрастающее число работ по прикладным задачам математического моделирования. Источниками информации должны служить базы данных научных публикаций (как минимум – Кокрейновская библиотека, MEDLINE, Научная электронная библиотека), клинические рекомендации (национальные и зарубежные).

Поддерживаемые врачебные решения можно разделить на следующие типы:

1. Диагностика (включая дифференциальную).
2. Прогнозирование:
 - клинических событий (с целью их первичной и вторичной профилактики);
 - эффективности лечения (в интересах его персонализации).

Задача диагностики в имитирующих иСППВР обычно решается применением методов машинного обучения «с учителем», задача прогнозирования – как методами «с учителем», так и «без учителя» (data mining).

Для этих двух типов задач актуальность определяется обычно разными факторами.

В случае диагностики актуальность разработки иСППВР обычно определяется недостатками имеющегося способа диагностики – он может быть неточный, дорогой, инвазивный, слишком длительный, требующий специальной квалификации врачей и т.д.

В задачах прогнозирования частой является ситуация полного отсутствия способа прогноза.

В соответствии с концепцией ДМ на этапе постановки задачи должны быть определены целевая популяция, способ контроля и оцениваемый исход.

Определение целевой популяции осуществляется путем формулирования критериев включения и исключения случаев. Именно для такой популяции разрабатывается иСППВР, и именно к такой популяции может впоследствии применяться разработка (в ДМ это называется обобщаемостью результатов исследования). Пример: результаты, полученные на добровольцах, пришедших на скрининг, не могут быть перенесены на общую популяцию.

При определении целевой популяции важно также учитывать условия оказания медицинской помощи будущей СППВР – скорая помощь, поликлиника, диспансер, больница, научный центр и т.п. Очевидно, что популяции систематически отличаются в зависимости от места оказания врачебной помощи. Например, преваленс идентифицируемого состояния при скрининге является низким, при диагностике – высоким.

Контролем в задачах диагностики следует выбрать «золотой стандарт» – наиболее точный из имеющихся способов идентификации, который следует использовать как верифицирующий изучаемые состояния. В задачах прогнозирования таким контрольным идентификатором выступает собственно прогнозируемое событие, которое должно определяться по заранее сформулированным критериям. Например, идентификация рецидива заболевания должна выполняться по критериям, принятым в клинических рекомендациях. В задачах поиска подгрупп пациентов контрольная идентификация не требуется (применяется машинное обучение «без учителя»).

Оцениваемый исход (состояние, нозология) в задачах диагностики и прогнозирования должен быть определен на основании принятых критериев

диагностики, чтобы верификация выполнялась надежно. Негативный пример: разрабатывается система анализа медицинских изображений для выявления патологических образований. Системе удалось выявить дополнительные (по отношению к результатам врача) случаи, которые были расценены врачами как вероятно позитивные, что представляется разработчиками как успех системы. На самом деле, верификация пока не выполнена, и нельзя исключить, что вновь выявленные случаи являются ложноположительными.

Важнейшим компонентом этапа постановки задачи является согласование с врачами подхода к разработке иСППВР – имитирующего или моделирующего. Имитирующие модели неинтерпретируемы и часто поэтому заведомо неприемлемы для врачей. Этим недостатком не обладают моделирующие рассуждения экспертные системы.

2. Разработка дизайна исследования

В контексте разработки имитирующих иСППВР под дизайном исследования следует понимать способ сбора данных. Для моделирующих иСППВР, основанных на знаниях, данные не нужны.

Первой задачей является определение способа формирования несмещенной (репрезентативной) выборки из целевой популяции. Рекомендуемые способы формирования выборки – случайный и сплошной. Эти способы могут быть использованы в исследованиях любого дизайна – одномоментных, проспективных, ретроспективных. Главное, чего нужно избегать – произвольное формирование выборки, которая с большой вероятностью будет смещенной. В идеале желательно организовывать многоцентровое исследование, с тем, чтобы данные были репрезентативны по отношению не к одному источнику (медицинской организации), а к нескольким организациям одного уровня медицинской помощи.

При разработке диагностических систем очень важно предусмотреть сбор позитивных и негативных случаев из общей целевой популяции, с учетом фактического преваленса идентифицируемого состояния. Частой ошибкой, приводящей к систематическому смещению результатов, является набор позитивных случаев из клиники, а негативных – из контрольной группы здоровых лиц. При разработке систем прогнозирования обязательна синхронизация когорты пациентов (клинические случаи должны быть наблюдаемы с одного и того же момента в истории заболевания, например, с момента постановки диагноза, с момента выполнения хирургической операции). В случае поиска подгрупп пациентов также важно понимать



происхождение данных и их репрезентативность. Если используются данные ранее выполненного клинического или популяционного исследования, то известны критерии включения и исключения пациентов. Если же используется просто цифровой архив медицинской организации, то качество этих данных будет заведомо спорным и зависящим от характеристик этой конкретной организации.

Для любой разработки, использующей данные, методы генерации данных должны быть гармонизированы или синхронизированы. Например, лабораторные тесты должны быть выполнены одинаковыми лабораторными методами, используемые клинические шкалы и опросники должны быть валидизированными, томографы – иметь близкие технические характеристики. Следует обратить внимание и на возможность исторического смещения в данных – изменения критериев оценки состояний, способов диагностики и т.п.

Важно определить, что является объектом исследования – пациент, его орган, курс лечения и т.д. Например, в задачах офтальмологии часто объектом исследования служит глаз, в онкологии – курс химиотерапии, что создает фактически непреодолимые трудности в анализе данных, смешивая независимые (пациентов) и зависимые наблюдения (парные органы, курсы терапии). В таких задачах рекомендуется зафиксировать объект по отношению к пациенту (например, всегда включать данные от «более больного» глаза), брать в анализ случайный курс терапии и т.п. Это позволит избежать нарушения предположения основных многомерных методов анализа данных о независимости объектов выборки, а также обоснованно присовокупить к данным организменные показатели (пол, возраст, вес, АД и проч.).

Заранее должен быть продуман способ обработки пропусков в данных. При использовании ретроспективных данных часто случаи с пропусками просто не включают в выборку, что вносит систематическое смещение в данные, а, следовательно, и в результаты.

Немаловажным компонентом дизайна является расчет необходимого объема выборки, который в подобных задачах в основном определяется приемлемой точностью оценок эффективности разрабатываемого продукта.

Разработка дизайна исследования заканчивается разработкой протокола сбора данных.

3. Сбор данных. В проспективном или одномоментном исследовании качество данных является более высоким, т.к. меньше пропусков в данных,

больше единообразие измерений и т.п. В ретроспективном дизайне эти проблемы неустранимы.

4. Построение «решателя» с использованием различных методов машинного обучения («с учителем» или «без учителя»), методов извлечения и представления экспертных знаний. Любая модель имеет смысл, если она с приемлемым качеством описывает реальность. Не останавливаясь здесь на методах, перечислим основные критерии (метрики) качества разрабатываемой модели:

а) статистически значимая сопряженность результатов применения модели и «золотого стандарта»;

б) оценки чувствительности (в машинном обучении обычно называется откликом (англ. *recall*) и специфичности – устойчивых, не зависящих от преваленса идентифицируемого состояния в выборке, операционных характеристик модели, и их доверительных интервалов (95%, а лучше – 99%);

в) оценки прогностической ценности положительного и отрицательного результата (ПЦПР и ПЦОР соответственно) – зависящих от преваленса операционных характеристик модели – и их ДИ (99%, 95%). В машинном обучении ПЦПР обычно называется английским термином *precision*. В случае, если выборка репрезентативна по отношению к целевой популяции в отношении преваленса (так происходит обычно при использовании сплошного или случайного метода формирования выборки), эта проблема практически отсутствует. Однако если в исследование отдельно набирались позитивные и негативные случаи, поправка на преваленс необходима. Прогностические ценности крайне важны, т.к. именно с ними работает врач, оценивая результат конкретного больного с учетом вероятности гипер- и гиподиагностики;

г) точность модели. Под этим обычно понимают отношение суммы истинно-положительных и истинно-отрицательных результатов к общему числу наблюдений в выборке. В машинном обучении эта метрика обычно называется *accuracy*. Иногда под точностью понимают среднее между Ч и С. Для точности также могут быть рассчитаны ДИ (99%, 95%);

д) в некоторых случаях для выбора оптимальной отрезной точки необходимым бывает проведение ROC-анализа. Критериями оптимальности выбора могут быть минимизация ошибок I или II рода, оптимизация их соотношения, максимизация их суммы (критерий Юдена) и т.п. Обычно в задачах диагностики оптимизируется Ч, в задачах скрининга – С. При прогнозировании часто предпочтение отдается гипердиагностике.



О том, какие значения эффективности можно считать успехом разработки, единого мнения не существует. Обычно полагают, что модель приемлема, если ее точность равна или превышает 85%. По нашему мнению, модель хороша, если ДИ для любой из оценок эффективности лежит в диапазоне 85%–100%. При этом очевидно, что если любой ДИ включает 50%, то модель вовсе неработоспособна, и проще бросать монетку.

5. Аналитическая (техническая) валидизация полученной модели. Может выполняться с использованием кросс-валидации, скользящего экзамена, независимой тестовой выборки (при построении экспертных систем обучающая выборка не нужна, но тестовая обязательна). Поскольку известно, что многие модели, полученные в результате применения машинного обучения, являются переобученными (т.е. настроенными на конкретные данные, использованные при обучении), то может применяться и анализ чувствительности к данным. Для этого проводится новое разделение исходной выборки на обучающую и тестовую (процедура *shuffle*), заново строится модель и оцениваются ее характеристики. Далее устанавливается точность совпадения результатов старой и новой моделей. Процесс может повторяться сколь угодно много раз. Заменой такого трудоемкого процесса является оценка неопределенности (ДИ) характеристик модели, о чем было сказано выше.

6. Программная реализация модели возможна в форме модуля локальных информационных систем, облачного сервиса, отдельного программного продукта (веб-приложения, мобильного приложения). Важно, чтобы в интерфейсе пользователю-врачу предоставлялись решающие правила на естественном языке, а также – обязательно – значения операционных характеристик модели. В качестве последних предпочтительны прогностические ценности – показатели, с которыми работает врач, показывающие степень уверенности системы в выдаваемом заключении, и диапазоны их неопределенности (ДИ).

7. Публикация результатов не менее, чем в двух профильных рецензируемых научных журналах – в журнале в соответствии с рассматриваемой медицинской проблемой и в журнале по медицинской информатике (например, в журнале «Врач и информационные технологии»). Как и всегда, желательно выбирать наиболее высокоимпактные профильные журналы, как международные, так и российские.

Для моделирующих (экспертных) СППВР актуальны этапы 1, 4–7, для информационно-справочных – этапы 1, 5–7.

На этом этап разработки и производства может считаться законченным, однако, как следует из *рис. 1* – это лишь первая стадия ЖЦ МТ. Успешная разработка диагностических и прогностических систем лишь означает, что разработанный на основе модели программный продукт работает правильно, однако ценность такой системы на данном этапе неизвестна.

II. Клинические исследования (клиническая валидация)

С позиций доказательной медицины (ДМ) независимо от актуального порядка регистрации ПО в национальной юрисдикции за производством МТ должны последовать клинические испытания СППВР. Они необходимы для создания доказательной базы данной МТ. Именно ей могут (или не могут, если испытания проведены плохо) доверять врачи, практикующие ДМ.

Клинические испытания – трудоемкий и длительный этап ЖЦ МТ. Финансовые затраты обычно ложатся на производителя МТ, который становится спонсором КИ. Доля клинических исследований III фазы (РКИ) в общем бюджете разработки фармацевтических продуктов достигает 70%. Вряд ли имеет смысл сопоставлять общие бюджеты разработки фармацевтических субстанций и ИТ продуктов, но средняя стоимость РКИ составляет 19 млн. долл. [14]. Таким образом, коммерческая ИТ-компания, разработавшая СППВР, должна профинансировать его клинические испытания. Она может это сделать за счет собственного бюджета либо получить грант в каком-либо фонде. РФФИ, РНФ в своих документах декларируют, что занимаются финансированием только фундаментальных исследований, в то время как клинические испытания – исключительно прикладные исследования. Фонды Сколково, РВК, Бортника финансируют прикладные исследования, от них можно пробовать получать финансирование. Если же продукт произведен в академическом учреждении, то в это случае должен быть доступен механизм получения финансирования клинических испытаний. Решением в таких ситуациях представляется прямое проектное государственное финансирование (путем госзаказа или госзакупки) через ведомство (Минздрав, Минобрнауки и др.) либо инициативное обращение академической организации за финансированием в институты инновационного развития.



В испытаниях диагностических СППВР необходимо оценить не правильность работы системы (что должно было быть ранее оценено при разработке и аналитической валидации), а ее эффективность и безопасность в сравнении с рутинно используемым методом. При этом проверке может быть подвергнута одна из гипотез:

- гипотеза превосходства по эффективности, безопасности – при предполагаемой замене старой МТ на СППВР либо при использовании СППВР в дополнение (параллельно или последовательно) к старому способу диагностики;

- гипотеза неуступающей эффективности (или эквивалентности) по эффективности, безопасности – при планировании замены инвазивного, дорогого или другого недостаточно хорошего диагностического теста на СППВР.

Для прогностических систем цель испытаний – это обычно доказательство преимущества персонализированного ведения пациента над рутинным (стандартным) ведением на основе оценки рисков исходов. Часто рутинная практика представляет собой отдельную МТ, совокупность МТ или даже отсутствие какой-либо МТ. До проведения испытания врачами на основе полученных при разработке

системы решающих правил должна быть сформулирована персонализированная тактика ведения пациентов, которая и будет представлять собой испытываемое вмешательство.

После формулировки гипотезы следует рассчитать необходимый объем выборки с целью обеспечения достаточной статистической мощности испытания. Он делается на основе ряда параметров, основным из которых является клинически значимая величина эффекта, которую должны определить врачи.

Приведем основные черты дизайна испытаний для разных типов поддерживаемых решений (таблица 1).

Прогностические СППВР подлежат испытанию в РКИ, проведение которых является организационно сложным. В частности, должна быть обоснованной их этичность, в том числе соблюдаться принцип эквипойза. Заметим, что в то же время для диагностических систем РКИ не только не нужны (о необходимости РКИ для диагностических методов неверно пишут многие авторы), но и невозможны, поскольку изучается одна группа – выборка из целевой популяции пациентов.

Чтобы избежать систематических смещений в клинических испытаниях, рекомендуется пользоваться

Таблица 1

Основные черты дизайна клинических испытаний для разных типов интеллектуальных СППВР

	Диагностическая СППВР	Прогнозирующая СППВР
Медицинская задача	Диагностика, скрининг	Профилактика (первичная, вторичная), выбор метода лечения
Дизайн испытания	Одномоментный (поперечный)	РКИ (проспективный сравнительный)
Популяция	Целевая	Целевая
Число выборок из популяции	Одна	Одна
Число групп	Одна	Две
Способ распределения на группы	Не требуется	Рандомизация
Испытываемое вмешательство	СППВР	Персонализированное ведение пациентов на основе прогноза СППВР
Контрольное вмешательство	Рутинно используемый способ диагностики	Рутинное (стандартное) ведение пациентов
Основные оцениваемые показатели	ДЧ, ДС, ПЦПР, ПЦОР и их ДИ	Относительный риск, изменение абсолютного риска, отношение угроз и их ДИ
«Ослепление» вмешательств	Невозможно	Хотя бы одинарное (ослепление пациентов)
«Ослепление» при оценке эффектов	Взаимное маскирование результатов сравниваемых диагностических тестов	Независимая оценка эффектов
Оценка выбывания	Анализ наборов данных с учетом и без учета неопределенных результатов	Анализ наборов данных per protocol и intention-to-treat
Минимизация конфликта интересов	Найм независимой исследовательской организации для проведения исследования; публикация исследователями результатов исследования независимо от его успешности и без необходимости получения разрешения спонсора на публикацию	



опросниками для оценки методологического качества РКИ [15] и одномоментных исследований диагностических тестов [16], а при подготовке публикации пользоваться шаблонами CONSORT [17] и STARD [18] соответственно.

Итак, клинические испытания должны закончиться оценкой эффективности и безопасности СППВР, публикации должны появиться в профильных научных изданиях.

III. Государственная регистрация

В соответствии с Основами законодательства Российской Федерации об охране здоровья граждан (статья 43) в практике здравоохранения можно использовать только разрешенные к применению методы профилактики, диагностики, лечения, скрининга. СППВР, будучи ПО, в соответствии с действующими регуляторными нормами в РФ рассматриваются как программные медицинские изделия. Как заявлено на официальном веб-сайте Росздравнадзора, регистрация медицинских изделий – государственная процедура, цель которой разрешить выпуск на российский рынок качественных и безопасных изделий. Из этого вероятно можно сделать вывод, что если пребывание на рынке (т.е. коммерциализация продукта) не предусмотрена, то регистрация не является обязательной. Одновременно частью 4 статьи 38 Федерального закона от 21.11.2011 № 323-ФЗ «Об основах охраны здоровья граждан в Российской Федерации» установлено, что на территории Российской Федерации разрешается обращение медицинских изделий, зарегистрированных в порядке, установленном Правительством Российской Федерации, уполномоченным им федеральным органом исполнительной власти. Что означает термин «обращение», в законе не определено. Вместе с тем, программные медицинские изделия необходимо подвергать обязательной государственной регистрации с присвоением ему определенного класса риска [19]. Для проведения же регистрации в соответствии с действующими документами РФ [20] и ЕАЭС [21] необходимо, помимо других испытаний, провести и клинические испытания с целью оценки безопасности и эффективности. Более ранний документ этого направления – Приказ Министерства здравоохранения РФ от 9 января 2014 г. № 2н «Об утверждении Порядка проведения оценки соответствия медицинских изделий в форме технических испытаний, токсикологических исследований, клинических испытаний в целях государственной регистрации медицинских

изделий» [22]. В США в настоящее время ведется дискуссия о том, как именно следует определять СППВР [23].

Реальная практика в настоящее время, по видимому, сводится к тому, что коммерческие СППВР проходят этап регистрации, в то время как свободно распространяемые – нет. В Государственном реестре медицинских изделий (<http://www.roszdravnadzor.ru/services/misearch>) нам не удалось обнаружить зарегистрированные СППВР, однако не исключено, что мы выполнили неоптимальный поиск.

Таким образом, пока необходимость и порядок регистрации программных СППВР остаются неопределенными.

IV. Клинико-экономический анализ

Клинико-экономический анализ (КЭА) выполняется обычно со следующими целями:

- для включения клинико-экономического обоснования МТ в клинические рекомендации;
- для включения МТ в ограничительные перечни МТ, гарантированно закупаемых государством либо иным плательщиком.

КЭА основан на результатах сравнительных клинических исследований. При этом если клинические испытания носят международный характер, т.е. обычно их результаты валидны во всех странах мира, то КЭА выполняется исключительно на национальном уровне. Это связано с тем, что организационно-экономические аспекты систем здравоохранения в разных странах мира несопоставимы – действуют разные системы страхования, лекарственного обеспечения, стандарты помощи и т.п.

КЭА выполняется обычно не путем анализа «сырых» данных, а методами моделирования (деревья решений, марковские модели и др.). В модели учитываются следующие виды затрат: прямые медицинские (постоянные, переменные), прямые немедицинские, не прямые, нематериальные. Основными методами КЭА являются оценка минимизации затрат (при равной эффективности и безопасности сравниваемых вмешательств) и анализ соотношения затрат и эффективности (при разных эффективности и безопасности). Соответственно, если по результатам клинического испытания СППВР предлагается использовать взамен неоптимального способа диагностики (дорогостоящего, инвазивного, длительного, требующего специальной квалификации и т.п.), и при этом она обладает не меньшей или эквивалентной эффективностью и безопасностью,



то может проводиться оценка минимизации затрат. Это послужит аргументом в пользу внедрения такой СППВР. Если же СППВР более эффективна, но дороже, чем рутинная технология, то следует проводить анализ соотношения эффекта и затрат. Если СППВР более эффективна, но дешевле рутинной технологии, то КЭА не требуется, и такая технология называется доминирующей.

V. Оценка медицинских технологий и дальнейшее продвижение

Комплексная экспертиза – оценка медицинских технологий (ОМТ) – включает (но может не ограничиваться) рассмотрение как научных обоснований эффективности и безопасности технологий, полученных в клинических исследованиях, так и результатов КЭА. В РФ ОМТ выполняется пока только для лекарственных препаратов при формировании Перечня ЖНВЛП, однако в перспективе она должна охватить все типы вмешательств, финансируемых государством. При успешных результатах ОМТ можно рассчитывать на включение СППВР в клинические рекомендации и финансирование их закупок государством или иными плательщиками (ведомствами, страховыми компаниями и др.).

Заключение

Любые СППВР, и прежде всего интеллектуальные, должны рассматриваться как медицинские технологии (независимо от требований регулятора). В их жизненном цикле можно выделить стадии разработки, клинической валидации, государственной регистрации, клинико-экономического анализа и оценки медицинских технологий. Клиническая валидация заключается в получении доказательств улучшения исходов пациентов при применении СППВР и ее безопасности по сравнению с рутинно применяющимися методами. Клинико-экономический анализ и последующая комплексная оценка медицинских технологий могут обосновать включение СППВР в клинические рекомендации и финансирование закупок СППВР государством или иными плательщиками.

Пример: интеллектуальная СППВР для персонализации ведения пациентов с болезнью Иценко-Кушинга (БИК) после успешного нейрохирургического лечения [24, 25]

Этап 1 – разработка и производство ИСППВР

Анализировались ретроспективно собранные данные 219 пациентов с БИК, которые получили

эффективное нейрохирургическое лечение. Выборка сформирована сплошным методом. Когорта синхронизирована по дате хирургического лечения. Частота рецидивов к трем годам после операции составила 19,5%, 95% ДИ [14,4%; 24,5%] (оценка методом Каплана-Мейера).

Для построения математических моделей использовались статистические методы и методы ИНС. Статистическими методами построить модели приемлемой точности не удалось. Построена эффективная ИНС для прогнозирования ремиссии/рецидива в течение трех лет после успешной нейрохирургической операции.

Операционные характеристики ИНС составили:

- чувствительность – 75% [60%; 86%];
- специфичность – 97% [93%; 99%];
- прогностическая ценность положительного результата – 85% [71%; 94%];
- прогностическая ценность отрицательного результата – 93% [89%; 96%];
- общая точность модели составляет 92% [87%; 95%].

На основе предсказательных ценностей сформулированы два решающих правила:

- «Если математической моделью прогнозируется ремиссия, то пациент с вероятностью 93% [89%; 96%] будет иметь ремиссию в течение 3 лет».
- «Если математической моделью прогнозируется рецидив, то пациент с вероятностью 85% [71%; 94%] будет иметь рецидив в срок до 3 лет».

Далее выполнена программная реализация модели – разработано свободно доступное веб-приложение – <http://medcalc.appspot.com>. Планируется также встроить программный продукт в госпитальную МИС.

Статьи по результатам разработки приняты в печать в профильном журнале по эндокринологии [24] и направлены в профильный журнал по информационным технологиям в медицине [25].

Этап 2 – Клиническое исследование

На основе построенных решающих правил может быть предложено модифицировать алгоритм послеоперационного ведения пациентов следующим образом: при прогнозе рецидива (таких пациентов около 20%) в течение трех лет после операции пациенту предлагается посещать врача не каждые 6 месяцев (так при стандартном ведении), а каждые 3 месяца. В то же время пациентам, для которых прогнозируется ремиссия (таких пациентов примерно 80%), предлагается посещать врача 1 раз в год. Таким образом, может быть реализована



идея персонализации послеоперационного ведения пациентов.

Для оценки эффективности и безопасности СППВР необходимо будет провести РКИ со следующими параметрами:

- популяция – все пациенты с успешной нейрохирургической операцией;
- вмешательство – послеоперационное ведение пациентов с использованием прогноза, предлагаемого СППВР (активный контроль каждые 3 мес. при прогнозе рецидива, каждые 12 мес. при прогнозе ремиссии);
- контроль – стандартное ведение пациентов (контроль каждые 6 мес.);
- оцениваемые исходы – любые госпитализации, связанные с основным заболеванием (инфаркт миокарда, ОНМК, тяжелые инфекции, болезни почек, низкотравматические переломы), летальность.

Проверяемые гипотезы:

1) Персонализированное ведение пациентов обладает не меньшей эффективностью и безопасностью, чем стандартное ведение.

2) Персонализированное ведение пациентов более эффективно и безопасно, чем стандартное ведение.

Этап 3 – Клинико-экономический анализ

Анализ может быть проведен путем моделирования на горизонте, например, 5 лет путем расчета коэффициента «затраты/эффективность» – отношения разности затрат при персонализированном и стандартном ведении пациентов к разности их эффектов. Учитываемые затраты могут быть определены на основе клинических рекомендаций или стандарта медицинской помощи. Обычно к ним как минимум относятся прямые медицинские затраты – на амбулаторные визиты, плановые и экстренные госпитализации, скорую и неотложную помощь. Также могут быть учтены немедицинские затраты, связанные с временной утратой трудоспособности, инвалидностью. Не исключено, что при КЭА обнаружится, что персонализированный подход представляет собой т.н. доминирующую технологию, превосходящую стандартную как по эффективности и безопасности, так и по затратам.

ЛИТЕРАТУРА



1. Гусев А.В., Зарубина Т.В. Поддержка принятия врачебных решений в медицинских информационных системах медицинской организации // Врач и информационные технологии. – 2017. – № 2. – С. 60–72.
2. Garg A.X., Adhikari N.K.J., McDonald H. et al. Effects of Computerized Clinical Decision Support Systems on Practitioner Performance and Patient Outcomes. A Systematic Review // JAMA. 2005; 293: 1223–1238.
3. Tan K., Dear P.R.F., Newell S.J. Clinical decision support systems for neonatal care. Cochrane Database of Systematic Reviews 2005, Issue 2. Art. № CD004211. DOI: 10.1002/14651858.CD004211.pub2.
4. Jamal A., McKenzie K., Clark M. The impact of health information technology on the quality of medical and health care: a systematic review // Health Inf Manag. 2009;38(3):26–37.
5. Black A.D., Car J., Pagliari C. et al. The Impact of eHealth on the Quality and Safety of Health Care: A Systematic Overview // PLoS Med. 2011, 8(1): e1000387. doi:10.1371/journal.pmed.1000387.
6. Bright T.J., Wong A., Dhurjati R. et al. Effect of Clinical Decision-Support Systems. A Systematic Review // Ann Intern Med. 2012;157: 29–43.
7. Moja L., Kwag K.H., Lytras T. et al. Effectiveness of Computerized Decision Support Systems Linked to Electronic Health Records: A Systematic Review and Meta-Analysis // Am J Public Health. 2014;104: e12–e22. doi:10.2105/AJPH.2014.302164.
8. Varghese J., Kleine M., Gessner S.D.I. et al. Effects of computerized decision support system implementations on patient outcomes in inpatient care: a systematic review // Journal of the American Medical Informatics Association, 25(5), 2018, 593–602. doi: 10.1093/jamia/ocx100.
9. Pombo N., Arasjo P., Viana J. Knowledge discovery in clinical decision support systems for pain management: a systematic review. Artif Intell Med. 2014 Jan; 60(1):1–11. doi: 10.1016/j.artmed.2013.11.005.
10. Tomaselli Muensterman E., Tisdale J.E. Predictive Analytics for Identification of Patients at Risk for QT Interval Prolongation: A Systematic Review. Pharmacotherapy. 2018 Aug;38(8):813–821. doi: 10.1002/phar.2146.



11. Arani L.A., Hosseini A., Asadi F., Masoud S.A., Nazemi E. Intelligent Computer Systems for Multiple Sclerosis Diagnosis: a Systematic Review of Reasoning Techniques and Methods. *Acta Inform Med.* 2018 Dec; 26(4): 258–264. doi: 10.5455/aim.2018.26.258–264.
12. Liu X., Faes L., Kale A.U. et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis // *The Lancet Digital Health.* – 2019. – Т. 1. – № 6. – С. e271–e297. 10.1016/S2589–7500 (19) 30123-2.
13. Software as a Medical Device (SAMd): Clinical Evaluation Guidance for Industry and Food and Drug Administration Staff. Document issued on December 8, 2017. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/software-medical-device-samd-clinical-evaluation-guidance-industry-and-food-and-drug-administration>. Доступ 15.12.2019.
14. Moore T.J., Zhang H., Anderson G., Alexander G.C. Estimated Costs of Pivotal Trials for Novel Therapeutic Agents Approved by the US Food and Drug Administration, 2015–2016 // *JAMA Intern Med.* 2018; 178(11): 1451–1457. doi: 10.1001/jamainternmed.2018.3931.
15. Реброва О.Ю., Федяева В.К., Хачатрян Г.П. Адаптация и валидизация вопросника для оценки риска систематических ошибок в рандомизированных контролируемых испытаниях // *Медицинские технологии. Оценка и выбор.* – 2015. – № 1 (19). – С. 9–17.
16. Реброва О.Ю., Федяева В.К. Оценка риска систематических ошибок в одномоментных исследованиях диагностических тестов: русскоязычная версия вопросника QUADAS // *Медицинские технологии. Оценка и выбор.* – 2017. – Т. 27. – № 1. – С. 11–14.
17. Moher D., Hopewell S., Schulz K.F. et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials // *BMJ* 2010;340: c869 doi: 10.1136/bmj.c869.
18. Bossuyt P.M., Reitsma J.B., Bruns D.E. et al. STARD2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies // *BMJ.* 2015 Oct 28; 351: h5527. doi: 10.1136/bmj.h5527.
19. Столбов А.П. О классификации рисков применения медицинского программного обеспечения в Евразийском экономическом союзе // *Врач и информационные технологии.* – 2019. – № 3. – С. 22–31.
20. Постановление Правительства РФ от 27 декабря 2012 г. № 1416 (ред. от 31.05.2018) «Об утверждении Правил государственной регистрации медицинских изделий». Справочная правовая система КонсультантПлюс. http://www.consultant.ru/document/cons_doc_LAW_140066/. Доступ 15.12.2019.
21. Решение Совета Евразийской экономической комиссии от 12.02.2016 № 29 «О Правилах проведения клинических и клинико-лабораторных испытаний (исследований) медицинских изделий». Справочная правовая система КонсультантПлюс. http://www.consultant.ru/document/cons_doc_LAW_197969/ Доступ 15.12.2019.
22. Приказ Министерства здравоохранения РФ от 9 января 2014 г. № 2н «Об утверждении Порядка проведения оценки соответствия медицинских изделий в форме технических испытаний, токсикологических исследований, клинических испытаний в целях государственной регистрации медицинских изделий». URL: <http://base.garant.ru/70631448>. Доступ 15.12.2019.
23. Clinical Decision Support Software Draft Guidance for Industry and Food and Drug Administration Staff. Document issued on September 27, 2019. URL: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-decision-support-software>. Доступ 15.12.2019.
24. Nadezhdina E.Y., Rebrova O. Yu., Grigoriev A.Y. et al. Prediction of recurrence and remission within 3 years in patients with Cushing disease after successful transnasal adenectomy // *Pituitary.* – 2019. – Vol. 22. – № 6. – P. 574–580. doi: 10.1007/s11102-019-00985-5.
25. Надеждина Е.Ю., Реброва О.Ю., Антюх М.С., Григорьев А.Ю. Прогнозирование рецидива у пациентов с болезнью Иценко-Кушинга после успешной эндоскопической трансназальной аденомэктомии: нейросетевая модель и ее программная реализация // *Врач и информационные технологии.* – 2019. – № 4. – С. 65–71.