

УДК 616-071:006.86(045)

В. Д. Кузовик, д-р техн. наук, проф.,
Л. А. Кошечая, канд. техн. наук, доц.

НЕКОТОРЫЕ МЕТОДОЛОГИЧЕСКИЕ АСПЕКТЫ ОЦЕНКИ ЭФФЕКТИВНОСТИ МЕДИЦИНСКИХ ВМЕШАТЕЛЬСТВ

Институт электроники и систем управления НАУ, e-mail: kmice@kv.ukrtel.net

Показана важность соблюдения требований международных стандартов при проведении клинических испытаний для обеспечения достоверности их результатов, что способствует взаимному признанию данных клинических испытаний, а с развитием интеграционных процессов, становится актуальным при разработке и внедрении стандартов медицины в практику.

Ключевые слова: клиническое исследование, дизайн исследования, статистический анализ, клиническая значимость, систематизированный обзор, мета-анализ.

Введение и постановка задачи. Усиливающаяся в последние годы коммерциализация медицины привела к появлению разнообразных медицинских методов и средств лечения, которые внедряются в клиническую практику всеми возможными способами. Часто маркетинговая декларация не соответствует реальности, а агрессивная рекламная политика фирм, занимающихся лечением, требует от врача знаний, которые бы позволили ему самому убедиться в корректности различных медицинских вмешательств.

Эффективность многих медицинских вмешательств воспринимается бесспорной, часто не из-за наличия надежных доказательств, а скорее по традиции. Так, например, по разным оценкам достоверно доказана эффективность лишь 5 –10 % применяемых в современной практике лекарственных средств [1].

В соматической медицине для оценки эффективности вмешательства применяют лабораторные эксперименты с привлечением электрофизиологических, биохимических и других методик, в результате которых можно оценить *косвенные эффекты*. К косвенным эффектам (косвенным критериям эффективности) относят положительные изменения какого-либо исследуемого показателя, например, снижение повышенного артериального давления, нормализацию уровня каких-либо биологических субстанций в крови, восстановление активности ферментов, усиление метаболизма глюкозы в лобных отделах мозга и т.п. Косвенные (их еще называют суррогатными) эффекты имеют значение тогда, когда научно доказано, что эти показатели действительно имеют отношение к процессу медицинского вмешательства и полученные данные можно обобщать и экстраполировать. Кроме косвенных, существуют и *прямые критерии эффективности*, которыми характеризуются выздоровление, снижение летальности и осложнений, сокращение срока госпитализации, улучшение качества жизни. Часто проверку эффективности осуществляют в структуре «до–после» исследований, занимающих в иерархии доказательности далеко не первое место [2].

Значимость любого медицинского исследования определяется рядом факторов, среди которых особое место принадлежит *научному обоснованию* этого исследования и возможности *практического применения* его результатов. Научная обоснованность медицинского исследования является *мерой доверия* к его результатам. Результаты исследования, не соответствующего строгим научным принципам, теряют универсальность, так как они получены в условиях данного клинического наблюдения, и поэтому выявить существующие объективные закономерности и взаимосвязи можно только стандартизируя и формализуя условия эксперимента. Стандарты научной обоснованности включают строго определенные требования к проведению исследования, обработке его результатов и оценке достоверности, без соблюдения которых результаты клинических исследований теряют смысл, а то и могут быть вредны.

Основной материал. В практике медицинской диагностики в последнее время интенсивно развиваются неинвазивные методы и средства оценки психофизиологического состояния человека. Научной группой кафедры биокибернетики и аэрокосмической медицины Национального авиационного университета разработан метод диагностирования психосоматического состояния операторов экстремальных видов деятельности (летчиков, полярников, спортсменов) с использованием *неинвазивного кефалознцефалографического комплекса*. Метод позволяет *количественно* определять выраженность эффектов лечебных воздействий при реализации соответствующих подходов к технологиям сбора, анализа, обобщения полученной медицинской информации, которые регламентируются стандартными правилами проведения клинических исследований (клинические исследования и клинические испытания – синонимы [3]). В связи с предстоящим внедрением в медицинскую практику данного комплекса встала проблема научной организации его клинических испытаний.

Необходимость создания единого подхода к оценке результатов медицинских исследований привела к выработке *единых критериев* научной обоснованности. «Золотым стандартом» проведения клинического исследования в этом смысле является клиническое исследование, выполненное в соответствии с правилами *ICH-GCP* [3]. Правила *GCP* — это свод положений, в соответствии с которыми планируются и проводятся клинические испытания, а также анализируются и сообщаются их результаты. Клиническое исследование, выполненное без соблюдения требований *ICH-GCP*, не является доказательным, научно обоснованным, а отражает лишь субъективное отношение исследователя к рассматриваемому явлению, чаще всего с учетом уже известных данных авторитетных научных исследований.

Одним из ключевых моментов исследований является их *достоверность*. Для обеспечения доказательности результатов, получаемых в исследованиях, необходимо выбрать адекватную *методику исследования* (дизайн исследования) и *методы статистического анализа*.

Целесообразно рассматривать организацию клинического испытания, как и любого другого испытания, поэтапно. Основные этапы включают в себя: планирование, проведение, обработку данных и принятие решения.

Планирование исследования. На этом этапе в соответствии со стандартом [3] составляется *протокол* – основной документ, содержащий предпосылку, обоснование, цель исследования и описывающий его дизайн, методологию и организацию исследования, включая статистические подходы к планированию и анализу исследования.

Дизайн исследования (англ. design – проектировать, конструировать) – это метод проектирования исследования, который зависит от его цели и выдвигаемой гипотезы. Дизайном предусматривается, чтобы каждый элемент исследования выполнялся в соответствии с надлежащими методологическими особенностями, минимизируя при этом погрешности результатов исследований.

С точки зрения *организации* можно планировать дизайны различных видов, например, наблюдательные исследования или экспериментальные (контролируемые, с преднамеренным вмешательством). Выбор дизайна исследования зависит от ряда факторов, основными из которых являются: уровень предыдущих знаний, изучаемый клинический исход и категория клинического вопроса.

При организации *контролируемых исследований* обязательно формируют одну группу пациентов, которая подвергается тому или иному медицинскому вмешательству (основная), и контрольную. Слово «контролируемые» подразумевает, что при организации таких исследований учтено действие факторов, способных повлиять на результат, и находится под контролем. Методика их проведения предполагает сравнение двух или более видов вмешательств в однородных группах пациентов. Однородность группы определяется

критериями отбора, определенными при планировании исследования и утвержденными в его протоколе.

Можно организовать *обсервационные исследования*, которые основаны на сборе и оценке результатов реальной клинической практики. В обсервационных исследованиях формируется новое клиническое знание и формулируется соответствующая гипотеза, а в клинических испытаниях производится проверка выдвигаемой гипотезы.

Хотя контролируемые исследования более предпочтительны, однако хорошо организованное обсервационное исследование может быть более информативным, чем плохо выполненное контролируемое [4].

Проведение исследования. После того как выбран дизайн исследования, необходимо принять определенные меры для его оптимизации с целью *минимизации систематической погрешности (смещения) оценки результатов* и получения *максимально возможной их точности*, а также обеспечения обоснованной возможности *распространения результатов на всю генеральную совокупность*.

Оценивая структуру испытания, необходимо контролировать, чтобы его результат не стал следствием какой-нибудь ошибки. Достоверные исследования должны быть свободны от смещений, возникающих при неправильной организации исследования: различий в сравниваемых группах, связанных со способом формирования выборки, влияния побочных факторов (сопутствующего медицинского вмешательства помимо исследуемого), выбывания пациентов из исследования, метода выявления исходов. Отсутствие *систематической погрешности* в испытании называют *обоснованностью (validity)*. Наличие такого типа погрешности может привести к полной непригодности результатов наблюдения. С целью компенсации смещения результата и получения надежных данных вносят *поправки*, учитывающие значение систематической погрешности, а также используют организационные меры.

Эффективным средством для борьбы со смещением является рандомизация. *Рандомизация* – это процесс случайного распределения пациентов в группы исследования с целью исключения возможной необъективности и вероятного смещения оценки результата. При помощи рандомизации исследователь независимо формирует группы.

Наиболее близкий к классическому научному эксперименту вид клинических испытаний, так называемый «золотой стандарт» процедуры проверки эффективности вмешательств – *рандомизированные контролируемые испытания* (рисунок). Рандомизированные испытания при правильном планировании обладают наибольшей доказательностью. Их делят на открытые и слепые (виды маскирования).

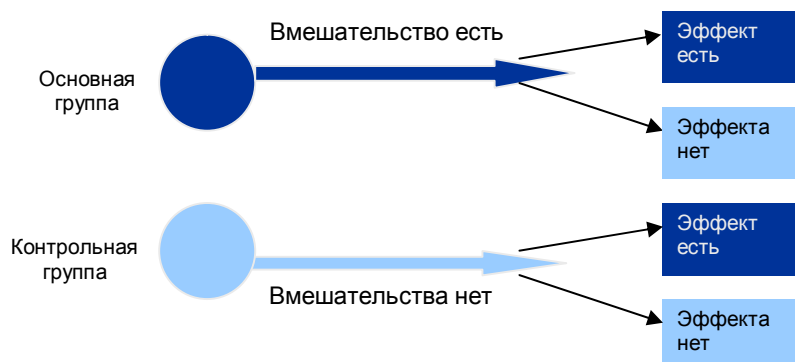


Схема рандомизированного контролируемого испытания

Открытое рандомизируемое контролируемое испытание — это исследование, при котором и пациент, и врач сразу после проведения рандомизации узнают о том, какой вид вмешательства будет применен. Вид маскирования может быть *простым* (пациенту не

известно о виде применяемого вмешательства), *двойным* (ни врач, ни пациент не знают, какое из вмешательств будет использовано), *тройным* (о виде применяемого вмешательства не знают ни пациент, ни врач, ни статистик, обрабатывающий результаты исследования). Процесс «ослепления» участников испытания предназначен для устранения смещения субъективной оценки результата.

К проблемам, сопровождающим рандомизированные исследования, можно отнести сложность организации процесса исследования, значительные временные и финансовые затраты, низкую обобщаемость получаемых результатов.

Чтобы уменьшить разницу между очевидно различающимися группами пациентов, включаемых в исследование, можно применить *стратификацию*, т.е. рандомизацию с учетом слоев (страт), на которые разделяется целевая совокупность по каким-либо признакам. Например, если на результат исследования может повлиять один из заданных параметров (возраст, уровень артериального давления, перенесенный инфаркт миокарда и др.) пациентов вначале делят на подгруппы.

Рандомизированные исследования являются универсальной технологией и могут применяться для проверки эффективности (безопасности) любых форм медицинских вмешательств. Специально проведенные исследования показали, что отсутствие рандомизации или ее неверное проведение приводят к переоценке эффекта до 150 %, либо к его недооценке на 90 % [5].

Обработка данных исследования. Еще на этапе планирования исследования составляется *план статистического анализа* – документ, который содержит подробное описание принципиальных подходов к анализу результатов и детальное описание методов выполнения статистического анализа первичных (основных) и вторичных (дополнительных) переменных, а также других данных.

Этап статистической обработки данных с появлением мощных статистических программных пакетов значительно упрощается. Однако на этом этапе возможны ошибки в первую очередь вследствие неправильно выбранного метода анализа. Довольно часто исследователи ограничиваются использованием *t*-критерия Стьюдента, подсчетом коэффициентов корреляции, расчетом частотных таблиц, упуская из виду существенные ограничения применения этих методов. Существует и другая крайность – применение сложных методов многомерной статистики, интерпретация результатов которых без специальной подготовки может привести к ошибочным выводам.

Для корректной статистической обработки данных клинических исследований целесообразно использовать многофакторный дисперсионный анализ или анализ вариаций (*ANOVA*), регрессионный анализ, анализ выживаемости, факторный анализ и др. При этом определяется *главная переменная*, по которой в конечной точке делается основной вывод клинических исследований об эффективности и/или безопасности сравниваемых вмешательств в зависимости от выдвинутой гипотезы, а также *вариабельность* значений главной переменной, характеризующая отклонение индивидуальных значений от среднего в группе испытуемых.

Таким образом, уже на этапе разработки плана клинического исследования необходимо определить объем выборки с учетом вариабельности изучаемых переменных и особенностей дизайна исследования, выбрать адекватный метод измерений (оценки) параметров, а также корректный метод статистического анализа.

Одной из целей клинического исследования является *возможность перенесения результатов, полученных на выборке, на всю генеральную совокупность*. Планируя распространить результаты исследований на всю генеральную совокупность, следует учитывать вероятность ошибок 1-го и 2-го рода.

Ошибка 1-го рода (уровень значимости α) – это вероятность, с которой утверждают о наличии и достоверности различий исследуемых показателей в основной и контрольной

группах. На практике уровень значимости часто берут равным 5 %. Для суждения об отсутствии достоверных различий показателей в группах необходим расчет *ошибки 2-го рода* (β), т. е. вероятности того, что на самом деле отличия показателей в группах существенны, но ошибочно принимаются как несущественные. Вероятность этой ошибки в медицинских исследованиях должна быть меньше 20 % [6]. Достичь такого значения β удается только при значительном количестве испытуемых, поэтому не найденные возможные различия, как правило, объясняются просто недостаточным количеством наблюдений.

Весомое значение при клинических исследованиях имеет *статистическая мощность исследования* — вероятность того, что в исследовании будет найдено статистически достоверное различие, если это различие действительно существует. Мощность исследования определяется выражением $1 - \beta$ и зависит от разброса данных и объема выборки. При этом требуемый размер выборки — это минимально достаточное количество испытуемых для получения клинически и статистически значимых результатов клинических испытаний.

Статистически значимо то, что действительно происходит с высокой вероятностью. *Клинически* значимо то, что своими показателями (например, снижением смертности) убеждает врача в необходимости изменить свою практику в пользу нового образа действий. В клинических испытаниях, проводимых на большой выборке испытуемых, можно статистически достоверно выявить даже небольшие эффекты от применения вмешательства. Таким образом, клинически значимый эффект — это минимальное клинически значимое различие между группами испытуемых в конечной точке по главной переменной.

Анализ полученных результатов и принятие решения. При анализе полученных результатов исследований указывают степень убедительности доказательств, используя *унифицированную шкалу оценки доказательств*, выдвигая последовательно гипотезы:

- *доказательства убедительны*: есть веские доказательства в пользу применения данного вмешательства;
- *относительная убедительность доказательств*: есть достаточно доказательств в пользу того, чтобы рекомендовать данное вмешательство;
- *достаточных доказательств нет*: имеющихся доказательств недостаточно для вынесения рекомендации, но рекомендации могут быть даны с учетом иных обстоятельств;
- *достаточно отрицательных доказательств*: имеется достаточно доказательств, чтобы рекомендовать отказаться от применения данного вмешательства в определенной ситуации;
- *веские отрицательные доказательства*: имеются достаточно убедительные доказательства того, чтобы исключить данное вмешательство из рекомендаций [4].

Для принятия соответствующего решения необходимо последовательно выдвинуть и проверить каждую из гипотез, оценивая при этом вид альтернативной гипотезы (односторонняя/двусторонняя), уровень значимости, мощность и дизайн статистического исследования, а также клинически значимый эффект.

На эти же вопросы с еще большей достоверностью отвечает и *систематизированный обзор*, представляющий собой серьезные научные исследования — процедуры статистического обобщения многих подобных рандомизированных контролируемых испытаний. По мере появления новых данных, обзор систематически обновляется. На основе систематизированных обзоров обычно проводится *мета-анализ* — обобщение и статистический анализ результатов отдельных клинических исследований. Обобщение качественно выполненных рандомизированных контролируемых испытаний в виде систематических обзоров занимает первое место в иерархии типов исследований по степени их научной доказательности [7].

Выводы. Неадекватное планирование, методически несовершенные исследования и статистические некорректные действия с анализируемыми данными искажают полученные

результаты, что делает их непригодными для практического здравоохранения и дальнейших научных изысканий, а также приводят к неправильным выводам. При оценке результатов вмешательства следует в соответствии с правилами *ICH-GCP* четко *формализовать задачу* исследования, выбрать *дизайн* исследования и *методы статистического анализа* соответствующими поставленным задачам, а не обосновывать выводы субъективными критериями. Если эти правила соблюдены, любой вид исследования можно назвать качественным независимо от того, будут ли результаты подтверждать или опровергать выдвинутую гипотезу.

Таким образом, результаты правильно проведенных клинических испытаний с использованием метода кефалозефалографии могут дать информацию о том, медицинское вмешательство действует (вредно или полезно), не действует (бесполезно) или о его влиянии ничего не известно, *но нельзя получить ответ на вопросы – каким образом и почему действует*. На эти вопросы могут ответить только *фундаментальные исследования происходящих в организме* процессов. Фундаментальные исследования, в свою очередь, не могут обойтись без *процедуры проверки эффекта* медицинского вмешательства, откуда следует, что основная задача при проведении клинических испытаний заключается в формулировании корректной гипотезы и выработке квалифицированных решений.

Список литературы

1. Интернет-ресурс <http://www.mediasphera.ru>.
2. Флетчер Р., Флетчер С., Вагнер Э. Клиническая эпидемиология. Основы доказательной медицины: пер. с англ. – М.: Медиа Сфера, 1998. – 352 с.
3. ГОСТ Р 52379:2005. Надлежащая клиническая практика. – Введен 01.04.2006.
4. Власов В. В. Эпидемиология. – М.: ГЭОТАР-МЕД, 2004. – 464 с.
5. Интернет-ресурс <http://www.medinfo.ru>.
6. Гланц С. Медико-биологическая статистика. – М.: Практика, 1999. – 460 с.
7. Власов В. В., Семернин Е. Н., Мирошенко П. В. Доказательная медицина и принципы методологии / Мир медицины. – 2001. – №11. – С. 22 – 24.

В. Д. Кузовик, Л. О. Кошева

Деякі методологічні аспекти оцінювання ефективності медичних втручань

Показано значущість дотримання вимог міжнародних стандартів під час проведення клінічних випробувань для забезпечення достовірності їх результатів, що сприяє взаємному визнанню даних клінічних випробувань, а з розвитком інтеграційних процесів стає актуальним при розробленні та впровадженні стандартів медицини у практику.

V. D. Kuzovik, L. O. Kosheva

Some methodological aspects of estimation effectiveness of medical interpererention

Shown the importance of compliance with international standards in the conduct of clinical trials to ensure the accuracy of their results, leading to mutual recognition of these clinical trials, and the development of integration processes, it becomes important in the development and implementation of standards in medical practice.