

К.П. Воробьев

Государственное учреждение «Луганский государственный медицинский университет»

Формальные критерии качества и ценности клинической информации

В работе выполнен анализ проблемы формализации качества и ценности клинической информации для последующего использования этих характеристик в клинических рекомендациях. Качественная клиническая информация характеризуется низким уровнем систематических ошибок исследования, а ее ценность определяется тем, в какой степени изучаемое воздействие влияет на критически важный для пациента клинический исход при минимальном уровне побочных эффектов и наиболее низкой стоимости. Низкий уровень систематических ошибок рандомизированного клинического испытания обеспечивается качественным отбором групп сравнения, использованием в качестве конечных точек прямых клинических исходов, маскирование вмешательств для участников исследования, применением адекватных методов статистического анализа. Оценка качества клинического исследования требует специальных знаний в области клинической эпидемиологии, поэтому были разработаны несколько десятков систем шкальных оценок доказательности и силы клинической информации. В последние 10 лет международная группа экспертов активно развивает и внедряет новую систему оценки доказательности и силы клинической информации GRADE. В настоящее время эта система принята ведущими международными и национальными медицинскими организациями как наиболее оптимальный принцип градации клинической информации для использования ее оценок в клинических рекомендациях. В работе приведена краткая характеристика системы GRADE и обоснована необходимость ее использования в текущей научной и клинической практике.

Ключевые слова: качество клинической информации, формализация, шкала доказательности.

Введение

В последнее десятилетие в украинском клиническом сообществе все реже возникает дискуссия о целесообразности использования современных принципов доказательной медицины (ДМ) в принятии клинического решения. В настоящее время все национальные ассоциации специалистов и органы здравоохранения Украины признают приоритет научно обоснованной клинической информации при создании согласованных клинических рекомендаций. Казалось бы «всеобщая и окончательная победа» принципов ДМ позволит быстро сформировать качественные национальные клинические рекомендации. Но в реальной медицинской практике проблема выбора той или иной лечебной и диагностической технологии только обостряется, поскольку происходит постоянный рост количества противоречивой клинической информации. Это связано с тем, что медицина во всех экономически развитых странах является полем наиболее прибыльного бизнеса. Поэтому производители медицинских средств прилагают огромные усилия не только в отношении клинических исследований своих продуктов, но и в их продвижении на медицинский рынок. Приоритет прибыли в медицинском бизнесе и несовершенство средств общественного и государственного контроля за новыми медицинскими продуктами являются причинами глубокого конфликта во всем мире. Об этом достаточно часто и много говорится в последние годы. Наиболее актуальная информация на эту тему представлена в пятом издании книги «Как разрушается медицина, и как мы можем это исправить» известного британского

врача, эпидемиолога и общественного деятеля Ben Goldacre (2013; www.bad-science.net). В первую очередь автор показывает, как глобальный фармакологический бизнес с общим бюджетом в 600 млрд дол. США разрушает этические принципы медицинской науки и практики. Причем основным полем этой разрушительной деятельности является клиническая наука. Под флагом ДМ производители медицинских средств предлагают медицинскому сообществу и пациентам клиническую информацию, которая получена в результате недобросовестной практики научных исследований.

В сложившейся ситуации для формирования независимого и объективного суждения об эффективности медицинских средств врачу необходимы глубокие знания методологии современного клинического исследования. Современная методология клинического исследования и институт журнальной клинической публикации находятся в постоянном и динамическом развитии. Это требует от экспертов, которые осуществляют оценку медицинских технологий, глубоких специальных знаний, которые недоступны обычному врачу даже в странах с устоявшимися традициями преподавания основ ДМ в университетах. Поэтому возникает потребность в формализации клинической информации в виде достаточно простых и общепонятных для врача и пациента категорий. Для этих целей различные национальные и международные врачебные ассоциации разработали собственные шкалы «уровней доказательности» клинической информации. При этом часто возникает путаница при оценке клинической

информации врачами на конференциях. В этой связи в украинском клиническом сообществе обозначились две проблемы: авторы докладов и публикаций достаточно редко объявляют принятую в работе классификацию уровней доказательности, а врачи как потребители клинической информации почти никогда не интересуются применяемыми в работе классификациями уровней доказательности. Чаще всего любое упоминание об уровне доказательности оказывает некое магическое действие на врачей, что заставляет поверить в абсолютную полезность лекарственного средства без оценки формальных классификационных характеристик.

Конечно, знание и понимание формальных классификационных категорий уровней доказательности не заменит знания клинической эпидемиологии (КЭ), которая необходима для формирования собственных суждений о качестве клинической информации. Но **понимание формальных критериев уровней доказательности клинической информации является наиболее качественным суррогатом знаний КЭ**, которые необходимы для обоснованного принятия клинического решения.

Цель: представить анализ категорий качества и ценности клинической информации, предоставляемой в научно-медицинских публикациях, и дать общую характеристику наиболее распространенной шкалы уровней доказательности клинической информации.

Источники информации

Формальная оценка клинической информации является логическим продол-

жением идеологии ДМ, которая предназначена для продвижения лучшей научно обоснованной информации в текущую клиническую деятельность. Соответственно наиболее содержательная информация по изучаемой теме исходит из стран с развитыми традициями изучения КЭ как методологической основы ДМ. Среди русскоязычных публикаций следует отметить обстоятельную обзорную работу группы российских эпидемиологов (Андреева Н.С. и соавт., 2012). Эта работа также исследована не только качественным анализом теории и существующих систем формальной оценки клинической информации, но и достаточно полным списком литературы для самостоятельного изучения проблемы.

Наиболее ранняя и наиболее полная работа по анализу методологии формальной оценки клинической информации выполнена и представлена на английском языке в США Национальным агентством по здравоохранению и качеству исследований (Agency for Healthcare Research and Quality — AHRQ; www.ahrq.gov) в 2002 г. в виде содержательных рекомендаций «Системы оценки градации и силы научных данных» (West S. et al., 2002). Этот документ создан с учетом лучших традиций, содержит подробный анализ состояния вопроса, представляет конкретные рекомендации по анализу научной информации, полученной в клинических исследованиях с различными дизайнами и для решения разных задач. Особая ценность этого документа для нашего исследования состоит в систематизированном анализе существовавших в 2001 г. формальных шкал оценки уровней доказательности и силы рекомендаций: в специальном приложении рекомендаций AHRQ (5B) на 28 страницах представлена сравнительная оценка известной на тот момент 41-й системы формальной оценки клинической информации.

В данной работе основное внимание будет уделено системе классификации и оценки качества рекомендаций, которая обозначается как шкала GRADE (Grading of Recommendations Assessment, Development and Evaluation; www.gradeworkinggroup.org). На протяжении последних лет эта система получает все большее распространение благодаря международному сотрудничеству ведущих экспертов в области ДМ, которые еще в 2000 г. поставили цель создать унифицированную шкалу оценки качества клинической информации. В связи с этим рабочая группа инициативы GRADE опубликовала множество работ по разъяснению основных положений новой системы оценки клинической информации, на которые представлены ссылки на соответствующей странице сайта (www.gradeworkinggroup.org/publications/index.htm). Отдельно рабочая группа GRADE выделяет две серии публикаций, в которых представлено систематизированное изложение основных положений новой системы оценки качества клинической информации (www.gradeworkinggroup.org/publications/JCE_series.htm). Прежде всего, к этим работам отно-

сятся публикации в 336 и 337 томах Британского медицинского журнала (British Medical Journal — BMJ) (2008), но еще более полная информация представлена в 15 публикациях (из 20 запланированных) на протяжении 2011–2013 гг. в Журнале КЭ (Journal of clinical epidemiology — JCE) (Guyatt G.H. et al., 2011g). Рабочая группа GRADE позаботилась о создании простой и информативной формы разъяснения основных положений новой системы оценки качества клинической информации в виде вопросов и ответов с расширенными толкованиями и гиперссылками на различные примеры в специальном разделе своего сайта (www.gradeworkinggroup.org/FAQ/index.htm).

Кроме всей этой информации, также существует ряд отдельных публикаций на английском языке, в которых представляется критическая оценка, популяризация и конкретные примеры применения системы GRADE. Эти и другие работы достойны самого пристального внимания со стороны отечественных экспертов, вовлеченных в оценку медицинских технологий и разработку национальных клинических рекомендаций. В связи с этим обстоятельством обращаем внимание читателя, что большая часть современной периодики, на основе которой подготовлена эта работа, находится в свободном доступе в сети интернет, что отражено в списке литературы данной публикации.

Основные характеристики клинической информации

Под клинической информацией подразумеваются первичные или обобщенные результаты медицинских исследований на людях. В клинических исследованиях ставятся задачи оценки факторов риска, изучаются диагностические тесты и эффективность отдельных средств лечения и профилактики. В наименовании работы нами выделены две главные категории клинической информации: качество и ценность. Для разработки формальных классификационных категорий качества и ценности клинической информации требуется вначале определить **ожидания потребителей клинической информации**. Понятие «ожидания» является ключевой категорией, которая позволит нам лучше понять, почему в современной клинической деятельности различные ее участники по-разному оценивают одну и ту же клиническую информацию.

В обыденной жизни у каждого человека есть собственные критерии качества по отношению к окружающим вещам и явлениям, то есть различные ожидания, которые могут в той или иной мере соответствовать общепринятым критериям качества. Чем сложнее объект, и чем меньше знаний у оценщика об основных свойствах объекта — тем труднее определить его качество. Поэтому у различных оценщиков, в зависимости от уровня полноты знаний о фундаментальных характеристиках объекта, складываются разные ожидания и требования к характеристикам объекта оценки и, соответственно, формируются

различные суждения о качестве объекта. Полнота знаний о клинической информации определяется знанием методологии современной клинической науки, которая развивается настолько быстро, что даже те, кто создает новые клинические знания, не всегда владеют всеми вопросами объективной оценки ее качества.

Таким образом, для объективной оценки качества и ценности клинической информации недостаточно бытовых представлений для формирования полноценных суждений об этих категориях. Необходимо достаточно хорошо знать теоретические основы клинических исследований. В современной медицине умение вынести объективные суждения о клинической информации является одной из важнейших компетенций врача, которую он обязан развивать на протяжении всей своей клинической деятельности. Эти компетенции черпаются из теории новой фундаментальной клинической науки — КЭ, предмет которой совершенно неизвестен врачам в странах постсоветского пространства. С другой стороны, только в последние годы отечественные исследователи были вынуждены начать освоение принципов КЭ, без знания которых невозможно не только проведение эффективных клинических исследований, но самое главное — невозможно подготовить качественный клинический отчет в виде журнальной публикации в рейтинговом медицинском журнале. Низкий уровень знаний в области КЭ формирует в нашем клиническом сообществе существенные упрощения и искажения требований к качеству новой клинической информации.

В современной клинической науке **качество клинической информации** определяется многими факторами, прежде всего внутренней валидностью исследования, которая зависит от того, в какой мере полученные результаты справедливы в отношении данной выборки (Флетчер Р. и соавт., 1998; с. 24). Важнейшим фактором качества клинического исследования является вероятность систематических ошибок в связи неправильным планированием, проведением и анализом результатов исследования. Добротная клиническая информация предполагает наиболее точный ответ на поставленный клинический вопрос, поэтому качество клинической информации повышается, если исследования выполнены на достаточном количестве наблюдений, имеют однонаправленный вектор изучаемого воздействия, а результаты укладываются в узкие доверительные интервалы и повторяются в независимых исследованиях. **Ценность клинической информации** определяется тем, в какой степени изучаемое воздействие влияет на критически важный для пациента клинический исход при минимальном уровне побочных эффектов и наиболее низкой стоимости.

Главным объектом оценки качества клинической информации является публикация в рецензируемом научно-медицинском журнале. В ограниченном объеме журнальной работы невозможно раскрыть

все стороны клинического исследования, которые позволят вынести окончательные и полные суждения о качестве новой клинической информации, но современный международный институт журнальной клинической публикации основывается на развитых правилах, которые позволяют обеспечить целостность клинического исследования. Под целостностью подразумеваются определенные гарантии рецензируемого медицинского журнала в том, что авторы использовали верные подходы для организации и анализа результатов клинического исследования (методология), а результат исследования базируется на реальных данных.

Если методологическое качество научного исследования можно оценить по формальным критериям журнальной публикации, то для оценки достоверности первичной информации и соблюдения заявленного протокола исследования требуется выполнение определенных правил в процессе подачи публикации и рецензирования статьи, вплоть до проведения специального расследования при подозрении в недобросовестной практике научного исследования. Потребитель клинической информации имеет ограниченные возможности в оценке целостности клинического исследования, но достаточно жесткий стандарт формата современной публикации по результатам клинического исследования позволяет читателю вынести целый ряд собственных независимых суждений о качестве клинической информации.

Для определения порога качества клинической информации и ее пригодности для решения конкретных клинических задач существуют общепринятые принципы, которые чаще всего обозначают как принцип «мусорной корзины». В книге Т. Гринхалх «Основы доказательной медицины»* глава 3.1 называется «Ученый и мусорная корзина». В начале этой главы автор утверждает, что «примерно 99% научных медицинских публикаций годятся не для практической деятельности, а для мусорной корзины» (Гринхалх Т., 2004; с. 53). Далее автор определяет ряд правил, позволяющих оценить качество клинической информации и ее полезность для медицинской практики. Аналогичным образом ведущий российский эксперт в области ДМ В.В. Власов излагает алгоритм отбора клинической публикации для использования в практике или для мусорной корзины (Власов В.В., 2001). По своей сути эти алгоритмы представляют собой основу оценки качества клинической информации. С учетом этих принципов проведем анализ факторов качества клинической информации, представляемой в журнальных публикациях.

Очень простым, но существенным ограничительным критерием качества клинической информации является ясно и однозначно выраженные клиническая

проблема и клиническая гипотеза в рамках этой проблемы. Формулировке гипотезы и цели исследования во введении должны предшествовать описание уровня предыдущих знаний и клиническое значение изучаемой проблемы. В отечественной периодике частыми являются публикации, в которых описываются некоторые клинические случаи без четко обозначенной цели исследования и без формулировки клинических вопросов, на решение которых направлено исследование. Объективная оценка уровня предыдущих знаний и формулировка изучаемой гипотезы определяют выбор типовой структуры клинического исследования — важнейшего параметра, от которого зависит оценка качества новой клинической информации. Упрощенный алгоритм предполагает «взвешивание доказательств» (Флетчер Р. и соавт., 1998; с. 302) в ряду различных типовых дизайнов исследований от рандомизированного контролируемого испытания (РКИ) до изучения серий случаев.

Безусловно, РКИ общепризнан наиболее точным и эффективным методом проверки изучаемой клинической гипотезы, но целый ряд факторов могут ограничивать возможность проведения исследований с такой структурой для решения определенных клинических вопросов. С другой стороны, клинические испытания требуют тщательного соблюдения целого ряда правил, нарушение которых является поводом для отрицательной оценки качества полученной клинической информации. Все тщательно продуманные правила типового РКИ разработаны с единственной целью — **снизить уровень систематической ошибки исследования**. Вместе с тем следует учитывать, что крупные РКИ чаще всего проводятся за счет средств производителей, заинтересованных в получении смещенного в их пользу результата клинического исследования. Соответственно, в таких случаях вероятность нарушения протоколов РКИ и смещенной оценки результатов исследований многократно возрастает. Для предотвращения влияния медицинских компаний на результат клинических исследований предусмотрен целый ряд мер на уровне международного научного сообщества и правительственных организаций экономически развитых стран. Эти меры предполагают предварительную регистрацию протоколов РКИ, независимый мониторинг выполнения протокола РКИ, попытки обязать публиковать все (в том числе и отрицательные) результаты корпоративных клинических исследований, обеспечение доступа к первичным данным исследований, широкое раскрытие конфликтов интересов всех участников клинических исследований. Несмотря на все эти меры, существует ряд методов, которые используются заинтересованными производителями для получения смещенной оценки результатов исследований. Поэтому при решении некоторых клинических вопросов информация, полученная в результате масштабных многоцентровых независимых наблюдательных клинических исследований текущей

клинической практики, может иметь более высокую ценность, чем результаты РКИ на ограниченном количестве клинических случаев с высокой вероятностью предвзятости при проведении и оценке результатов клинического исследования. Эти закономерности явно учитываются в современных развитых системах оценки качества и ценности клинической информации.

Все же для решения клинически важных вопросов следует полагаться на результаты качественных РКИ. Представим основные критерии, которые будут определять качество РКИ и соответствующей клинической информации.

Четкое и однозначное **определение клинических исходов**, под которыми подразумеваются исходы, имеющие наибольшее значение для больных (критически значимые для пациентов исходы), такие как вероятность смерти, наличия заболевания, симптоматики, которая снижает качество жизни пациента, сохранения способности к привычной деятельности, ограничений некоторых функций. Смещение целей РКИ к оценке суррогатных исходов, таких как данные лабораторных тестов и других косвенных признаков заболеваний, снижает ценность и качество клинической информации (Guyatt G.H. et al., 2011e).

Использование **стандартного метода лечения в контрольной группе**. Это обязательное требование методологии сравнительных исследований также является этическим правилом РКИ. Также следует помнить правило исходного признания нулевой гипотезы, в соответствии с которым стандартный и изучаемый метод имеют одинаковую эффективность и безопасность. В реальности это правило выполняется редко, но оно должно быть четко обозначено в протоколе испытания и обязательно получить соответствующую оценку этического комитета. В противном случае повышается вероятность смещенного результата исследования в пользу заинтересованных лиц.

Без однозначного и точного алгоритма **критериев и методов отбора клинических групп** возникает высокая вероятность систематической ошибки исследования. Для РКИ важнейшим условием снижения систематической ошибки является случайное (рандомизированное) формирование групп исследования с использованием надежного алгоритма. Не четко определенные критерии отбора пациентов в группы исследования являются основанием для отказа от использования соответствующей клинической информации. При формальной оценке клинической информации результаты, полученные в нерандомизированных проспективных клинических исследованиях, обладают более низким уровнем качества.

Однозначное **определение когорт исследования** является необходимым условием для оценки внешней валидности исследования, то есть для оценки приемлемости клинической информации по отношению к конкретному клиническому случаю. Следует учитывать, что для сниже-

*В оригинале эта наиболее известная в мире и популярная среди врачей книга о ДМ называется «Как читать научные статьи». Книга посвящена, в первую очередь, обучению врача поиску и оценке качества клинической информации.

ния уровня систематической ошибки изучаемые группы по целому ряду параметров являются подобными и узкоселективными (пол, возраст, форма, стадия и тяжесть заболевания, социальный достаток, уровень образования и др.). Поэтому ценность клинической информации может, при прочих равных условиях, существенно снижаться, если известно, что по некоторым признакам, от которых зависит исход и течение заболевания, группы клинического исследования и конкретный клинический случай в текущей практике существенно различаются.

Обеспечение так называемого **слепого метода** (син.: маскирование) позволяет исключить психосоматическое влияние на состояние пациента основного и стандартного методов лечения, снижает возможности исследователей влиять на результат исследования и является важным приемом для снижения систематической ошибки исследования.

Мера клинического эффекта является важнейшей и обязательной характеристикой клинической информации. Поэтому адекватное использование **методов статистического анализа** первичной информации является неотъемлемой и важнейшей характеристикой качества и ценности клинического исследования. В первую очередь, исследователь обязан обобщать **размеры выборки опытной и контрольной групп**, которые определяются такими факторами, как: распространенность заболевания, сила изучаемого воздействия, дисперсия оцениваемых в исследовании параметров и др. Для снижения уровня случайной ошибки исследования также должны быть адекватно подобраны и точно описаны методы статистического анализа, поскольку существует целый ряд ограничений и правил для использования тех или иных методов статистического анализа. Также следует учитывать формат представления изучаемых характеристик, который должен позволить читателю оценить как уровень статистической значимости соответствующих тестов, так и величину клинического эффекта. Эти правила требуют специальных глубоких знаний, а их соблюдение должно оцениваться подготовленными специалистами. Использование неадекватных статистических методов и неявное представление выборочных характеристик исследования в сравниваемых группах является основанием для отказа от использования клинической информации, представленной в публикации.

Качественная клиническая информация также должна содержать **эпидемиологические цифровые характеристики** полученных в исследовании клинических эффектов. Эти характеристики специфичны для различных типов исследований (диагностики, оценки прогноза, эффективности воздействий) и выражаются такими параметрами, как чувствительность и специфичность диагностического теста, отношения рисков в сравниваемых группах воздействия, шансы возникновения изучаемых исходов и другие обобщающие ре-

зультаты исследования, которые направлены на оценку величины клинического эффекта и его дисперсии. Современным стандартом публикационной практики предусматривается представление этих параметров с указанием **диапазона оговоренного доверительного интервала**, значения которого позволяют судить как о статистической значимости, так и о клинической ценности результата исследования (Guyatt G.H. et al., 2011c).

Наиболее высокий уровень качества клинической информации обычно содержат **результаты метаанализа клинических исследований**, который должен быть выполнен по специальному алгоритму для каждого типа исследований. Именно такие обобщающие исследования позволяют получить наиболее качественную клиническую информацию.

Даже столь упрощенное представление основных вопросов методологии клинического исследования демонстрирует сложность объекта под названием «клиническое исследование». Учет различных факторов качества и ценности клинического исследования требует не просто специальных знаний, но еще и достаточно трудоемкого анализа первичных результатов для формирования собственных суждений. С учетом этого вполне понятно стремление научной медицинской общественности к формализации критериев качества и ценности клинической информации для обобщенных и сравнительных оценок различных вмешательств с последующим выбором альтернатив для использования в клинической практике.

История развития систем оценки качества клинической информации

Первый опыт разработки формальных критериев качества клинической информации принадлежит Канадской рабочей группе по профилактической медицине (Canadian Task Force on the Periodic Health Examination; www.answers.com/topic/canadian-task-force-on-preventive-health-care). Это была правительственная организация в лице 10 заместителей министров здравоохранения соответствующих провинций Канады. В группу также были включены эпидемиологи, исследователи в области здравоохранения и другие специалисты.

В 1976–1978 гг. целевая группа разработала методологию для оценки научных данных за и против эффективности любого профилактического вмешательства. Новизна методологии оценки рекомендаций, которую разработала Канадская целевая группа, заключалась в том, что **научные данные ставились выше консенсусных решений**. Эта методология стала не только международным стандартом для оценки профилактических мероприятий, но и явилась основой для последующих систем оценки качества клинической информации. Кратко рассмотрим градацию доказательств, которую предложила рабочая группа в 1978 г. и которая заложила основные принципы современных методов

формальной оценки качества и ценности клинической информации. Для этой системы характерна пятиуровневая градация доказательности:

1. Доказательства, полученные, по крайней мере, в одном и более должным образом спланированных РКИ.

2. Доказательства, полученные в хорошо спланированных контролируемых испытаниях без рандомизации.

3. Доказательства, полученные в хорошо спланированных когортных исследованиях или аналитических исследованиях случай — контроль, предпочтительно из более чем одного центра или от нескольких исследовательских групп.

4. Доказательства, полученные в исследованиях без вмешательства с историческим контролем или при сравнении результатов, полученных из разных мест.

5. Мнения уважаемых авторитетов, основанные на клиническом опыте, описательных исследованиях или на докладах экспертных групп.

В сочетании с этими принципами доказательности научных данных, целевая группа разработала трехуровневую классификацию силы рекомендаций:

1. Хорошие (good) доказательства для того, чтобы принять рекомендацию о регулярных медицинских осмотрах.

2. Надежные (fair) доказательства для принятия рекомендации, что определенные условия (condition) будут учтены в регулярных медицинских осмотрах.

3. Недостаточные (poor) доказательства в отношении включения или исключения рекомендации при периодических медицинских осмотрах, что может быть заключено на основе следующих данных:

- надежные доказательства для того, чтобы не рекомендовать использовать вмешательство в регулярных медицинских осмотрах;
- значительные доказательства, чтобы рекомендовать исключить вмешательство из регулярных медицинских осмотров.

Дальнейшая история возникновения, развития и текущего использования различных шкал формальной оценки клинической информации достаточно подробно описана в аналитической публикации российских эпидемиологов (Андреева Н.С. и соавт., 2012). Среди англоязычных работ наиболее полная информация о существующих в 2001 г. системах оценки качества клинической информации изложена в рекомендациях AHRQ. В этих рекомендациях представлен систематизированный сравнительный анализ 41 шкалы оценки качества клинической информации (West S. et al., 2002). Очевидно, именно это разнообразие, путаница и несовершенство существующих систем оценки качества клинической рекомендации стало основанием возникновения рабочей группы GRADE. Эта инициативная группа возникла в 2000 г. как неформальное сотрудничество специалистов, которые **поставили цель создать единую классификацию оценки качества научных данных, используемых в клинических рекоменда-**

циях. Уже в 2004 г. рабочая группа GRADE впервые представила свои предложения по использованию новой системы оценки качества клинической информации (Atkins D. et al., 2004).

Рабочая группа представляет интересы 32 стран, но почти половину (78 из 182) рабочей группы составили специалисты США, Великобритании и Канады (www.gradeworkinggroup.org/about_us.htm). Последовательная и независимая деятельность этой группы позволила к настоящему времени сделать систему GRADE стандартом представления качества научной информации в клинических рекомендациях (Andrews J.C. et al., 2013b).

Таким образом, еще в 1978 г. был заложен определенный принцип оценки качества клинической информации, который развивался в течение последующих лет. Также важно заметить, что за прошедшие 35 лет в странах англосаксонской традиции сформировалось понимание необходимости подобных шкальных оценок клинической информации. Это достаточно большой промежуток времени, в течение которого изменялись взгляды медицинского сообщества соответствующих стран на ценность клинической информации. Кроме того, параллельно в этих странах велась систематизированная до- и последипломная подготовка врачей в области КЭ. Можно предположить, что фактор времени не имеет существенного значения для внедрения формальных критериев качества клинической информации в национальных украинских клинических рекомендациях, поскольку будем использовать уже проверенную и общепризнанную методологию. Для успешной имплементации новой методологии разработки и использования клинических рекомендаций более важен **фактор достаточной подготовки врача и разработчиков рекомендаций в области КЭ**. Без такой подготовки внедрение стандартов оценки качества клинической информации в Украине столкнется с противодействием и непониманием. Как бы мы не поднимали самооценку своих знаний в области современной эпидемиологии, следует учесть, что ни в России, ни в Украине, ни в других странах постсоветского пространства нет современных школ КЭ и критического числа квалифицированных экспертов, которые могли бы создавать рабочие группы при решении задач медицинской науки и здравоохранения. Об этом свидетельствует тот факт, что из 182 членов рабочей группы GRADE нет ни одного специалиста из стран постсоветского пространства. Данное обстоятельство также свидетельствует о глубоком отставании, в частности Украины, от процессов, связанных с внедрением современных технологий ДМ.

Основные положения системы оценки качества клинических рекомендаций GRADE

При разработке новой системы оценки качества клинической информации рабо-

чая группа GRADE постаралась минимизировать следующие недостатки предшествующих аналогов:

- неявные определения уровня доказательности и силы рекомендаций;
- неявные суждения о том: какие результаты клинических исследований имеют важное значение, каково качество доказательств для каждого изучаемого клинического исхода, каково общее качество доказательств, каков баланс между пользой и вредом, каковы затраты и дополнительные преимущества рекомендации;
- в качестве основных суждений о качестве доказательств часто используется только дизайн исследований и не учитываются важные клинические исходы или другие факторы, такие как дозозависимый эффект, публикационная систематическая ошибка, степень повторяемости основного результата или влияние плацебо на результат исследования;
- общий вывод о доказательности клинической рекомендации основывается на фактах положительного влияния изучаемого средства без учета низкоуровневых дизайнов исследований, в которых могут быть представлены убедительные альтернативные данные в отношении значимых клинических исходов;
- часто не учитывается баланс польза/вред и возможные особые условия использования рассматриваемого вмешательства;
- часто не рассматриваются прямо соотношение между преимуществами рекомендаций для здоровья и стоимостью их использования;
- процесс обобщения доказательств не имеет единого алгоритма и прозрачности;
- рекомендации часто используются локально из-за трудностей в понимании различных систем градации качества клинических рекомендаций, что снижает эффективность работы экспертов и не позволяет расширить международное сотрудничество в различных отраслях медицины.

Исходя из этих недостатков существующих систем формальной оценки качества и ценности клинической информации рабочая группа GRADE начала систематизированную работу по созданию новой универсальной классификации.

Основными категориями шкалы оценки качества клинической информации в системе GRADE являются **уровень достоверности клинической информации** (levels of evidence) и **уровень убедительности рекомендаций** (grades of recommendation), который определяется как сила рекомендаций (strength of recommendation) (Atkins D. et al., 2004). **Сила рекомендаций** — степень превышения выгоды от преимуществ метода над рисками и стоимостью. Рекомендации считаются **сильными** (strong) в случае полной уверенности в том, что ожидаемая польза от применения вмешательства

превосходит его нежелательные последствия, и **слабыми** (weak), если преимущества, риски и стоимость тонко сбалансированы или в отношении этого баланса существуют неопределенности (Guyatt G.H. et al., 2008a). С учетом возрастающей роли учета предпочтений пациентов в принятии клинических решений врач обязан, при условии полного информирования о сильной рекомендации, предложить альтернативные — слабые рекомендации (www.gradeworkinggroup.org/FAQ/index.htm).

Двоичная классификация класса рекомендаций по силе обеспечивает четкие ориентиры для пациентов, врачей и политиков. Сильные рекомендации означают, что наиболее информированные пациенты будут выбирать рекомендуемое воздействие, а врачи обязаны будут дать содержательную информацию пациенту об этом методе. Слабые рекомендации означают, что выбор пациентов может меняться в зависимости от их предпочтений, и врачи обязаны убедиться, что помощь пациентам оказывается в соответствии с их ценностями и предпочтениями. Сила рекомендации определяется балансом между желательными и нежелательными последствиями альтернативных стратегий воздействия, качества доказательств, изменчивости ценностей и предпочтений пациентов, а также использованием разных ресурсов (Guyatt G.H. et al., 2011a).

При определении уровня достоверности клинической информации GRADE исходит, прежде всего, **из характеристик дизайна исследования**. Изначально считается, что дизайн РКИ обеспечивает наиболее высокий уровень доказательности, но в некоторых случаях рейтинг достоверности РКИ может быть снижен. Например, заявление дизайнера РКИ не обеспечивает высшего уровня достоверности при наличии явных различий между сравниваемыми группами, отсутствии маскирования вмешательства для участников клинического испытания или существовании других явных и вероятных источников возникновения систематической ошибки.

В результате использования специальных алгоритмов GRADE выделяет следующие уровни достоверности клинической информации: **высокий** (high), **средний** (moderate), **низкий** (low) и **очень низкий** (very low). Высокий уровень предполагает, что существует очень малая вероятность того, что дальнейшие исследования изменят нашу уверенность в оценке эффекта. Средний уровень предполагает, что дальнейшие исследования, вероятно, окажут значительное влияние на нашу уверенность в оценке эффекта и могут изменить оценку. Низкий уровень означает, что дальнейшие исследования скорее всего окажут важное влияние на нашу уверенность в оценке эффекта и, вероятно, изменят оценку. Очень низкий уровень означает, что любая оценка эффекта является очень неопределенной (Balsheem H. et al., 2011).

Для практического использования могут применяться две формы обозначения силы и убедительности рекомендаций: словесная и традиционная символическая. Во втором случае уровни достоверности от высокого к очень низкому обозначаются буквами «А», «В», «С» и «D» или закрашенными кружочками в таблицах рядом с рекомендациями (чем больше из четырех кружочков закрашено — тем выше доказательность рекомендаций). Сильные и слабые рекомендации могут обозначаться цифрами «1» и «2» соответственно.

Система GRADE гибко учитывает методологический уровень научных исследований при определении уровня достоверности научных данных, например РКИ с существенными недостатками относятся к категории умеренного уровня доказательств, а РКИ с несколькими серьезными недостатками — к категории низкого уровня доказательств. Важнейшим нововведением системы GRADE явились новые научно обоснованные **принципы снижения или повышения уровня доказательности** клинической информации. При определении уровня доказательности первоначально РКИ относятся к высокому уровню, наблюдательные исследования — к низкому, прочие клинические данные — к очень низкому уровню доказательности. Далее используется ряд следующих принципов, которые снижают или повышают уровень доказательности клинической информации (Guyatt G.H. et al., 2008c).

Принцип снижения уровня достоверности клинической информации учитывает следующие факторы: серьезные и очень серьезные дефекты качества исследования; несогласованность (inconsistency) результатов различных исследований (Guyatt G.H. et al., 2011d); умеренная или существенная неопределенность направления клинического эффекта; неопределенные результаты или недостаточное количество данных; высокая вероятность публикационной систематической ошибки (публикуются определенные результаты или только те, которые финансируются промышленностью) (Guyatt G.H. et al., 2011f; i). Редкие данные (sparse data) как причина снижения уровня достоверности классифицируются в двух ситуациях: 1) очень мало пациентов в исследовании для получения надежных результатов; 2) доверительные интервалы достаточно широки (включают пользу и ее отсутствие или наличие вреда). При различиях характеристик пациентов клинического исследования и тех, для которых предназначены будущие клинические рекомендации (низкая внешняя валидность), также рекомендуется снизить уровень достоверности клинической информации. Снижению уровня доказательности также способствует раннее окончание лечения при улучшении и несообщение о клинически важных исходах (Guyatt G.H. et al., 2011h).

Принцип повышения уровня достоверности клинической информации основан на учете следующих факторов: уровень градации достоверности повышается на единицу, если имеются убедительные доказательства клинического эффекта по данным не менее чем двух наблюдательных исследований, которые выражаются в высоком (>2) или низком (<0,5) относительном риске при маловероятном влиянии вмешивающихся факторов; уровень градации достоверности повышается на 2 балла, если получены очень существенные эффекты в виде очень высокого (>5) или очень низкого (<0,2) относительного риска, если данные получены на основе прямых доказательств (оценка значимых клинических исходов), без существенной вероятности влияния ошибок исследования на достоверность данных; на один балл также повышается оценка уровня достоверности, если в исследовании показан дозозависимый эффект или если имевшиеся в исследовании вмешивающиеся факторы могли снизить клинический эффект (Guyatt G.H. et al., 2008b).

Таким образом, уровень достоверности наблюдательных исследований повышается, если в различных исследованиях получены клинически значимые и согласующиеся (consistent) оценки величины воздействия. Также повышается уровень доказательности наблюдательного исследования, если эксперты определяют, что вмешивающийся фактор снизил оценку результата воздействия. Низкоуровневые наблюдательные исследования могут модифицировать наши представления о результатах РКИ. Например, градация РКИ, в котором показана эффективность длительного применения стероидов при саркоидозе, без указания данных о токсическом эффекте, была снижена по результатам наблюдательных исследований, в которых показаны токсические эффекты воздействия (Grutters J.C., van den Bosch J.M., 2006).

В целом перечисленные критерии могут изменить (понижить или повысить) достоверность данных на один или два уровня. Этот принцип существенно изменяет прежнюю парадигму и его можно оценить как **революционный в оценке достоверности клинической информации**. Например, результаты наблюдательного исследования (исследование без вмешательства) изначально рассматриваются как доказательства низкого уровня. Но если в двух таких исследованиях получены убедительные данные и показан дозозависимый эффект, то эти данные сразу переходят на высокий уровень доказательности клинической информации при условии соответствия исследования определенным требованиям. Трудно переоценить важнейшие следствия таких перемен. Если раньше, независимо от методологического качества и полученных научных фактов, все результаты наблюдательных исследований укладывались на запасные полки клинической информации и воспринимались не иначе, как научные гипотезы, то теперь появляются серьезные мотивации проведения недорогих по бюджету наблюдательных исследований. При относительно низких затратах в сравнении с дизайном РКИ результаты таких исследований могут стать убедительными аргу-

ментами в клинических рекомендациях для изменения текущей клинической практики.

В связи с этими новыми принципами и высокой вероятностью публикационной систематической ошибки (Guyatt G.H. et al., 2011f) при рассмотрении результатов исследований были существенно повышены требования к идеальному РКИ, претендующему на высокий уровень доказательности клинической информации. Недостатки в следовании протокола испытания, неадекватность маскирования и случайного формирования групп исследования, субъективная оценка клинических исходов — все это способствует систематической ошибке, что может проявляться в существенных различиях результатов разных РКИ (неоднородность и варибельность результатов).

Следующий принцип GRADE предполагает определение уровня достоверности данных для критически важных и важных клинических исходов (Guyatt G.H. et al., 2011b). Если критически важные исходы получены в исследованиях с низким уровнем достоверности, то, несмотря на важные эффекты воздействия в другом исследовании с более высоким уровнем достоверности, общий уровень достоверности рекомендаций снижается (Guyatt G.H. et al., 2013a; b).

Достаточно много и часто в системе GRADE упоминается о **ценностях и предпочтениях пациентов**, которые, например, могут повлиять на уровень оценки достоверности рекомендаций из-за выбора различных путей разных когорт критически значимых исходов. Например, людей молодого возраста существенно заботит ожидаемая продолжительность жизни, в то время как лиц пожилого возраста может больше волновать предстоящее качество жизни. Соответственно для этих двух групп критически важные клинические исходы будут различны. Чем больше ценности и предпочтения изменяются в популяции, или чем больше неопределенность в ценностях и предпочтениях, тем выше вероятность того, что слабые рекомендации будут иногда оправданными (Guyatt G.H. et al., 2008a).

Оценка стоимости также влияет на степень убедительности рекомендаций. Клинические рекомендации неизбежно связаны с оценкой распределения ресурсов, которые обычно называют затратами. В различных условиях существуют разные отношения к тому, как влияют расходы на клиническое решение врача. Например, годовой запас дорогого препарата может соответствовать зарплате медсестры в США или зарплате тридцати медсестер в Китае. Соответственно, рекомендации по применению этого препарата могут существенно различаться в США и Китае. Вопросы, связанные с использованием ресурсов, являются в высшей степени политическими и часто приводят к конфликту интересов в группах экспертов, поэтому конкретный контекст имеет решающее значение для рассмотрения использования ресурсов (Guyatt G.H. et al., 2008b). Общий принцип учета стоимости в реко-

мендациях можно сформулировать следующим образом: чем выше затраты на вмешательство, то есть чем больше ресурсов потребуется — тем меньше вероятность того, что сильная рекомендация является оправданной (Guyatt G.H. et al., 2008a; Brunetti M. et al., 2013).

Особое внимание рабочая группа GRADE уделяет результатам **исследований точности диагностических тестов**. Точность диагностического теста является суррогатной характеристикой для оценки важных клинических исходов, которые могут быть использованы в качестве стандарта диагностики. Рассмотрение достоверности данных зависит от того, насколько мы уверены в правильной классификации важных для пациента клинических исходов. Рабочая группа GRADE продолжает заниматься развитием рекомендаций, которые обеспечат дальнейшее понимание классификации качества доказательств для диагностических исследований (Tatsioni A. et al., 2005; Bossuyt P.M. et al., 2006; Leeflang M.M. et al., 2008; Schünemann H.J. et al., 2008a; b; Horvath A.R., 2009), но члены рабочей группы GRADE также предлагают вполне конкретные алгоритмы оценки уровня доказательности и силы рекомендаций диагностических тестов для использования этих градаций в создании новых клинических рекомендаций (Schünemann H.J. et al., 2008a; b; Hsu J. et al., 2011).

Считается, что принципы GRADE сложны для реализации (Guyatt G.H. et al., 2013c), поэтому для эффективного их использования, обеспечения прозрачности процесса подготовки рекомендаций и улучшения повторяемости работы разных экспертов на сайте Кокрановского сотрудничества предлагается бесплатное программное обеспечение (<http://ims.cochrane.org/revman/gradepro>), а на сайте рабочей группы GRADE — курс видеобучения по работе с этими программами (www.gradeworkinggroup.org/toolbox/index.htm).

Распространение шкалы GRADE в современной практике разработки клинических рекомендаций

Существуют ли приемлемые альтернативы шкалы GRADE для использования в клинических рекомендациях? Несмотря на общее признание системы GRADE, некоторые ассоциации врачей используют в настоящее время иные шкалы, например USPSTF (US Preventive Services Task Force and the Grading of Recommendations) в рекомендациях для заболеваний щитовидной железы у беременных применяют несколько иную шкалу доказательности. Несмотря на отличия этой шкалы, USPSTF ссылается в своих рекомендациях на уровни силы рекомендаций шкалы GRADE (De Groot L. et al., 2012). Эти рекомендации имеют градации: А — строго рекомендовано; В — менее строго, но несет пользу пациентам; С — неопределенные рекомендации из-за слабой доказательности; D — рекомен-

дации против. В тексте рекомендации указывается, что уровни «А» и «В» соответствуют сильным и слабым рекомендациям по шкале GRADE.

Для оценки эффективности шкалы GRADE также проводятся научные исследования. В сравнительном исследовании повторяемости оценок экспертов при градации качества и силы клинической информации продемонстрирован вдвое более высокий уровень повторяемости выводов экспертов, использующих систему GRADE (Mustafa R.A. et al., 2013). Результаты этого исследования свидетельствуют о перспективах широкого внедрения новой системы взамен существующих.

Процесс установления нового стандарта шкалы доказательности лучше всего проявляется в текущих согласованных клинических рекомендациях различного уровня, но прежде всего, в международных рекомендациях (Andrews J.C. et al., 2013a). Междисциплинарное сообщество Surviving Sepsis Campaign (SSC), которое объединяет более 15 крупнейших международных обществ и ассоциаций Европы, Северной Америки и Японии, уже много лет ведет работу над совершенствованием клинических рекомендаций по лечению сепсиса (www.survivingsepsis.org). В 2012 г. SSC представило обновленные рекомендации по ведению больных тяжелым сепсисом и септическим шоком. На нескольких страницах этих рекомендаций в разделе Grading of Recommendations подробно описываются принципы шкалы GRADE с деталями применительно к сепсису (Dellinger R.P. et al., 2013). Еще ранее, в 2008 г., при разработке предыдущего варианта рекомендаций по лечению сепсиса члены рабочей группы применили новую оригинальную технологию достижения консенсуса среди членов экспертной группы путем последовательных уточняющих опросов в рамках концепции системы оценки качества клинической информации (Jaeschke R. et al., 2008).

Подобным образом составлены международные рекомендации по лечению острых повреждений почек (Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group, 2012). При этом принятию этого стандарта в международном урологическом журнале предшествовала статья, которая раскрывала преимущество системы GRADE как инструмента создания клинических рекомендаций (Canfield S.E., Dahm P., 2011). Кроме международных организаций, различные национальные агентства оценки медицинских технологий признают шкалу GRADE как наиболее удобный и эффективный инструмент при создании клинических рекомендаций. Например, Шотландская межколлегияльная организация по разработке клинических рекомендаций (Scottish Intercollegiate Guidelines Network — SIGN) в 2009 г. официально признала приоритет методологии GRADE при подготовке клинических рекомендаций (Applying the GRADE methodology to SIGN guidelines: core principles; www.sign.ac.uk/pdf/gradeprincipals.pdf).

Это частные случаи. Более важно, как отреагировали на инициативу GRADE ве-

дущие игроки международного медицинского сообщества. В 2006 г. BMJ предложил использовать систему GRADE при публикации клинических рекомендаций. В 2008 г. на сайте рабочей группы GRADE в специальном разделе, где указаны партнеры, были представлены 25 крупнейших и авторитетных организаций, таких как Всемирная организация здравоохранения, Американская коллегия врачей, Американская торакальная ассоциация, всемирно известный популярный электронный журнал UpToDate (www.uptodate.com), Кокрановское сотрудничество и др. (Guyatt G.H. et al., 2008d). В конце 2013 г. в этом же разделе сайта рабочей группы представлено уже 75 различных авторитетных организаций, которые поддерживают рекомендации GRADE (www.gradeworkinggroup.org/society/index.htm). Все это свидетельствует об установлении де-факто нового стандарта оценки формальных критериев клинической информации, знание которого **становится важнейшей компетенцией современного врача**.

Место альтернативных шкал оценки клинической информации

Качество доказательств — это непрерывная категория, и любая дискретная классификация является определенной степенью упрощения. Чем больше точек определенной функции будет описано, тем более качественно и полно представляется функция для восприятия. Поэтому в организациях, которые традиционно занимаются фундаментальными вопросами ДМ (кафедры КЭ и биостатистики), для большей полноты описания качества клинической информации используются более сложные классификации и соответствующие шкалы. Например, в Оксфордском центре доказательной медицины используется собственная шкала уровней доказательности («The Oxford 2011 Levels of Evidence»; www.cebm.net/mod_product/design/files/CEBM-Levels-of-Evidence-2.1.pdf). Эта шкала выделяет пять уровней доказательности, которым соответствуют: систематизированные обзоры РКИ, РКИ, нерандомизированные контролируемые исследования, исследование серий случаев, а самый низкий уровень шкалы основывается на изучении механизмов заболеваний и мнениях экспертов. В комментариях к описанию этой шкалы упоминается шкала GRADE, которая, по мнению экспертов Оксфордского центра, применяется для более узких практических целей (<http://www.cebm.net/index.aspx?o=8111>). То есть существование иных классификационных шкал оценки клинической информации может быть оправдано необходимостью использования более сложных категорий в рамках научных исследований, например при выполнении сравнительного анализа многих исследований по одному и тому же клиническому вопросу. Для эффективного взаимопонимания и точности оценок при создании клинических рекомендаций следует использовать шкалу GRADE.

Насколько вероятно возникновение альтернативных согласованных шкал оценки клинической информации для широкого использования в клинических рекомендациях?

Позволим себе предположить, что возникновение иных шкал оценки качества клинической информации будет не альтернативой, а дополнением в связи с двумя обстоятельствами. Первое — при решении различных клинических вопросов (прогноз, скрининг, эффективность лечения, диагностический тест, фармакоэкономический анализ) необходимо использовать различные описательные категории. Поэтому существует вероятность того, что по мере развития КЭ как науки и по мере повышения образовательного уровня врачей в этой области у потребителей информации возникнет необходимость описывать клиническую информацию более сложными понятиями. В этом случае потребуется более сложная градация качества и ценности клинической информации. Второй аспект, с которым связана вероятность возникновения новых дополнительных шкал описания клинической информации, обусловлен необходимостью формализации количественных характеристик результатов исследования.

Мера явления — важнейший, и, вероятно, единственный критерий, по которому врач и пациент делает свой клинический выбор. У каждого индивидуума существует своя грань, которая разделяет явление на белое/черное, хорошее/плохое, приемлемое/неприемлемое. Именно **мера клинического явления позволяет сделать осознанный выбор**. В соответствии с принципами существующей шкалы GRADE, этот выбор делают эксперты, используя только некоторые количественные характеристики клинической информации. Если в шкале оценки качества лечебного воздействия будет включен уровень доказательности, число больных, которых необходимо лечить, чтобы получить определенный клинический исход (Number needed to treat — NNT), или другая числовая характеристика клинического эффекта и количественная характеристика безопасности лекарственного средства, то такой показатель будет нести более полную информацию, чем существующие классификационные характеристики GRADE. Аналогично может быть представлен некий новый сложный критерий качества диагностического теста, в котором будет отражена не только вероятность систематической ошибки в результате недостатков дизайна исследования, но и количественные характеристики вероятности ложноотрицательного и ложноположительного результата при использовании диагностического теста. Но для применения таких сложных классификаций в клинических рекомендациях требуется существенное повышение общего уровня знаний в области понимания методологии клинического исследования.

Заключение

В прошлом году группа российских экспертов подготовила обзор о существующих

шкалах формализации клинической информации с целью «помочь созданию аналогичных систем оценки в России». А в заключении этого замечательного сравнительного исследования наиболее рейтинговых систем формальной оценки качества клинической информации российские эксперты, принимая приоритет системы GRADE, отмечают, что: «Возможно, потребуется некоторая модификация системы в процессе ее адаптации к условиям, сложившимся в клинической практике и здравоохранении России». С нашей точки зрения, целью подобных аналитических исследований является не создание новых формальных критериев клинической информации и не их адаптация к реальным условиям национальной медицины, а выбор наиболее качественных.

Наше клиническое и научное медицинское сообщество существенно отстает в уровне знаний принципов ДМ и методологии клинических исследований. С учетом этих обстоятельств легко представить результат творчества наших экспертов по созданию собственных формальных критериев клинической информации. Следует также учесть, что подобная работа требует развитых традиций национальных школ КЭ, достаточного финансирования деятельности экспертных групп, серьезной работы по обсуждению конкурентных проектов и прочих обязательных условий, которых нет ни в России, ни в Украине. Единственное, что нам остается — это сделать правильный выбор и использовать шанс начать активно изучать методологические основы ДМ через необходимость понимания новых классификационных характеристик клинической информации. А что касается попыток адаптации существующих клинических стандартов к условиям национальной системы здравоохранения, то гораздо более правильно **поднять уровень здравоохранения до приемлемого минимума**, при котором начнут работать международные клинические рекомендации. Выбор шкалы градации качества клинической информации в национальных клинических рекомендациях уже не является проблемой выбора по своей сути. Выбор уже сделан всем международным медицинским сообществом, и нам только остается осознанно принять этот выбор и начать использовать формальную оценку качества клинической информации по системе GRADE в клинических рекомендациях и текущей медицинской практике.

Список использованной литературы

Андреева Н.С., Реброва О.Ю., Зорин Н.А. и др. (2012) Системы оценки достоверности научных доказательств и убедительности рекомендаций: сравнительная характеристика и перспективы унификации. Медицинские технологии. Оценка и выбор, 4: 10–24 (http://medpro.ru/groups/sistema_otsenki_dostovernosti_nauchnykh_dokazatelstv_ubeditelnosti_rekomendatsii_sravniteln).

Власов В.В. (2001) Введение в доказательную медицину. Медиа Сфера, Москва, 392 с.

Гринхальх Т. (2004) Основы доказательной медицины. ГЭОТАР-МЭД, Москва, 240 с.

Флетчер Р., Флетчер С., Вагнер Э. (1998) Клиническая эпидемиология: основы доказательной медицины. Медиа Сфера, Москва, 350 с.

Andrews J.C., Guyatt G.H., Oxman A.D. et al. (2013a) GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. J. Clin. Epidemiol., 66(7): 719–725.

Andrews J.C., Schünemann H.J., Oxman A.D. et al. (2013b) GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. J. Clin. Epidemiol., 66(7): 726–735.

Atkins D., Best D., Briss P.A. et al.; GRADE Working Group (2004) Grading quality of evidence and strength of recommendations. BMJ, 328(7454): 1490.

Balshem H., Helfand M., Schünemann H.J. et al. (2011) GRADE guidelines: 3. Rating the quality of evidence. J. Clin. Epidemiol., 64(4): 401–406 ([http://www.jclinepi.com/article/S0895-4356\(10\)00332-X/fulltext](http://www.jclinepi.com/article/S0895-4356(10)00332-X/fulltext)).

Bossuyt P.M., Irwig L., Craig J., Glasziou P. (2006) Comparative accuracy: assessing new tests against existing diagnostic pathways. BMJ, 332(7549): 1089–1092.

Brunetti M., Shemilt I., Pregno S. (2013) GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. J. Clin. Epidemiol., 66(2): 140–150 ([http://www.jclinepi.com/article/S0895-4356\(12\)00134-5/fulltext](http://www.jclinepi.com/article/S0895-4356(12)00134-5/fulltext)).

Canfield S.E., Dahm P. (2011) Rating the quality of evidence and the strength of recommendations using GRADE. World J. Urol., 29(3): 311–317.

De Groot L., Abalovich M., Alexander E.K. et al. (2012) Management of thyroid dysfunction during pregnancy and postpartum: an Endocrine Society clinical practice guideline. J. Clin. Endocrinol. Metab., 97(8): 2543–2565.

Dellinger R.P., Levy M.M., Rhodes A. et al.; Surviving Sepsis Campaign Guidelines Committee including the Pediatric Subgroup (2013) Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. Crit. Care Med., 41(2): 580–637.

Goldacre B. (2013) Bad Pharma: How Medicine is Broken, and How We Can Fix it. Faber & Faber, 448 p.

Grutters J.C., van den Bosch J.M. (2006) Corticosteroid treatment in sarcoidosis. Eur. Respir. J., 28(3): 627–636.

Guyatt G.H., Oxman A.D., Akl E.A. et al. (2011a) GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J. Clin. Epidemiol., 64(4): 383–394 ([http://www.jclinepi.com/article/S0895-4356\(10\)00330-6/fulltext](http://www.jclinepi.com/article/S0895-4356(10)00330-6/fulltext)).

Guyatt G.H., Oxman A.D., Kunz R. et al. (2011b) GRADE guidelines: 2. Framing the question and deciding on important outcomes. J. Clin. Epidemiol., 64(4): 395–400 ([http://www.jclinepi.com/article/S0895-4356\(10\)00331-8/fulltext](http://www.jclinepi.com/article/S0895-4356(10)00331-8/fulltext)).

Guyatt G.H., Oxman A.D., Kunz R. et al. (2011c) GRADE guidelines: 6. Rating the quality of evidence — imprecision. J. Clin. Epidemiol., 64(12): 1283–1293 ([http://www.jclinepi.com/article/S0895-4356\(11\)00206-X/fulltext](http://www.jclinepi.com/article/S0895-4356(11)00206-X/fulltext)).

Guyatt G.H., Oxman A.D., Kunz R. et al.; GRADE Working Group (2008a) Going from evidence to recommendations. BMJ, 336(7652): 1049–1051 (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2376019>).

Guyatt G.H., Oxman A.D., Kunz R. et al.; GRADE Working Group (2008b) Incorporating considerations of resources use into grading recommendations. BMJ, 336(7654): 1170–1173 (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2394579>).

Guyatt G.H., Oxman A.D., Kunz R. et al.; GRADE Working Group (2008c) What is «quality of evidence» and why is it important to clinicians? BMJ, 336(7651): 995–998 (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2364804>).

Guyatt G.H., Oxman A.D., Kunz R. et al.; GRADE Working Group (2011d) GRADE guide-

lines: 7. Rating the quality of evidence — inconsistency. *J. Clin. Epidemiol.*, 64(12): 1294–1302 ([http://www.jclinepi.com/article/S0895-4356\(11\)00182-X/fulltext](http://www.jclinepi.com/article/S0895-4356(11)00182-X/fulltext)).

Guyatt G.H., Oxman A.D., Kunz R.; GRADE Working Group (2011e) GRADE guidelines: 8. Rating the quality of evidence — indirectness. *J. Clin. Epidemiol.*, 64(12): 1303–1310 ([http://www.jclinepi.com/article/S0895-4356\(11\)00183-1/fulltext](http://www.jclinepi.com/article/S0895-4356(11)00183-1/fulltext)).

Guyatt G.H., Oxman A.D., Montori V. et al. (2011f) GRADE guidelines: 5. Rating the quality of evidence — publication bias. *J. Clin. Epidemiol.*, 64(12): 1277–1282 ([http://www.jclinepi.com/article/S0895-4356\(11\)00181-8/fulltext](http://www.jclinepi.com/article/S0895-4356(11)00181-8/fulltext)).

Guyatt G.H., Oxman A.D., Santesso N. et al. (2013a) GRADE guidelines: 12. Preparing summary of findings tables—binary outcomes. *J. Clin. Epidemiol.*, 66(2): 158–172 ([http://www.jclinepi.com/article/S0895-4356\(12\)00032-7/fulltext](http://www.jclinepi.com/article/S0895-4356(12)00032-7/fulltext)).

Guyatt G.H., Oxman A.D., Schünemann H.J. et al. (2011g) GRADE guidelines: a new series of articles in the *Journal of Clinical Epidemiology*. *J. Clin. Epidemiol.*, 64(4): 380–382 ([http://www.jclinepi.com/article/S0895-4356\(10\)00329-X/fulltext](http://www.jclinepi.com/article/S0895-4356(10)00329-X/fulltext)).

Guyatt G.H., Oxman A.D., Sultan S. et al. (2013b) GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. *J. Clin. Epidemiol.*, 66(2): 151–157 ([http://www.jclinepi.com/article/S0895-4356\(12\)00025-X/fulltext](http://www.jclinepi.com/article/S0895-4356(12)00025-X/fulltext)).

Guyatt G.H., Oxman A.D., Sultan S. et al.; GRADE Working Group (2011h) GRADE guidelines: 9. Rating up the quality of evidence. *J. Clin. Epidemiol.*, 64(12): 1311–1316 ([http://www.jclinepi.com/article/S0895-4356\(11\)00184-3/fulltext](http://www.jclinepi.com/article/S0895-4356(11)00184-3/fulltext)).

Guyatt G.H., Oxman A.D., Vist G. et al. (2011i) GRADE guidelines: 4. Rating the quality of evidence — study limitations (risk of bias). *J. Clin. Epidemiol.*, 64(4): 407–415 ([http://www.jclinepi.com/article/S0895-4356\(10\)00413-0/fulltext](http://www.jclinepi.com/article/S0895-4356(10)00413-0/fulltext)).

Guyatt G.H., Oxman A.D., Vist G.E. et al.; GRADE Working Group (2008d) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, 336(7650): 924–926 (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2335261>).

Guyatt G.H., Thorlund K., Oxman A.D. et al. (2013c) GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles—continuous outcomes. *J. Clin. Epidemiol.*, 66(2): 173–183.

Horvath A.R. (2009) Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *Clin. Chem.*, 55(5): 853–855 (<http://www.clinchem.org/content/55/5/853.long>).

Hsu J., Brozek J.L., Terracciano L. et al. (2011) Application of GRADE: making evidence-based recommendations about diagnostic tests in clinical practice guidelines. *Implement. Sci.*, 6: 62 (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3126717>).

Jaeschke R., Guyatt G.H., Dellinger P. et al.; GRADE Working Group (2008) Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. *BMJ*, 337: a744.

Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group (2012) KDIGO Clinical Practice Guideline for Acute Kidney Injury. *Kidney inter.*, 2(Suppl.): 1–138.

Leefflang M.M., Deeks J.J., Gatsonis C., Bossuyt P.M.; Cochrane Diagnostic Test Accuracy Working Group (2008) Systematic reviews of diagnostic test accuracy. *Ann. Intern. Med.*, 149(12): 889–897.

Mustafa R.A., Santesso N., Brozek J. et al. (2013) The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. *J. Clin. Epidemiol.*, 66(7): 736–742.

Schünemann H.J., Oxman A.D., Brozek J. et al. (2008a) GRADE: assessing the quality of evidence for diagnostic recommendations. *Evid. Based Med.*, 13(6): 162–163.

Schünemann H.J., Oxman A.D., Brozek J.; GRADE Working Group (2008b) Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ*, 336(7653): 1106–1110 (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2386626>).

Tatsioni A., Zarin D.A., Aronson N. et al. (2005) Challenges in systematic reviews of diagnostic technologies. *Ann. Intern. Med.*, 142(12 Pt 2): 1048–1055.

West S., King V., Carey T.S. et al. (2002) Systems to Rate the Strength of Scientific Evidence. Evidence Report/Technology Assessment No. 47 (Prepared by the Research Triangle Institute—University of North Carolina Evidence-based Practice Center under Contract No. 290-97-0011). AHRQ Publication No. 02-E016. Rockville, MD: Agency for Healthcare Research and Quality. April 2002 (<http://www.thecre.com/pdf/ahrq-system-strength.pdf>).

Формальні критерії якості та цінності клінічної інформації

К.П. Воробйов

Резюме. У роботі виконано аналіз проблеми формалізації якості та цінності клінічної інформації для подальшого використання цих характеристик у клінічних рекомендаціях. Якісна клінічна інформація характеризується низьким рівнем систематичних помилок дослідження, а її цінність визначається тим, якою мірою досліджуваний засіб впливає на критично важливий для пацієнта клінічний результат при мінімальному рівні побічних ефектів і найбільш низькій вартості. Низький рівень систематичних помилок рандомізованого клінічного випробування забезпечується якісним відбором груп порівняння, використанням у ролі кінцевих точок прямих клінічних результатів, маскуванням втручань для учасників дослідження, застосуванням адекватних методів статистичного аналізу. Оцінка якості клінічного дослідження потребує спеціальних знань в галузі клінічної епідеміології, тому були розроблені кілька десятків систем шкальних оцінок доказовості й переконливості клінічної інформації. В останні 10 років міжнародна група експертів активно розвиває і впроваджує нову систему оцінки доказовості й сили клінічної інформації GRADE. На сьогодні ця система прийнята провідними міжнародними та національними медичними організаціями як найбільш оптимальний принцип градації клінічної інформації для використання її оцінок у клі-

нічних рекомендаціях. У роботі наведено коротку характеристику системи GRADE та обґрунтовано необхідність її використання в поточній науковій та клінічній практиці.

Ключові слова: якість клінічної інформації, формалізація, шкала доказовості.

Formal criteria for quality and value of clinical information

K.P. Vorobyov

Summary. In this paper we analyzed the problem of formalizing the quality and value of clinical information for later use of these characteristics in clinical recommendations. Good clinical information is characterized by low bias of studies, and its value is determined by the extent to which the studied treatment impacts the critical for patient clinical outcome with minimal side effects and the lowest value. The low level of systematic errors in randomized clinical trial is provided by quality selection of groups for comparisons, application of direct clinical outcomes as endpoints, masking of interventions for study participants, using appropriate methods for statistical analysis. Estimation of quality of clinical research requires special knowledge in clinical epidemiology, therefore a dozen of the scale estimation systems were developed to estimate the quality and strength of evidence. In the last ten years the international team of experts is actively developing and implementing a new system for rating quality of evidence in systematic reviews and guidelines and grading strength of recommendations in guidelines GRADE. Currently, this system has been adopted by leading international and national medical organizations as the best principle of gradation of clinical information for use of its assessment in clinical recommendations. The paper gives a brief description of the GRADE system and the necessity of its use in the current research and clinical practice.

Key words: quality of clinical information, formalization, scale of evidence.

Адрес для переписки:

Воробьев Константин Петрович
91045, Луганск, квартал 50-летия
Обороны Луганска, 1 Г
ГУ «Луганский государственный
медицинский университет», кафедра
анестезиологии, реаниматологии
и неотложных состояний ФПО
E-mail: hbo_vorobyov@mail.ru
<http://www.vkp.dsip.net>

Получено 16.12.2013