

Доказательный подход в разработке оценочных средств для аккредитации выпускников вузов

**Малахова
Татьяна Николаевна**

главный специалист Методического центра аккредитации специалистов ФГАОУ ВО Первый МГМУ им. И.М. Сеченова (Сеченовский университет) Минздрава России
akkredcentrmgmu@yandex.ru

**Чельшкова
Марина Борисовна**

доктор педагогических наук, главный специалист Методического центра аккредитации специалистов ФГАОУ ВО Первый МГМУ им. И.М. Сеченова (Сеченовский университет) Минздрава России
akkredcentrmgmu@yandex.ru

**Семёнова
Татьяна
Владимировна**

кандидат медицинских наук, директор Департамента медицинского образования и кадровой политики в здравоохранении Министерства здравоохранения Российской Федерации
info@rosminzdrav.ru

Ключевые слова: аккредитация, доказательный подход, компетентностный подход, множественный разветвляющийся кейс.

Введение

Перспективные направления развития науки в сфере образовательных измерений связаны с попытками получить ответы на вопросы: на основании каких соображений испытуемый выбрал один из ответов, предложенных в заданиях с множественным выбором? Как и каким путём в мышлении испытуемых происходит переход от знания к умениям его применять для решения жизненных или профессиональных проблем? В чём причины выбора неправильных ответов? Как обобщить результаты выполнения отдельных заданий, чтобы получить более широкие представления о способностях испытуемых?

Для ответа на эти и другие не менее важные вопросы были развиты многочисленные методы, опирающиеся на аппарат классической и современной теории тестов. С целью получения различных более сложных выводов, чем оценка по результатам выполнения теста с множественным выбором, привлекались вероятностные модели, обладающие возможностями прогноза, и статистические методы, расширяющие возможность интерпретации стандартных оценок. Однако, несмотря на возможности теорий, моделей и методов, со временем стало ясно, что понимание сущности процессов, происходящих в мышлении испытуемых при выборе ответов в стандартизованных тестах, остаётся за гранью возможностей тех, кто анализирует результаты аттестации в контексте проблем совершенствования качества образования. Более того, стала очевидной сама ограниченность и недостаточность традиционной аттестационной оценки.

На протяжении первого десятилетия XXI века для специалистов по оценке качества образования всё очевиднее становился вывод о том, что при аттестации необходима комплексная оценка подготовленности испытуемых, преодолевающая ограниченность традиционных тестов. В качестве одного из направлений развития идей комплексности оценивания был выбран переход в аттестации к многостадийным измерениям, другое направление получило название доказательного подхода в образовании¹.

Сама идея доказательного подхода (evidence-based approach) возникла в медицине в 90-е годы XX века. В этот период врачи нередко сталкивались с ситуацией, когда результаты большого количества различных исследований, клинической практики или экспериментальной работы приводили к отличающимся рекомендациям по выбору видов лечения для сходных симптомов у больных. Идея подхода родилась у доктора Арчи Кокрейна, британского эпидемиолога, боровшегося за эффективность здравоохранения. Он призывал общественность платить только за то лечение, которое было эмпирически обосновано как эффективное (Enkin, 1992). Кокрейн был активным сторонником использования доказательного подхода в лечении пациентов. В других публикациях исследователей появление доказательного подхода связывают с именем Яна Чалмерса, который в 1993 г. предложил методику создания «систематических обзоров», содержащих систематизированные результаты обширного анализа различных исследований с тем, чтобы помочь практикующим врачам в выборе оптимальных методов лечения для каждого пациента².

Благодаря интенсивному развитию информационных технологий в сфере создания больших баз данных (Big Data), позволявших собрать и интегрировать огромные массивы информации, в первое десятилетие XXI века в ряде стран стало

уделяться значительное внимание доказательному подходу не только в медицине, но и в социологии, и в борьбе с преступностью. Со временем доказательный подход нашёл своё место в образовании, в частности, в сфере образовательных измерений. Здесь доказательный подход приобрёл форму поиска обоснований для оценивания тех или иных результатов выполнения заданий тестов.

В статье предлагаются основные положения доказательного подхода в аккредитационном тестировании и рассматривается пример его использования при создании множественного кейса, предназначенного для аккредитации выпускников медицинских вузов России.

Теоретические и методические основы разработки инструментария в рамках доказательного подхода

Со временем, во втором десятилетии XXI века, работы в области доказательного подхода в сфере образовательных измерений приобрели характер методологии, которую в целом можно определить как теорию создания методов инновационного оценивания и применения его результатов для принятия управленческих решений по совершенствованию качества образования. Методология доказательного подхода, ориентированная на создание нового поколения оценочных средств, должна гарантировать, что пути представления формата доказательств, их сбора и интерпретации совместимы с основной целью оценивания и обеспечивают координацию работы различных специалистов по созданию инструментария, таких как статистики, авторы заданий, администраторы процессов предъявления измерителей в аттестации и проектировщиков их интерфейса.

Идеологами доказательного подхода была разработана совокупность принципов, которая в виде, трансформированном для сферы образования, включает:

- Принцип обоснованности решений, принимаемых в образовании по результатам измерений, их направленность на улучшение методов, практик и программ обучения.

- Принцип осознанности возможностей и ограничений различных видов доказательств или обоснований в образовательных измерениях.

¹ Semenova T., Sizova Z., Zvonnikov V., Masalimova A., Ersozlu Z. The Development of Model and Measuring Tool for Specialists Accreditation// EURASIA J. Math., Sci Tech. Ed 2017; 13(10):6779–6788.

² Mazurek M.B., Fineout-Overholt E. Evidence-based practice in nursing & healthcare: a guide to best practice / Bernadette Mazurek Melnyk. — 2nd ed. 624 p.

■ Принцип необходимости поиска применимости методов доказательств и обоснований в реальных обстоятельствах оценивания.

■ Принцип совместимости путей и способов сбора доказательств с основной целью оценивания.

■ Принцип доверия и принятия доказательств или обоснований пользователями результатов измерений.

Представленная совокупность принципов служит регламентирующей основой для построения дизайна инструментария в рамках доказательного подхода, охватывающего, в целом, все его виды, от стандартизированных тестов и практических заданий до портфолио и оценочных кейсов. Вне зависимости от видов инструментария проектирование его дизайна должно опираться на основную идею доказательного подхода, нацеленного на поиск достаточных обоснований того, что испытуемый справился с заданием не случайно, а вполне осознанно применил свои знания для решения практических или профессиональных задач. Методики разработки таких заданий стали называть «центрированным на доказательствах дизайном» («Evidence-Centered Design — ECD»). Они легли в основу создания инновационного инструментария, привлёкшего внимание многих крупнейших служб тестирования в конце второго десятилетия XXI века³.

Рост интереса к центрированному на доказательствах дизайну оценочных средств, наблюдающийся в наши дни, возник не случайно. Возможности доказательного подхода, позволяющие глубже проникнуть в суть процессов, происходящих при выполнении заданий испытуемыми, и отслеживать как испытуемый способен применить полученные знания в решении практических проблем, оказались во многом созвучны с идеями компетентностного подхода к трактовке качества результатов образования, с его прагматической ориентацией на оценивание способности испытуемых применять знания в профес-

³ Burke E., Kleeman J. Assessing for Situational Judgment Designing, deploying and getting value from Situational Judgment Assessments, White Paper, 2018, 43 p. Режим доступа: www.questionmark.com (дата обращения: 10.12.2018 г.).

сиональных или жизненных проблемных ситуациях⁴.

Создавая компетентностно-ориентированные задания в рамках доказательного подхода, их автор должен:

■ сформулировать компетентностно-ориентированную жизненную или профессиональную проблему;

■ трансформировать проблему в форму совокупности заданий с множественным выбором или в ситуационное задание, содержащее, по возможности, некоторую дилемму или допускающие более и менее эффективные решения;

■ собрать наиболее релевантные данные для обоснования возможных ответов испытуемых в электронном формате;

■ критически оценить данные, которые были собраны, чтобы определить их достоверность, надёжность и применимость, а затем синтезировать эти данные для использования в качестве критериев при оценивании ответов испытуемых.

Конечно, каждая из этих позиций представляется достаточно трудной тем авторам, кто привык за считанные минуты разрабатывать задания с множественным выбором, проверяющие понятийный аппарат. Понятно, что подобные понятийные задания не могут быть компетентностно-ориентированными и, тем более, не годятся для реализации центрированного на доказательствах подхода, поскольку бессмысленно ждать от испытуемого обоснования выбора того или иного определения из числа предложенных в ответах. Выбор формата заданий во многом определяет природа оцениваемого конструкта. Формат заданий зависит от того, какие поведенческие или профессиональные качества должны продемонстрировать владение этим конструктом и какие задания или ситуации должны выявить его.

Чаще всего обращаются к заданиям с множественным выбором, при разработке которых в рамках центрированного на доказательствах дизайна используются несколько приёмов. Один из наиболее распространённых способов состоит в со-

⁴ Semenova T., Sizova Z., Chelyshkova M., Dorozhkin E., Malygin A. Fairness and Quality of Data in Healthcare Professionals' Accreditation // Modern Journal of Language Teaching Methods. Режим доступа: www.mjltm.org (дата обращения: 10.12.2018).

четании формы заданий с множественным выбором и последующих обоснований, имеющих формат мини-эссе или других способов предоставления испытуемыми своего развернутого ответа. В этом случае в задании испытуемого просят после выбора ответа, который он считает верным, обосновать свой выбор и привести аргументы.

Трудность формулировок подобных заданий состоит в том, что приведенная ситуация или направления аргументации, предлагаемые в задании, не должны служить подсказкой по выбору верного ответа. То есть недопустимым считается случай, когда испытуемый, прочитав вторую часть задания, требующую обоснования, возвращается назад и меняет опцию выбранного ответа. Конечно, чтобы такие задания формулировать, необходимо иметь достаточный опыт. Иногда авторы вносят в число ответов несколько проблемных опций, порождающих дилемму и требующих обоснования проблемной ситуации для её разрешения. В одних случаях, испытуемому дают на выбор ответы, которые обычно предполагают привлечение взаимоисключающих аргументов к решению сформулированной ситуации. В других случаях, испытуемого просят опознать аргументы и ранжировать их в своём обосновании подхода к решению проблемной ситуации, выявить различные точки зрения в какой-либо статье, привести доводы в защиту или для опровержения определённой точки зрения и т.д. Наиболее удачно доказательный подход реализуется при разработке кейсов и ситуационных заданий, содержащих проблемы, профессионального или практического характера, которые позволяют испытуемому продемонстрировать способность к применению выбранного ответа и обосновать в процессе применения свой выбор.

В общем случае кейс можно определить как структурированную и смоделированную ситуацию, отображающую конкретную профессиональную проблему и требующую от испытуемых проявления совокупности навыков (аналитических, критических, информационных, сравнительных, оценочных и т.д.) для получения оптимальных решений в условиях, когда

контекст проблемы может варьировать⁵. Вариация контекста, осуществляемая испытуемым в процессе выполнения кейса, требует от него объяснений и обоснований, составляющих ядро доказательного подхода в образовательных измерениях. Ситуационное задание, в отличие от кейса, обладает жёстко детерминированным контекстом, поэтому столь широких возможностей для реализации доказательного подхода не предоставляет. Лишь в том случае, когда ситуационное задание сопровождается заданиями с выбором ответов, от испытуемых можно требовать обоснование выбора того или иного ответа.

Работа по созданию множественных (оценивающих несколько переменных) ситуационных заданий ведётся в Методическом центре аккредитации специалистов здравоохранения на базе Первого Московского государственного медицинского университета имени И.М. Сеченова уже более года. В работе принимают участие представители различных вузов России, готовящие будущих специалистов здравоохранения. Обращение к множественным (многомерным) ситуационным заданиям было вызвано необходимостью оценивания уровня освоения совокупности трудовых функций, представленных в профессиональных стандартах. Идея многомерности потребовала выделения несколько подмножеств вопросов к каждому ситуационному заданию, охватывающих в совокупности от 3-х до 5-ти различных переменных. К вопросам приводятся задания с множественным выбором, разработанные для каждого вопроса кейса.

В основу содержания ситуационных заданий были положены клинические рекомендации, содержащие результаты обоснованного и осмысленного анализа лучших результатов клинических исследований для выбора методов лечения конкретной болезни. Однако мы живём в эпо-

⁵ Методические рекомендации по оцениванию специалистов здравоохранения при аккредитации (Выпуск 4) / [Ж.М. Сизова, В.И. Звонников, М.Б. Чельщикова: под ред. Ж.М. Сизовой]: ФГАОУ ВО Первый Московский государственный медицинский университет имени И.М. Сеченова Минздрава России (Сеченовский университет), Методический центр аккредитации специалистов — Москва: Издательство Первого МГМУ имени И.М. Сеченова. 2018 — 56 с.

ху, когда медицина стремительно развивается, каждый год в мире разрабатываются новые методы лечения, создаются научные медицинские школы и в практику лечения поступают новейшие лекарственные средства и современное оборудование. Очевидно, что все эти нововведения затрудняют использование единых подходов к лечению и требуют адекватного развития оценочных средств для аккредитации специалистов здравоохранения. Одним из направлений такого развития может стать создание множественных, разветвляющихся кейсов, где выбор траектории разветвления должен обосновывать испытуемый в рамках доказательного подхода.

Структура такого кейса представлена на рисунке.

При выполнении такого кейса на первом этапе все испытуемые получают единую проблемную ситуацию и общий контекст. На втором этапе испытуемые получают для выбора клинические рекомендации и перечень научных школ, к опыту которых обращаются при подготовке специалистов здравоохранения в различных вузах России.

Однако не следует думать, что каждая научная школа имеет принципиальные отличия в подходах к ведению больных с одной и той же нозологией, что одну и ту же болезнь лечат по-разному в различных регионах Российской Федерации. Представители научных школ являются членами профессионального сообщества по специальности, поэтому любая научная школа свои взгляды на ту или иную медицинскую проблему, пути её решения так или иначе соотносит с известными клиническими рекомендациями. Несмотря на возможные различия в подходах к диагностике и тактике ведения пациентов различных нозологий, глобальный подход к решению той или иной проблемы является общим.

Важно отметить, что клинические рекомендации — это чрезвычайно «живые» документы, которые регулярно обновляются с учётом новых сведений и результатов рандомизированных клинических исследований. Позиции, которые излагаются в клинических рекомендациях, сопровождаются строгим ранжированием уровня достоверности доказательств того или иного метода



Рис. Структура множественного разветвляющегося кейса

диагностики или лечения⁶. Клинические рекомендации, имеющие 1 уровень достоверности доказательств, представлены проспективными рандомизированными контролируемые исследованиями; достаточным количеством исследований с участием большого количества пациентов и получением большого количества данных. Кроме того, этот уровень достоверности может опираться на результаты крупных мета-анализов и, как минимум, на одно хорошо организованное рандомизированное контролируемое исследование.

В то же самое время позиции, имеющие уровни достоверности доказательств 2 и 3, также являются убедительными. Так, в частности, уровень достоверности доказательств 2 основан на результатах проспективных рандомизированных исследований; нескольких исследований с небольшим количеством пациентов; одного хорошо организованного проспективного исследования когорты; ограниченных мета-анализов, но проведённых на хорошем уровне и непрезентативных результатах исследования в отношении целевой популяции.

Уровень достоверности доказательств 3 — это данные нерандомизированных контролируемых исследований и исследований с недостаточным контролем, ретроспективных или наблюдательных исследований, серии клинических наблюдений; могут быть использованы рандомизированные клинические исследования с 1 значительной или, как минимум, 3 незначительными методологическими ошибками. Возможны противоречивые данные, не позволяющие сформировать окончательную рекомендацию по тактике ведения пациента.

Но существуют и рекомендации иного рода, имеющие на сегодняшний день неоднозначную трактовку — 4 уровень достоверности доказательств, который основан на мнении эксперта или данных из отчёта экспертной комиссии, экспериментально подтверждённые и теоретически обоснованные.

⁶ Guyatt G., Oxman A.D., Aki E.A., Kunz R. et al. GRADE guidelines: 1. Introduction — GRADE evidence profiles and summary of findings tables / J Clin Epidemiol 2011; 64 (4): 383–394. Режим доступа: https://www.essentialvidenceplus.com/product/ebm_loe.cfm?show=grade (дата обращения: 10.12.2018 г.).

Вот в такой ситуации допустимо вариативное решение проблемы с учётом выбора той или иной научной школы. Есть множество примеров подобного подхода, одним из которых может служить ситуация с лечением гипертонических кризов.

Отсутствие на сегодняшний день чётких клинических рекомендаций по этой проблеме во всём мире объективно обосновано невозможностью проведения рандомизированных клинических исследований у этой категории больных, а поэтому невозможностью определить рекомендации по тактике ведения больных с гипертоническими кризами, как класс уровень достоверности доказательств 1. В таких случаях для принятия решения используются, так называемые, согласованные документы, в разработке которых принимают различные группы экспертов и представители различных научных школ.

Аналогично, уровень убедительности рекомендаций А основан на высоком уровне доказательности (как минимум 1 убедительная публикация 1 уровня доказательности, показывающая значительное превосходство пользы над риском). Уровни убедительности рекомендаций В (средний) и С (слабый) имеют свои критерии и используются соответственно как «метод/терапия второй линии» и «нет возражений против данного метода/терапии». Уровень убедительности рекомендаций D свидетельствует о превосходстве риска выбранной тактики лечения над его пользой.

Затем, после выбора траектории, наступает этап обоснования выбора, когда испытуемые могут проявить свои знания новейших методик лечения, появившихся новых лекарств и технологий. После обоснования сделанного выбора каждый испытуемый получает набор заданий для выполнения, ориентированный по содержанию на клинические рекомендации или определённую научную школу.

Таким образом, множественный разветвляющийся кейс позволяет отобразить уникальность отдельных научных школ и учесть клинические рекомендации, призванные помочь врачу принять оптимальные решения о рациональной помощи в различных клинических ситуациях

В целом, применение кейсов и ситуационных заданий в оценочных аккредита-

ционных процедурах позволяет испытуемым:

- продемонстрировать свои компетенции, навыки и умения по практическому применению полученных знаний;
- проявить свои способности по работе со значительными объёмами информации, в том числе применить умения собирать дополнительную информацию, анализировать её, интегрировать и интерпретировать для поставленной проблемы;
- применить навыки критического мышления и системного мышления;
- продемонстрировать навыки принятия решений на основе анализа ситуации;
- проявить навыки проведения диагностики, выявления причин возникшей ситуации, анализа её связей с внешними факторами в реальной жизни и т.д.

Рекомендации для авторов по разработке оценочных средств для аккредитации в рамках доказательного подхода

В целом, при разработке дизайна оценочного средства, центрированного на до-

казательствах, авторам можно порекомендовать придерживаться схемы действий, представленной в таблице.

Конечно, ряд действий авторов, из числа приведённых в таблице, достаточно традиционен. Другие действия, связанные с введением доказательного подхода, носят для них непривычный характер. Например, необычно, и непросто для авторов собирать обосновывающую графическую информацию, заранее предвидя пути обоснования, которые могут выбрать испытуемые. Поэтому доказательный подход нельзя внедрить в работу тестовых структур прямым директивным решением. Ему нужно учиться, осваивая постепенно отдельные его возможности и положения. Несмотря на трудности его следует внедрять в массовые оценочные процедуры, поскольку обращение к методологии центрированного на доказательствах подхода означает продвижение в когнитивном, психометрическом, и технологическом инструментарии, создаваемом для образовательных измерений.

Ключевые действия авторов оценочных средств в рамках доказательного подхода

Провести анализ содержательной области	Выявить — что является наиболее важным в этой области? Определить — какие задания являются ключевыми для оценки владения содержанием этой области, умениями или компетенциями? Обдумать — как представить собрать обосновывающую графическую информацию, которая может быть привлечена испытуемыми при обосновании ответов (формулы, карты, рисунки и т.д)?
Провести моделирование области	Выявить — как операционализировать ключевые аспекты содержания в терминах измеряемого конструкта? Определить — как структурировать содержание оценочного средства?
Выстроить концептуальную структуру оценивания	Выявить — какой инструментарий нужно применить для этого оценивания? Решить — как разработать спецификацию оценочного средства?
Провести апробацию	Выявить — как выбрать и представить задания? Определить — как провести оценочную процедуру? Решить — как проверять ответы испытуемых (в частности, как проверять обосновывающую часть)? Решить — как осуществить представление заданий испытуемым, сбор данных и анализ ответов?
Провести предъявление	Определить — когда предъявить задания испытуемым и как сообщить результаты выполнения?