

К. Т. МОМЫНАЛИЕВ

ДОКАЗАТЕЛЬНАЯ ЛОГИКА

Как построить, провести и защитить клинико-лабораторное испытание ИВД перед экспертизой



ДОКАЗАТЕЛЬСТВА
КАЧЕСТВО
ДОВЕРИЕ
РЕЗУЛЬТАТ



КЛИНИКО-ЛАБОРАТОРНЫЕ ИСПЫТАНИЯ
МЕДИЦИНСКИХ ИЗДЕЛИЙ
ДЛЯ ДИАГНОСТИКИ IN VITRO

Куват Момыналиев

Доказательная логика

<https://litres.ru/74378887>

SelfPub; 2026

Аннотация

Как доказать, что результат медицинского изделия для диагностики *in vitro* действительно заслуживает доверия? Хороших аналитических характеристик для этого недостаточно. Ошибка в выборе образцов, метода сравнения, дизайна исследования или статистического подхода может сделать убедительные на первый взгляд результаты непригодными для доказательства заявленного назначения.

Эта книга — практическое руководство по клинико-лабораторным испытаниям и доказательной оценке МИ ИВД. Читатель пройдет весь путь от назначения изделия и научной обоснованности до аналитической и клинической эффективности, выбора дизайна, расчета выборки, статистической обработки и итогового вывода. Отдельно разобраны экспресс-тесты, ПЦР, NGS, микробиология, инфекционные заболевания, онкология, скрининг и медицинский ИИ. В основе книги — практические примеры, расчеты, схемы, таблицы и разбор типичных ошибок.

Главный принцип прост: вывод об изделии не может быть шире доказательств, которыми он подтвержден.

Содержание

Глава	6
Об авторе	7
Аннотация	9
Предисловие	10
Часть I. Международные и методологические основы клинико-лабораторных испытаний медицинских изделий для диагностики <i>in vitro</i>	14
Глава 1. Международная система регулирования медицинских изделий для диагностики <i>in vitro</i>	17
Глава 2. Международная методологическая рамка и ее отражение в регулировании МИ ИВД	27
Глава 3. Назначение изделия и его клиническая роль	40
Глава 4. Доказательная модель оценки медицинских изделий для диагностики <i>in vitro</i>	50
Глава 5. Жизненный цикл медицинского изделия для диагностики <i>in vitro</i>	60
Глава 6. Источники доказательств: исследования, литература и данные реальной практики	73
Глава 7. Критическая оценка и взвешивание клинических доказательств	86
Глава 8. Документирование, демонстрация и генерация доказательств	100

Часть II. Научная обоснованность медицинских изделий для диагностики <i>in vitro</i>	112
Глава 9. Понятие научной обоснованности	115
Глава 10. Связь анализа с клиническим состоянием, физиологическим или патологическим процессом	127
Глава 11. Хорошо изученный и новый анализ: различия в доказательной стратегии	140
Глава 12. Источники данных для установления научной обоснованности	152
Глава 13. Литературный поиск при установлении научной обоснованности	166
Глава 14. Систематические обзоры в подтверждении научной обоснованности	177
Глава 15. Критическая оценка и анализ данных о научной обоснованности	190
Глава 16. Документирование научной обоснованности в составе доказательной документации	204
Часть III . Аналитическая эффективность медицинских изделий для диагностики <i>in vitro</i>	220
Глава 17. Общие принципы подтверждения аналитической эффективности	223
Глава 18. Правильность и прецизионность	235
Глава 19. Линейность и диапазон измерений	256
Глава 20. Предел обнаружения и предел	277

количественного определения

Глава 21. Аналитическая специфичность и
интерференция 297

Глава 22. Стабильность и сохранение
аналитической пригодности 318

Глава 23. Прослеживаемость и калибровка 337

Конец ознакомительного фрагмента. 353

Доказательная логика

Глава

Об авторе

Куват Темиргалиевич Момыналиев — доктор биологических наук, специалист в области биомедицины, молекулярной биологии, медицинских изделий для диагностики *in vitro* и их регуляторной оценки.

Более 25 лет занимается научной, экспертной и организационной деятельностью в сфере биотехнологий и медицинских технологий. Работал в научных и экспертных организациях России и Казахстана, руководил исследовательскими подразделениями и Национальным центром биотехнологии Республики Казахстан.

В настоящее время — руководитель Научно-практического центра испытаний медицинских изделий ФГБУ «ВНИИИМТ» Росздравнадзора, профессор Московского государственного технического университета им. Н. Э. Баумана.

С 2017 года профессионально занимается экспертизой и регулированием медицинских изделий. Основные направления работы — клинико-лабораторные испытания и доказательная оценка медицинских изделий для диагностики *in vitro*, молекулярно-генетические технологии, лабораторная диагностика, стандартные образцы, качество и безопасность медицинских технологий.

Автор научных публикаций, участник российских и международных проектов в области биомедицины и регулирования

ния медицинских изделий.

Аннотация

Как доказать, что результат медицинского изделия для диагностики *in vitro* действительно заслуживает доверия? Хороших аналитических характеристик для этого недостаточно. Ошибка в выборе образцов, метода сравнения, дизайна исследования или статистического подхода может сделать убедительные на первый взгляд результаты непригодными для доказательства заявленного назначения.

Эта книга — практическое руководство по клинико-лабораторным испытаниям и доказательной оценке МИ ИВД. Читатель пройдет весь путь от назначения изделия и научной обоснованности до аналитической и клинической эффективности, выбора дизайна, расчета выборки, статистической обработки и итогового вывода. Отдельно разобраны экспресс-тесты, ПЦР, NGS, микробиология, инфекционные заболевания, онкология, скрининг и медицинский ИИ. В основе книги — практические примеры, расчеты, схемы и разбор типичных ошибок.

Главный принцип прост: вывод об изделии не может быть шире доказательств, которыми он подтвержден.

Предисловие

Клинико-лабораторные испытания медицинских изделий для диагностики *in vitro* занимают особое место в системе их оценки. Такие изделия, как правило, не воздействуют на пациента напрямую, но их результат может определять дальнейшее медицинское решение: подтвердить или исключить диагноз, выбрать терапию, оценить динамику заболевания, установить инфекционный статус, определить чувствительность микроорганизма к антимикробному препарату, выявить риск или направить пациента на дополнительное исследование.

Поэтому клинико-лабораторное испытание МИ ИВД не может быть простой лабораторной проверкой набора реагентов, анализатора, тест-системы, картриджа, питательной среды, диска с антимикробным препаратом, программного модуля или алгоритма. В центре такой оценки находится вопрос о том, можно ли доверять результату конкретного изделия в заявленном применении, на заявленном материале, у заявленного пользователя и в заданном клинико-лабораторном сценарии.

На практике именно здесь возникает множество ошибок. Изделие испытывается на удобных образцах, а вывод распространяется на широкое назначение. Метод сравнения принимается за абсолютную истину, хотя он сам имеет огра-

ничения. Недействительные, сомнительные, пограничные и невыданные результаты исчезают из анализа. Общий процент совпадений скрывает слабую работу изделия в критически важной подгруппе. Данные по одной матрице переносятся на другую. Хорошая аналитическая характеристика ошибочно принимается за достаточное подтверждение клинической применимости результата.

Настоящее издание написано как практическое руководство по доказательной логике клинико-лабораторных испытаний МИ ИВД. Его задача — не заменить нормативные документы, стандарты или регуляторные процедуры, а помочь читателю понять, как выстраивать испытание так, чтобы его результаты были проверяемыми, интерпретируемыми и пригодными для экспертной оценки.

В настоящем издании последовательно проводится один принцип: испытание должно идти от назначения изделия к допустимому выводу. Сначала необходимо понять, какое применение заявлено, какую роль играет результат, какой материал используется, какая популяция или группа образцов должна быть представлена, какой метод сравнения или референтная рамка применимы, какие показатели нужно рассчитать и какие критерии приемлемости должны быть установлены заранее. Только после этого можно говорить о программе испытаний, объеме выборки, статистической обработке, протоколе, акте и экспертном заключении.

Эта практическая задача определяет и структуру книги.

Главы построены по единому методическому принципу: постановка методологического или регуляторного вопроса, его разбор, практические и расчетные примеры, анализ характерных ошибок и формулирование выводов. Такое единообразие позволяет рассматривать разные виды изделий и этапы доказательной оценки в сопоставимой логике, а книгу можно использовать как для последовательного чтения, так и в качестве справочника, обращаясь напрямую к нужной главе.

Рисунки, схемы и другие повторяющиеся элементы выполняют ту же функцию. Они служат навигационными и аналитическими инструментами: схемы показывают последовательность действий и решений, позволяют сопоставлять требования, показатели и доказательные ситуации, а расчетные примеры демонстрируют применение рассматриваемых подходов. Такое единообразие оформления помогает проследить логику анализа от исходного вопроса до допустимого вывода: что оценивается, какими данными это подтверждается, какие ограничения необходимо учитывать и какой вывод позволяют сделать полученные доказательства. При этом конкретное содержание каждой главы определяется назначением изделия, типом результата и клиническим контекстом его применения.

Издание ориентировано на специалистов, которым необходимо выполнить лабораторную работу и одновременно понимать ее доказательный смысл. В клинико-лабораторных испытаниях МИ ИВД правильная постановка исследования

и корректность расчетов имеют значение только в составе прослеживаемой доказательной цепочки: от назначения изделия и клинического материала до статистического анализа, учета проблемных результатов, ограничений и итогового вывода.

Автор надеется, что издание будет полезно специалистам испытательных лабораторий, производителям, разработчикам, экспертам, специалистам по регистрации, клиническим лабораторным специалистам, врачам лабораторной диагностики, микробиологам, молекулярным биологам, биоинформатикам, специалистам по качеству, преподавателям и обучающимся, а также всем, кто участвует в подготовке, проведении, документировании и оценке клинико-лабораторных испытаний медицинских изделий для диагностики *in vitro*.

Часть I. Международные и методологические основы клинико-лабораторных испытаний медицинских изделий для диагностики in vitro

Аннотация части

Часть I вводит читателя в доказательную логику оценки медицинских изделий для диагностики in vitro (МИ ИВД). Для таких изделий недостаточно показать, что реагент, прибор, программный модуль или тест-система технически способны получить лабораторный результат. Важно понять, какое медицинское значение имеет этот результат, почему выбранный анализ связан с определенным состоянием или клинической задачей, насколько надежно конкретное изделие получает результат и может ли этот результат использоваться в заявленном клиническом сценарии.

Особенность изделий для диагностики in vitro состоит в том, что они обычно не воздействуют на пациента напрямую, но их результат может определять дальнейшее медицинское решение. Положительный, отрицательный, количественный, молекулярный, морфологический или алгоритмический результат может повлиять на постановку диагноза, исключе-

ние заболевания, выбор терапии, мониторинг лечения, прогноз, скрининг, эпидемиологическое решение или маршрутизацию пациента. Поэтому регуляторная оценка таких изделий не может ограничиваться проверкой технологии. Она должна отвечать на более широкий вопрос: можно ли доверять результату изделия как основанию для заявленного медицинского применения.

В этой части рассматривается общая система понятий, без которой невозможно корректно планировать клиническо-лабораторные испытания МИ ИВД и готовить доказательную базу. Сначала описывается международная и евразийская логика регулирования изделий для диагностики *in vitro*, затем раскрывается значение назначения изделия, доказательная модель оценки, жизненный цикл изделия, источники доказательств и роль литературного поиска. Завершает часть различие трех видов доказательной работы: документирования уже имеющихся данных, демонстрации их применимости и генерации новых доказательств.

Цель части

Цель настоящей части состоит в том, чтобы сформировать у читателя целостное представление о доказательной оценке медицинских изделий для диагностики *in vitro* и показать, какое место в этой оценке занимают клиническо-лабораторные испытания.

После изучения этой части читатель должен понимать, почему исходной точкой оценки является назначение изде-

лия; как связаны между собой научная обоснованность, аналитическая эффективность и клиническая эффективность; какие источники данных могут использоваться для формирования клинических доказательств; почему литературные данные, аналитические исследования и клинико-лабораторные испытания выполняют разные доказательные функции; и каким образом международная доказательная логика соотносится с подготовкой документов в евразийской и российской регуляторной практике.

Глава 1. Международная система регулирования медицинских изделий для диагностики *in vitro*

Результат МИ ИВД как основание медицинского решения.

Международная система регулирования медицинских изделий для диагностики *in vitro* сформировалась вокруг особой роли лабораторного результата в медицинском решении. В отличие от имплантируемого изделия, хирургического инструмента, аппарата для терапии или средства, непосредственно воздействующего на организм пациента, изделие для диагностики *in vitro* чаще всего не изменяет состояние пациента само по себе. Его влияние возникает опосредованно — через информацию, которую получает врач, лабораторный специалист, пациент или система здравоохранения (рис.1.1). Именно эта информация может стать основанием для постановки диагноза, исключения заболевания, выбора терапии, мониторинга лечения, оценки прогноза, скрининга, маршрутизации пациента или эпидемиологического решения.

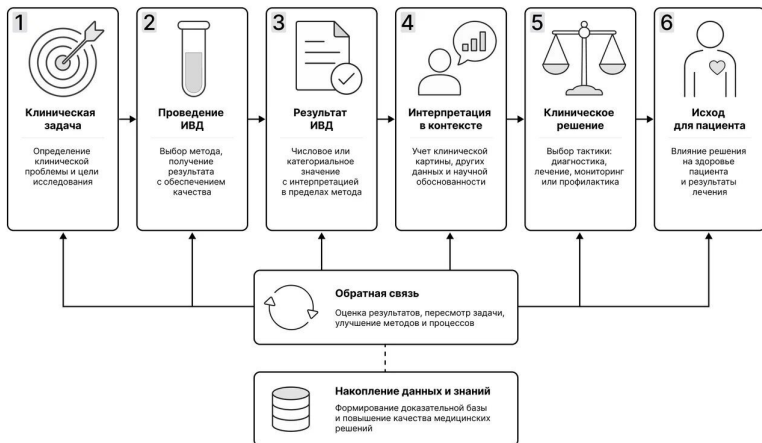


Рисунок 1.1. Результат лабораторного исследования как основа медицинского решения.

Эта особенность делает регулирование изделий для диагностики *in vitro* самостоятельной областью. На первый взгляд может показаться, что такие изделия безопаснее многих других медицинских изделий: они используются вне организма человека, работают с образцом, а не с пациентом, и не оказывают непосредственного физического воздействия. Однако регуляторная логика исходит не только из характера контакта с пациентом. Она учитывает последствия результата. Ошибочный лабораторный вывод может привести к пропуску заболевания, ненужному лечению, неправильно-

му выбору препарата, задержке терапии, необоснованному повторному вызову пациента, неверной оценке эпидемиологической ситуации или ошибочной маршрутизации клинического случая. Поэтому для изделий для диагностики *in vitro* риск часто связан не с самим процессом исследования, а с тем решением, которое принимается на основании результата.

Именно поэтому такое изделие нельзя оценивать только как техническую систему. Если тест-система правильно определяет аналит, прибор стабильно регистрирует сигнал, а программа корректно формирует числовой результат, это еще не означает, что изделие пригодно для медицинского применения. Техническая работоспособность является необходимым условием, но она не равна клинической пригодности. Для регуляторной оценки важно понять, какой медицинский смысл имеет результат, насколько надежно он получается данным изделием и может ли он использоваться именно в той ситуации, для которой изделие заявлено изготовителем.

Медицинское изделие для диагностики in vitro — это изделие, предназначенное изготовителем для исследования *in vitro* образцов биологических материалов человека с целью получения информации о физиологическом или патологическом состоянии, заболевании, предрасположенности к определенному состоянию, совместимости тканей, прогнозе реакции на терапию, выборе лечения или контроле

ле лечения. К таким изделиям относятся реагенты, калибраторы, контрольные материалы, приборы, анализаторы, наборы реагентов, программное обеспечение и иные изделия, если они предназначены для получения медицински значимой информации при исследовании образцов человеческого происхождения.

Клиническая роль результата — функция, которую результат медицинского изделия для диагностики *in vitro* выполняет в заявленном медицинском применении: диагностика или исключение состояния, скрининг, мониторинг, прогноз, стратификация риска, выбор терапии, либо иная предусмотренная назначением задача. Клиническая роль определяет контекст интерпретации результата и влияет на характер доказательств, необходимых для подтверждения заявленного применения изделия. Один и тот же анализ может иметь разные клинические роли. Поэтому доказательства, полученные для одной роли результата, не могут автоматически переноситься на другую без оценки их применимости.

Ключевым объектом оценки в диагностике *in vitro* становится как метод, так и результат в его клиническом контексте. Один и тот же лабораторный сигнал может иметь разное значение в зависимости от того, где и зачем он используется. Например, выявление нуклеиновой кислоты инфекционного агента у пациента с выраженными симптомами острой инфекции, у бессимптомного донора, у пациента после ле-

чения или в рамках эпидемиологического мониторинга не является одной и той же доказательной ситуацией. В первом случае результат помогает подтвердить клиническое подозрение и принять решение о лечении. Во втором — используется для защиты реципиента и системы донорства, где цена пропуска случая особенно высока. В третьем — может отражать остаточный след, персистенцию или неактивный процесс. В четвертом — служит не столько индивидуальному клиническому решению, сколько общественному здравоохранению.

Аналогичная логика действует в онкологии. Выявление мутации в опухолевом материале может иметь диагностическое значение, помогать уточнить молекулярный подтип опухоли, служить основанием для выбора таргетной терапии или использоваться для мониторинга минимальной остаточной болезни. Формально речь может идти об одном и том же молекулярном событии, но доказательная задача меняется в зависимости от клинической роли результата. Если результат влияет на выбор терапии, регуляторная оценка должна учитывать способность теста обнаружить вариант, а также то, что именно этот вариант имеет значение для терапевтического решения в данной нозологии, материале и популяции.

Поэтому в международной методологии доказательной оценки изделий для диагностики *in vitro* удобно выделять три взаимосвязанных вопроса. Первый — почему опреде-

ляемый анализ или иной объект исследования имеет медицинское значение. Второй — может ли конкретное изделие надежно выявлять, измерять или характеризовать этот объект. Третий — пригоден ли получаемый результат для заявленного медицинского применения. Этим вопросам соответствуют три взаимосвязанных элемента доказательной оценки: научная обоснованность, аналитическая эффективность и клиническая эффективность.

Такая трехкомпонентная модель представляет собой методологическую рамку оценки доказательств и не должна восприниматься как дословное воспроизведение терминологии каждого нормативного акта. В конкретной юрисдикции она реализуется через установленные требования к безопасности, качеству и эффективности изделия, его аналитическим и клиническим данным и составу регистрационного досье.

Эти три уровня нельзя подменять друг другом. В этом состоит одна из главных причин, по которой международное регулирование МИ ИВД уделяет такое внимание назначению изделия, клинической роли результата и границам доказательной базы.

Международная гармонизация регулирования медицинских изделий развивалась через деятельность GHTF и затем IMDRF. Эти структуры не заменяют национальные или региональные регуляторные системы и не выдают разрешения на обращение изделий. Их значение состоит в другом: они

формируют общий методологический язык регулирования. Благодаря таким документам разные юрисдикции используют сходные подходы к классификации изделий, оценке риска, клиническим доказательствам, системе менеджмента качества, жизненному циклу, пострегистрационному наблюдению и оценке изменений.

GHTF — *Международная целевая группа по гармонизации регулирования медицинских изделий. Она сформировала базовые подходы к классификации, оценке соответствия, системе менеджмента качества, клинической оценке и надзору на рынке.*

IMDRF — *Международный форум регуляторов медицинских изделий. Он развивает и обновляет международные подходы к регулированию медицинских изделий, включая изделия для диагностики in vitro. IMDRF не заменяет национальные и региональные регуляторные системы, но его документы используются как методологический ориентир при построении таких систем*

Риск ошибочного результата.

Для изделий для диагностики in vitro особенно важен риск-ориентированный характер регулирования. В этой области риск определяется как физическим воздействием изделия, так и последствиями ошибочного результата. Ложно-отрицательный результат теста на инфекцию может привести к задержке лечения или дальнейшему распространению заболевания. Ложноположительный результат может стать

причиной ненужной изоляции, тревоги, повторных исследований или необоснованного лечения. Ошибка теста для подбора онкологической терапии может лишить пациента подходящего препарата или, наоборот, направить его на лечение, от которого он не получит ожидаемой пользы. Ошибка в неонатальном скрининге может привести либо к пропуску состояния, где раннее вмешательство критично, либо к необоснованному повторному вызову семьи и нагрузке на систему подтверждающей диагностики.

Клинические доказательства как основа оценки МИ ИВД

Поэтому объем доказательств должен соответствовать роли результата. Тест для ориентировочного лабораторного использования, скрининговая система для донорской крови, молекулярный тест для выбора таргетной терапии и программный алгоритм, автоматически относящий образец к категории риска, не требуют одинакового доказательного объема. Но во всех случаях действует одна и та же логика: результат должен быть научно обоснован, аналитически надежен и клинически пригоден для заявленного применения.

Эта международная логика имеет прямое значение для евразийского и российского регулирования. Нормативные акты и установленные формы определяют, как должны быть представлены программа испытаний, результаты исследований, регистрационное досье и материалы по клиническим доказательствам. Однако наличие формы не означает авто-

матической доказательности. Форма задает структуру, а убедительность возникает тогда, когда в ней последовательно раскрыта связь между назначением изделия, клинической ролью результата, использованными источниками данных, проведенными исследованиями и итоговым выводом.

Поэтому при подготовке программы клинико-лабораторных испытаний нельзя начинать только с перечня процедур. Сначала необходимо понять, что именно заявлено в назначении изделия. Если изделие предназначено для скрининга, программа должна учитывать скрининговую популяцию, маршрут первично положительного результата и подтверждающий этап. Если изделие предназначено для мониторинга, программа должна учитывать динамику результата во времени. Если изделие включает программный алгоритм, необходимо понимать, является ли программа только техническим помощником или она формирует клинически значимый вывод. Эти вопросы вытекают из международной доказательной логики, но практически реализуются уже в структуре конкретных документов, представляемых при оценке соответствия.

Международная логика доказательной оценки.

Иначе говоря, международная система регулирования медицинских изделий для диагностики *in vitro* важна не только как внешняя нормативная среда. Она задает способ мышления о доказательствах. Изделие оценивается не как технология в отрыве от медицины, а как источник результата,

способного повлиять на клиническое, лабораторное или организационное решение. Эта логика будет проходить через все последующие разделы издания: от научной обоснованности анализа и аналитической эффективности до клинико-лабораторных испытаний, программных алгоритмов, регистрационного досье и пострегистрационного наблюдения.

Итог главы

Международная система регулирования изделий для диагностики *in vitro* исходит из того, что лабораторный результат должен быть технически получен, а также медицински значим, надежен и пригоден для заявленного применения. Поэтому доказательная оценка строится вокруг связи между назначением изделия, клиническим смыслом анализа, аналитической надежностью конкретного изделия и клинической пригодностью результата. Именно эта логика лежит в основе современной оценки медицинских изделий для диагностики *in vitro*.

Глава 2. Международная методологическая рамка и ее отражение в регулировании МИ ИВД

Клинические доказательства формируются на нескольких уровнях.

Клинические доказательства для медицинских изделий для диагностики *in vitro* формируются не только как результат отдельных испытаний. Они являются частью более широкой методологической логики, которая сложилась в международной практике регулирования медицинских изделий и затем была воспринята региональными и национальными системами регулирования. Для понимания этой логики важно сначала рассмотреть, почему появились международные инициативы по гармонизации регулирования, какую роль сыграли GHTF и IMDRF, а затем показать, как эти подходы отражаются в регулировании ЕАЭС и Российской Федерации.

Во второй половине XX века медицинские изделия активно развивались технологически и быстро выходили на международные рынки. При этом разные страны использовали собственные подходы к регулированию: отличались правила

классификации, требования к системе качества, состав регистрационных материалов, подходы к клинической оценке, пострегистрационному наблюдению и контролю изменений. Для производителей это означало необходимость готовить сходные по смыслу, но разные по форме материалы для разных рынков. Для регуляторов — трудности при сопоставлении подходов и обмене опытом.

Международный уровень: методологический язык доказательств.

На этом фоне была создана GHTF — Global Harmonization Task Force, Международная целевая группа по гармонизации регулирования медицинских изделий. Ее задача состояла не в создании единого мирового регулятора, а в сближении принципов регулирования. GHTF объединяла представителей регуляторных органов и регулируемой отрасли и разрабатывала документы, посвященные общим подходам к классификации медицинских изделий, оценке соответствия, системе менеджмента качества, клинической оценке, надзору после вывода изделия на рынок и действиям при выявлении проблем безопасности.

Значение GHTF состояло в том, что она предложила общий язык регулирования медицинских изделий. В ее документах закреплялась идея, что объем доказательств и глубина оценки должны зависеть от назначения изделия, риска, характера результата и возможных последствий ошибки. Для медицинских изделий для диагностики *in vitro* это особенно

важно, поскольку риск таких изделий часто связан не с непосредственным воздействием на пациента, а с последствиями ошибочного результата: пропуском заболевания, необоснованным лечением, неверным выбором терапии, неправильной маршрутизацией пациента, ошибочным скрининговым решением или задержкой медицинской помощи.

Позднее развитие этой работы продолжил IMDRF — International Medical Device Regulators Forum, Международный форум регуляторов медицинских изделий. IMDRF был создан как площадка взаимодействия регуляторных органов, развивающая подходы, ранее заложенные GHTF. IMDRF ориентирован прежде всего на взаимодействие самих регуляторов и на современные задачи регулирования: программное обеспечение как медицинское изделие, искусственный интеллект, клинические доказательства, уникальную идентификацию изделий, пострегистрационное наблюдение, управление изменениями, кибербезопасность и международное сближение требований.

Документы GHTF и IMDRF не являются регистрационными правилами конкретной страны или региона. Они не заменяют правовые акты ЕАЭС или Российской Федерации и сами по себе не дают разрешения на обращение изделия. Их роль методологическая: они помогают регуляторам, производителям, испытательным организациям и экспертам одинаково понимать, какие данные нужны для оценки изделия, как связать доказательства с назначением, как учитывать

риск и как поддерживать доказательную базу в течение жизненного цикла изделия.

В настоящем издании необходимо различать методологический уровень доказательной оценки и юридически обязательные требования конкретной регуляторной системы. Понятия и модели GHTF, IMDRF, международных стандартов и профессиональных руководств используются для объяснения логики формирования и оценки доказательств. Юридические обязанности изготовителя, заявителя, испытательной организации и иных участников определяются применимыми нормативными правовыми актами ЕАЭС или Российской Федерации. Поэтому используемые далее формулировки о необходимости определенных доказательств следует понимать с учетом их источника. В одних случаях соответствующее требование непосредственно установлено нормативным актом, в других оно представляет собой методологический принцип, позволяющий обосновать достаточность и применимость доказательной базы.

Для МИ ИВД международная методологическая рамка особенно важна потому, что такие изделия оцениваются не только как лабораторная технология. Изделие может корректно выявлять аналит, регистрировать сигнал или формировать числовой результат, но этого недостаточно для вывода о медицинской пригодности. Необходимо понять, какое медицинское значение имеет результат, насколько надежно он формируется конкретным изделием и может ли исполь-

зоваться в том клиническом сценарии, который заявлен изготовителем.

Как было указано в главе 1, доказательную оценку МИ ИВД можно представить через три взаимосвязанных элемента: научную обоснованность, аналитическую эффективность и клиническую эффективность. Эта модель далее используется как методологическая рамка издания. Она позволяет распределять различные источники данных по их доказательной функции и одновременно сопоставлять международные подходы с конкретными требованиями ЕАЭС и Российской Федерации.



Рисунок 2.1. Уровни формирования клинических доказательств

Евразийский уровень: нормативная форма доказательств.

В евразийском регулировании эта логика отражается через несколько ключевых блоков. Решение Совета ЕЭК № 29 задает рамку проведения клинических и клинико-лабораторных испытаний медицинских изделий. Для МИ ИВД программа клинико-лабораторных испытаний выполняет не только организационную функцию. С доказательной точки зрения она представляет собой перевод назначения изделия в исследовательский дизайн: заявленное применение определяет выбор клинического материала, исследуемой популяции, метода сравнения или референтной рамки, оцениваемых показателей, статистического подхода и границ последующего вывода.

Решение Совета ЕЭК № 46 связано с регистрацией и экспертизой медицинских изделий. В контексте клинических доказательств его значение состоит в том, что досье должно представить доказательную базу как согласованную систему. Эксперт оценивает как наличие отдельных документов, так и то, насколько они подтверждают заявленное назначение изделия. Если в досье имеются литературные данные, аналитические отчеты и результаты клинико-лабораторных испытаний, но они относятся к разным материалам, разным по-

пуляциям или разным версиям изделия, доказательная база будет восприниматься как разрозненная.

Решение Совета ЕЭК № 27 задает общие требования безопасности и эффективности. Для изделий диагностики *in vitro* эти требования должны пониматься через результат изделия и последствия его применения. Безопасность и эффективность здесь не сводятся только к отсутствию физического вреда от изделия. Они включают вопрос о том, насколько результат может быть использован без создания неприемлемого риска ошибочного медицинского решения. Иначе говоря, ошибочный лабораторный результат сам становится источником риска, который должен быть учтен при формировании доказательств.

Методические рекомендации ЕАЭС по клиническим доказательствам и установленные формы документов важны потому, что переводят эту общую логику в практическую структуру. Однако форма сама по себе не гарантирует качества. Один и тот же раздел может быть заполнен формально или доказательно. Формальное заполнение перечисляет сведения. Доказательное заполнение показывает связь: что заявлено в назначении, какие данные это подтверждают, какие ограничения остаются и почему итоговый вывод не выходит за пределы фактически полученных или обоснованно применимых данных.

Именно здесь возникает практическая задача для испытательной организации и разработчика. Необходимо не под-

гонять исследование под форму, а выстроить исследование так, чтобы форма отражала реальную доказательную логику. Если изделие заявлено для определенного клинического материала, программа должна включать этот материал или обосновывать использование иных источников данных. Если изделие заявлено для определенной популяции, данные должны быть применимы к этой популяции. Если результат используется для выбора терапии, доказательства должны быть связаны не только с обнаружением маркера, но и с его терапевтическим значением.

Российский уровень: национальная процедура оценки соответствия.

В Российской Федерации национальная процедура государственной регистрации медицинских изделий определяется Правилами государственной регистрации медицинских изделий, утвержденными постановлением Правительства Российской Федерации от 30 ноября 2024 г. № 1684. Правила устанавливают порядок представления регистрационных материалов, экспертизы качества, эффективности и безопасности медицинского изделия, проведения регистрационных процедур и внесения изменений в документы регистрационного досье. Требования к проведению соответствующих видов испытаний конкретизируются нормативными актами Министерства здравоохранения Российской Федерации. На момент подготовки настоящего издания происходит обновление нормативной основы этой процедуры: с 1

сентября 2026 года вступает в силу приказ Минздрава России от 13 мая 2026 г. № 421н, устанавливающий новый порядок проведения оценки соответствия медицинских изделий в форме технических испытаний, токсикологических исследований и клинических испытаний в целях государственной регистрации. Одновременно утрачивает силу ранее действовавший приказ Минздрава России от 30 августа 2021 г. № 885н. Для оценки соответствия, заявления о проведении которой были представлены до вступления нового порядка в силу, предусмотрены переходные положения.

Для медицинских изделий для диагностики *in vitro* национальная процедура определяет установленный порядок проведения испытаний и оформления их результатов. Однако нормативная форма документа не заменяет доказательной логики исследования. Программа, первичные данные и итоговые документы должны позволять проследить связь между назначением изделия, исследованным клиническим материалом, популяцией, методом сравнения, полученными результатами и сделанным выводом.

Критически важно различать наличие данных и наличие доказательств. Данные приобретают доказательное значение только в отношении конкретного вопроса, для решения которого они были получены или применимость к которому была обоснована. Публикация может подтверждать связь анализа с заболеванием, но не работу конкретного изделия. Аналитическое исследование может подтверждать ха-

рактеристики метода, но не заменяет оценку клинической эффективности. Клинико-лабораторное испытание, в свою очередь, подтверждает только ту область применения, которая фактически охвачена его дизайном и полученными данными.

Поэтому клинические доказательства нужно рассматривать как линию рассуждения, а не как механическую сумму документов. Эта линия начинается с назначения изделия. Затем показывается клинический смысл анализа или объекта исследования. Далее подтверждается, что конкретное изделие надежно получает результат. После этого доказывається, что результат пригоден для заявленной медицинской задачи. И только в конце формулируется вывод о достаточности доказательств, который должен быть не шире фактически подтвержденного применения.

Такой подход особенно важен для сложных изделий: молекулярных панелей, систем высокопроизводительного секвенирования, онкологических companion-тестов, изделий для неонатального скрининга, микробиологических тестов со сложной референтной рамкой и программных алгоритмов. В этих случаях доказательная база почти всегда формируется из разных источников. Часть сведений может быть взята из литературы, часть — из аналитических исследований, часть — из клинико-лабораторных испытаний, часть — из архивного материала, специально сформированных панелей или данных реальной практики. Международная,

евразийская и российская доказательная логика помогает распределить эти источники по их доказательной функции и не смешивать разные уровни подтверждения.

Например, изделие может быть предназначено для выявления инфекционного агента у пациентов с симптомами заболевания. В этом случае доказательная задача связана с подтверждением результата в соответствующей клинической популяции и заявленном материале. Если тот же тест заявляется для скрининга бессимптомных лиц или доноров, доказательная задача меняется. Важнее становятся специфичность, ложноположительная нагрузка, подтверждающий маршрут и последствия пропуска случая. Анализ может быть тем же, но клиническая роль результата будет иной.

Аналогичная ситуация возникает в онкологии. Наличие публикаций о связи мутации с опухолевым процессом может подтверждать научную обоснованность, но не доказывает автоматически пригодность конкретного изделия для выбора терапии. Для такого назначения необходимо показать, что изделие надежно выявляет соответствующее изменение в заявленном материале, что результат интерпретируется в нужном клиническом контексте и что эта находка действительно имеет значение для терапевтического решения.

Клинические доказательства связаны с жизненным циклом изделия. При первичной регистрации они обосновывают возможность обращения изделия в заявленной области применения. После регистрации они используются при

оценке изменений, обновлении инструкции, анализе жалоб, пострегистрационном контроле и накоплении новых данных. Если меняется тип материала, целевая популяция, алгоритм интерпретации, программная версия, порог, база знаний или клинический контекст применения, необходимо оценить, сохраняется ли применимость ранее представленных доказательств.

Это не означает, что при каждом изменении требуется новое клиническое испытание. Объем дополнительной оценки зависит от характера изменения и его влияния на результат. Иногда достаточно документированного обоснования или верификации. Иногда требуется дополнительный анализ имеющихся данных. Иногда необходимо новое клинико-лабораторное исследование. Главный принцип остается прежним: доказательства должны соответствовать фактическому изделию и его заявленному применению.

В итоге международная методология, евразийское регулирование и российская национальная процедура следует рассматривать как разные уровни одной практической задачи. Международные документы дают понятийную и методологическую рамку доказательной оценки. Правила ЕАЭС устанавливают требования к обращению, регистрации, экспертизе и клинико-лабораторным испытаниям в рамках Союза. Российское законодательство определяет национальную процедуру государственной регистрации и оценки соответствия в пределах ее применения. Эти уровни различаются

по юридическому статусу и сфере действия, но при оценке МИ ИВД сходятся в главном практическом вопросе: позволяют ли представленные данные обоснованно подтвердить безопасность, качество и эффективность конкретного изделия в пределах заявленного назначения.

Итог главы

Международная методологическая рамка, сформированная GHTF, IMDRF, международными стандартами и руководствами, задает общий язык доказательной оценки медицинских изделий для диагностики *in vitro*. GHTF заложила базовые подходы к гармонизации регулирования медицинских изделий, а IMDRF продолжил эту работу с акцентом на современные регуляторные задачи. В евразийской практике эта логика отражается в Решениях Совета ЕЭК № 29, № 46 и № 27, методических рекомендациях и формах документов. В Российской Федерации она реализуется через национальную процедуру регистрации, проведения испытаний, оформления результатов и экспертизы. Доказательная база МИ ИВД должна подтверждать безопасность, качество и эффективность изделия в пределах заявленного назначения и оставаться согласованной с фактическим изделием, инструкцией, изменениями и данными применения в течение жизненного цикла.

Глава 3. Назначение изделия и его клиническая роль

Назначение определяет доказательства.

Назначение изделия является исходной точкой доказательной оценки медицинского изделия для диагностики *in vitro*. Именно из него становится понятно, какой результат должно выдавать изделие, в каком материале этот результат получают, для какой группы пациентов или обследуемых лиц он предназначен, кто будет использовать изделие и какое медицинское решение может быть принято на основании результата. Поэтому назначение нельзя воспринимать как краткую формальную фразу из инструкции. В диагностике *in vitro* оно фактически задает границы всей доказательной базы.



Рисунок 3.1. Назначение изделия как центр доказательной оценки

Одна и та же технология может иметь совершенно разное доказательное значение в зависимости от того, какую клиническую роль ей придает изготовитель. Например, молекулярный тест для выявления инфекционного агента может использоваться для диагностики острой инфекции у пациента с симптомами, для скрининга доноров, для подтверждения результата другого метода, для мониторинга терапии или для эпидемиологического наблюдения. Во всех случа-

ях лабораторная мишень может быть одной и той же, но доказательная задача будет разной. Для диагностики у симптомного пациента важно показать пригодность результата в соответствующей клинической ситуации. Для скрининга доноров особенно значимы высокая чувствительность, специфичность и последствия пропуска случая. Для мониторинга терапии важна как детекция, так и воспроизводимость динамики результата во времени.

Именно поэтому назначение изделия связывает техническое описание изделия с медицинской задачей. Если в документации указано только, что изделие “предназначено для выявления” определенного аналита, этого часто недостаточно для доказательной оценки. Необходимо понимать, зачем этот аналит выявляется: для постановки диагноза, исключения заболевания, скрининга, мониторинга, прогноза, стратификации риска, выбора терапии или другой клинической функции. Пока эта роль не определена, невозможно корректно выбрать клинический материал, популяцию, метод сравнения, конечные точки и статистический план.

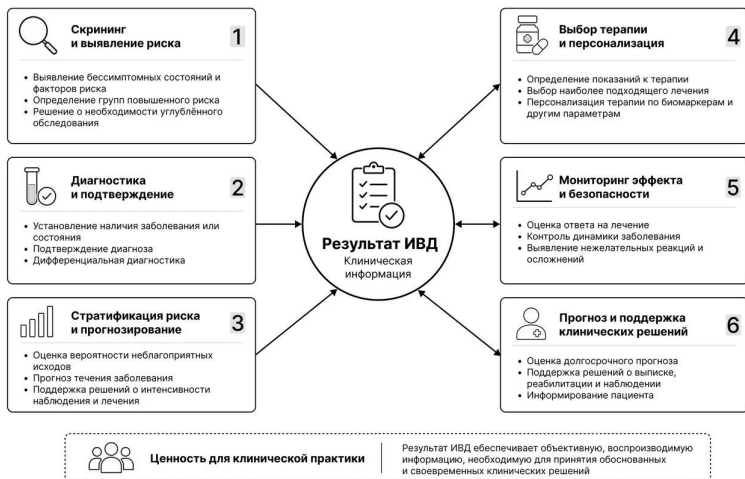


Рисунок 3.2. Клинические роли результата МИ ИВД

Не всякое применение уже доказано.

Для доказательной оценки назначение изделия должно быть сформулировано с достаточной конкретностью, позволяющей определить границы заявленного применения. Для МИ ИВД недостаточно учитывать только аналит и принцип метода. Существенное значение имеют тип исследуемого образца, целевая популяция, предполагаемый пользователь, условия применения, формат результата и его клиническая роль. Например, изделие для определения глюкозы в

венозной плазме в условиях клиничко-диагностической лаборатории и изделие для самостоятельного контроля глюкозы пациентом дома решают разные задачи, даже если оба связаны с одним аналитом. В первом случае основное внимание будет уделяться лабораторной надежности результата и его соответствию клиническому применению в медицинской организации. Во втором — добавляются вопросы пользователя, условий самостоятельного применения, интерпретации результата и риска неправильного действия пациента.

То же относится к онкологическим изделиям. Тест, выявляющий изменение в гене EGFR, может использоваться для молекулярной характеристики опухоли, определения клинически значимого варианта, поддержки выбора таргетной терапии, выявления механизмов резистентности или иной заявленной задачи. Эти применения нельзя считать доказательно эквивалентными только потому, что определяется один и тот же ген.

Если результат предназначен для использования при выборе терапии, доказательная база должна подтверждать его заявленную предиктивную роль в соответствующем клиническом контексте. Данных о том, что определенный вариант встречается при соответствующем типе опухоли, для этого недостаточно. Необходимо обосновать клиническое значение выявляемого варианта для соответствующего терапевтического решения и подтвердить способность конкретного изделия надежно определять этот вариант в заявленном ма-

териале и популяции. Такая связь между результатом МИ ИВД и терапевтическим решением становится частью самого назначения изделия и, следовательно, должна учитываться при формировании доказательной базы и дизайна соответствующих исследований.

Клиническая роль результата особенно важна потому, что она определяет цену ошибки. Ложноотрицательный результат скринингового теста для тяжелого состояния может привести к пропуску пациента, которому требуется срочное вмешательство. Ложноположительный результат в массовом скрининге может создать значительную нагрузку на систему подтверждающей диагностики и привести к необоснованной тревоге. Ошибка теста для выбора терапии может изменить лечебный маршрут. Ошибка мониторингового теста может привести к неверной оценке динамики заболевания. Поэтому одно и то же значение чувствительности, специфичности или воспроизводимости может оцениваться по-разному в зависимости от назначения изделия.

Кто, где и на каком материале использует изделие.

Назначение также определяет, какие доказательства могут быть перенесены, а какие должны быть получены заново. Если анализ хорошо изучен в одной клинической роли, это не означает, что он автоматически обоснован для другой. Связь маркера с заболеванием может быть достаточной для диагностики, но недостаточной для прогноза. Данные по пациентам с уже установленным заболеванием могут не под-

ходить для скрининга бессимптомной популяции. Результаты, полученные на сыворотке, не всегда можно переносить на плазму, цельную кровь, сухое пятно крови, тканевой материал или цифровое изображение. Каждый такой перенос требует обоснования.

В этом смысле назначение изделия является описанием предполагаемого применения, а также инструментом ограничения доказательных выводов. Если испытание проведено только на одном типе материала, итоговый вывод не должен без дополнительного обоснования распространяться на другие материалы. Если исследована только определенная популяция, нельзя автоматически заявлять более широкий круг пользователей или пациентов. Если использовалась одна версия программного алгоритма, доказательства должны быть связаны именно с этой версией или должно быть показано, что изменения не влияют на клинически значимый результат.

Особенно наглядно это проявляется в изделиях с программным или алгоритмическим компонентом. Если программа только передает результат в лабораторную информационную систему или печатает отчет, ее роль вспомогательная. Но если она фильтрует варианты секвенирования, рассчитывает индекс риска, классифицирует изображение, выделяет положительные клетки или автоматически относит образец к категории “норма”, “повторить” или “положительный”, ее функция становится частью формирования

клинически значимого результата. В таком случае назначение должно отражать лабораторный этап, а также роль программного компонента в формировании вывода. Иначе доказательная оценка окажется неполной.

Для программы клинико-лабораторных испытаний назначение изделия выполняет роль исходного технического задания. Именно из него вытекает, какие образцы нужно включить, какие группы пациентов должны быть представлены, какой метод сравнения или референтная рамка будет уместной, какие показатели нужно рассчитывать и какие выводы можно будет сделать. Если назначение сформулировано расплывчато, программа испытаний неизбежно становится уязвимой. Она может проверить не то применение, которое фактически заявлено, или получить данные, которые нельзя уверенно связать с итоговым выводом.

Для регистрационного досье значение назначения еще шире. Досье должно показывать, что все элементы доказательной базы относятся к одному и тому же заявленному применению. Литературные данные должны подтверждать клинический смысл анализа именно в этой роли. Аналитические исследования должны быть выполнены на релевантном материале и в условиях, соответствующих применению изделия. Клинические или клинико-лабораторные испытания должны охватывать заявленную популяцию и клинический сценарий. Итоговый вывод должен быть не шире, чем фактически подтвержденные данные.

Границы назначения — это границы вывода

Поэтому чрезмерно широкое назначение изделия часто становится источником методологических проблем. Производитель может стремиться описать изделие максимально широко, чтобы расширить область применения. Но чем шире назначение, тем больше доказательная нагрузка. Если изделие заявлено для нескольких материалов, нескольких популяций, нескольких клинических ролей или разных условий применения, каждая из этих составляющих должна быть подтверждена либо непосредственно, либо через обоснованный перенос данных. В противном случае возникает разрыв между заявленным применением и доказательной базой.

Более сильной стратегией часто является точное и доказательно устойчивое назначение. Оно может быть уже по объему, но лучше поддержано данными. Такое назначение легче перевести в программу испытаний, легче связать с клиническим материалом и популяцией, легче защитить в итоговом отчете и регистрационном досье. В дальнейшем, при накоплении новых данных, область применения может быть расширена, но такое расширение должно сопровождаться соответствующим доказательным обоснованием.

Иначе говоря, назначение изделия в диагностике *in vitro* играет роль доказательного центра. Оно связывает аналит, метод, материал, пользователя, популяцию, клинический сценарий, формат результата и медицинское действие. От него зависит, какие данные будут считаться релевантными,

какие исследования необходимо провести, какие ограничения должны быть указаны и какой итоговый вывод может быть сделан. Поэтому работа над доказательной базой начинается не с выбора таблицы для отчета и не с расчета чувствительности, а с точного понимания того, для чего именно предназначено изделие и какую роль его результат должен играть в медицинской практике.

Итог главы

Назначение изделия определяет границы всей доказательной оценки медицинского изделия для диагностики *in vitro*. Оно показывает, какой результат получает изделие, в каком материале, для какой популяции, каким пользователем, в каких условиях и для какого медицинского решения. Одна и та же технология может требовать разных доказательств в зависимости от клинической роли результата. Поэтому программа испытаний, регистрационное досье и итоговый вывод о клинических доказательствах должны строиться от назначения изделия и не выходить за пределы фактически подтвержденного применения.

Глава 4. Доказательная модель оценки медицинских изделий для диагностики *in vitro*

Доказательства должны складываться в систему.

Доказательная оценка медицинского изделия для диагностики *in vitro* строится не вокруг одного исследования и не вокруг одного числового показателя. Она представляет собой последовательное обоснование того, что результат изделия имеет медицинский смысл, получается надежно и может использоваться в заявленном клиническом сценарии. Поэтому в центре оценки находится не только технология, а связь между назначением изделия, определяемым объектом, аналитической работой изделия и клиническим значением результата.

В практике нередко возникает соблазн упростить эту логику. Если анализ давно известен и включен в клинические рекомендации, кажется, что изделие уже достаточно обосновано. Если метод показывает хорошую воспроизводимость и низкий предел обнаружения, может возникнуть впечатление, что его клиническая пригодность очевидна. Если в испытании получены высокие чувствительность и специфичность, иногда считают, что все вопросы доказательной базы закрыты. Однако для изделий диагностики *in vitro* та-

кие упрощения опасны. Каждый из этих элементов отвечает только на свой вопрос и не заменяет остальные.

Три вопроса доказательной оценки.

Именно поэтому современная доказательная модель МИ ИВД обычно рассматривается через три взаимосвязанных блока: научную обоснованность, аналитическую эффективность и клиническую эффективность. Они образуют единую цепочку, но каждый блок имеет самостоятельное значение.



Рисунок 4.1. Три элемента доказательной модели

Первый элемент — научная обоснованность — относит-

ся к смыслу самого объекта исследования. Если изделие определяет вирусный антиген, генетический вариант, концентрацию биохимического показателя, клеточную характеристику, морфологический признак или вычисляемый индекс, необходимо показать, что этот объект связан с заболеванием, физиологическим или патологическим процессом, риском состояния, прогнозом, ответом на лечение или иной медицинской задачей. Без такой связи результат остается лабораторным фактом, но не становится клиническим доказательством.

Например, можно разработать очень точный метод измерения нового белкового маркера. Но если не показано, что этот маркер устойчиво связан с конкретным заболеванием или клиническим решением, аналитическая точность метода не создает медицинской ценности. В этом смысле научная обоснованность предшествует всем последующим действиям. Она отвечает на вопрос: зачем вообще определять этот объект у пациента или в образце.

Второй элемент — аналитическая эффективность — относится уже не к самому аналиту как клиническому явлению, а к конкретному изделию. Даже если аналит хорошо известен и давно используется в медицине, каждое новое изделие должно показать, что оно надежно получает результат. Для количественного теста это может включать правильность, прецизионность, диапазон измерений, предел количественного определения и прослеживаемость. Для качествен-

ного теста особенно важны предел обнаружения, специфичность, устойчивость вблизи порога и поведение слабopоложительных образцов. Для молекулярного теста — способность выявлять заявленные варианты, устойчивость к матриксным влияниям, интерференции и риск ложного результата. Для программного изделия — корректность вычислительной обработки, версия алгоритма и устойчивость классификации.

Здесь важно не смешивать клиническую значимость анализа с работой конкретного изделия. То, что вирусная нагрузка имеет значение для мониторинга инфекции, еще не доказывает, что конкретная тест-система правильно измеряет ее во всем заявленном диапазоне. То, что мутация имеет значение для выбора терапии, еще не доказывает, что конкретный молекулярный тест надежно обнаруживает ее в заявленном типе опухолевого материала. Аналитическая эффективность должна быть подтверждена именно для оцениваемого изделия и именно в тех условиях, в которых оно будет применяться.

Третий элемент — клиническая эффективность (clinical performance) — связывает результат конкретного изделия с заявленным клиническим состоянием, физиологическим или патологическим процессом либо иной медицински значимой характеристикой. На этом уровне оценивается уже не только способность изделия получить аналитический сигнал, категорию или числовое значение, а способность ре-

результата выполнять заявленную клиническую функцию.

Конкретное содержание клинической эффективности определяется назначением изделия. Для диагностического теста это может быть способность правильно классифицировать наличие или отсутствие определенного состояния. Для скринингового изделия — способность выявлять лиц, требующих дальнейшего обследования, с учетом характеристик целевой популяции и предусмотренного подтверждающего маршрута. Для мониторингового изделия — способность результата надежно отражать клинически значимое изменение во времени. Для изделия, результат которого используется при выборе терапии, клиническая эффективность должна оцениваться применительно к заявленной предиктивной роли соответствующего результата.

Клиническая эффективность и клиническая польза

Клиническую эффективность (*clinical performance*) следует отличать от понятия клинической пользы (*clinical utility, clinical benefit*). Клиническая эффективность характеризует способность результата изделия быть связанным с определенным клиническим состоянием, физиологическим или патологическим процессом либо иной медицинской характеристикой в соответствии с назначением изделия. Клиническая польза относится к более широкому вопросу о последствиях использования полученной диагностической информации в медицинской практике: влияет ли результат на последующее обследование, выбор или изменение лечения,

маршрутизацию пациента либо иной элемент медицинской помощи и какое значение имеет такое использование для пациента или системы здравоохранения.

Эти понятия взаимосвязаны, но не тождественны. Доказательство клинической пользы не следует рассматривать как универсальный самостоятельный четвертый этап оценки эффективности каждого медицинского изделия для диагностики *in vitro*. Необходимость и характер ее оценки определяются назначением изделия, клинической ролью результата и применимой регуляторной рамкой.

Эти три элемента связаны между собой через назначение изделия. Назначение определяет, какой анализит или объект исследования оценивается, в каком материале, у какой популяции, каким пользователем и для какого клинического решения. Поэтому доказательная модель не является абстрактной схемой. Она всегда применяется к конкретному изделию и конкретному заявленному применению.

На практике доказательная модель помогает увидеть разрывы, которые иначе могут остаться скрытыми. Например, в досье могут быть хорошие публикации о клиническом значении маркера и убедительные аналитические данные по изделию, но отсутствовать клинические данные, относящиеся к заявленной популяции. В другом случае клинико-лабораторное испытание может быть проведено качественно, но только на одном типе материала, тогда как инструкция заяв-

ляет несколько материалов. В третьем случае изделие может быть испытано в одной версии программного обеспечения, а в регистрационных материалах фигурирует уже другая версия алгоритма, способная изменить классификацию результата.

Один тип доказательств не заменяет другой.

Такие ситуации не всегда означают, что изделие непригодно. Но они показывают, что доказательная линия неполна или требует ограничения. В сильной доказательной базе каждый блок выполняет свою функцию. Литература и рекомендации могут подтверждать научную обоснованность. Аналитические исследования подтверждают способность конкретного изделия получать результат. Клинические и клинико-лабораторные испытания показывают пригодность результата в заявленном клиническом сценарии. Данные реальной практики могут поддерживать устойчивость выводов после внедрения. Но эти источники нельзя просто сложить в один массив. Нужно показать, какой вопрос закрывает каждый источник и как все они вместе поддерживают назначение изделия.

Схема 4.1 — Что доказывает каждый элемент доказательной модели

— **Научная обоснованность** — *Что подтверждает:* Клинический смысл анализа или объекта исследования;
Чем не является: Не доказывает работу конкретного изделия.

— **Аналитическая эффективность** — **Что подтверждает:** Надежность получения результата конкретным изделием; **Чем не является:** Не доказывает автоматически клиническую пригодность результата.

— **Клиническая эффективность** — **Что подтверждает:** Способность результата конкретного изделия отражать заявленное клиническое состояние или иную медицински значимую характеристику и выполнять предусмотренную назначением клиническую функцию; **Чем не является:** Не заменяет научную обоснованность и аналитические данные, е тождественна клинической пользе.

— **Данные реальной практики** — **Что подтверждает:** Поведение изделия после внедрения и устойчивость выводов; **Чем не является:** Не исправляют слабый исходный дизайн испытаний.

— **Итоговый отчет по клиническим доказательствам** — **Что подтверждает:** Связь всех доказательных блоков с назначением изделия; **Чем не является:** Не должен расширять выводы за пределы данных.

Объем и характер клинических доказательств зависят от назначения изделия, клинической роли результата и связанных с ошибочным результатом рисков. Для изделия с ограниченной вспомогательной ролью достаточный объем данных может быть относительно небольшим. Для изделия, результат которого используется при выборе терапии, скри-

нинге тяжелого состояния, обеспечении безопасности донорского материала или принятии иного критически значимого решения, доказательная нагрузка, как правило, существенно выше.

Вывод не может быть шире доказательств.

При этом модель не требует одинакового набора исследований для всех изделий. Она требует другого: чтобы объем и характер доказательств соответствовали назначению и риску. Если доказательства получены косвенно, нужно обосновать их применимость. Если использованы архивные образцы или специальные панели, нужно показать их роль и ограничения. Если часть заявленного назначения не покрыта данными, это должно быть отражено в выводе, инструкции или плане дополнительных исследований.

Итогом доказательной оценки должен быть не перечень положительных результатов, а обоснованный вывод о достаточности клинических доказательств. Такой вывод должен быть привязан к назначению изделия: для какого анализа, материала, популяции, пользователя, условий применения, версии изделия и клинической роли результат можно считать подтвержденным. Если эта привязка отсутствует, даже большой объем данных может не образовывать убедительной доказательной базы.

Итог главы

Доказательная модель оценки медицинского изделия для диагностики *in vitro* строится на трех взаимосвязанных эле-

ментах: научной обоснованности, аналитической эффективности и клинической эффективности. Эти элементы отвечают на разные вопросы и не заменяют друг друга. Научная обоснованность показывает, почему определяемый аналит или иной объект исследования имеет медицинское значение. Аналитическая эффективность подтверждает способность конкретного изделия надежно получать заявленный результат. Клиническая эффективность характеризует связь результата конкретного изделия с заявленным клиническим состоянием или иной медицински значимой характеристикой и его способность выполнять предусмотренную назначением клиническую функцию. Сильная доказательная база возникает тогда, когда все три элемента согласованы с назначением изделия и поддерживают вывод в пределах фактически подтвержденного применения.

Глава 5. Жизненный цикл медицинского изделия для диагностики *in vitro*

Доказательства не заканчиваются регистрацией.

Жизненный цикл медицинского изделия для диагностики *in vitro* охватывает весь путь изделия: от первоначальной идеи и формулирования назначения до разработки, испытаний, регистрации, внедрения в практику, пострегистрационного наблюдения, внесения изменений и прекращения обращения. Для доказательной оценки это имеет принципиальное значение. Клинические доказательства не возникают одномоментно в конце разработки и не исчерпываются одним отчетом. Они формируются, уточняются, поддерживаются и при необходимости пересматриваются на протяжении всего периода существования изделия.

На раннем этапе жизненного цикла ключевым решением становится формулирование назначения изделия. Именно здесь определяется, какой анализит или иной объект исследования будет определяться, в каком клиническом материале, для какой популяции, каким пользователем, в каких условиях и для какого медицинского решения. Если назначение сформулировано слишком широко или неопределенно, последующая доказательная работа становится уязвимой. Про-

грамма испытаний может не охватить все заявленные сценарии, аналитические характеристики могут быть выбраны без учета реального риска, а клинические данные могут оказаться неприменимыми к тому использованию, которое указано в документации.

Эта проблема часто проявляется не сразу. На этапе разработки широкая формулировка назначения может казаться удобной: она оставляет изделию больше возможностей для применения. Но при подготовке программы испытаний становится ясно, что каждое расширение назначения требует доказательств. Если изделие заявлено для нескольких типов материала, нужно показать применимость к каждому из них или обосновать перенос данных. Если изделие предназначено для разных клинических ролей, например для диагностики и мониторинга, доказательная база должна поддерживать обе роли. Если изделие включает программный модуль, влияющий на результат, его версия и функция также становятся частью доказательной идентичности изделия.

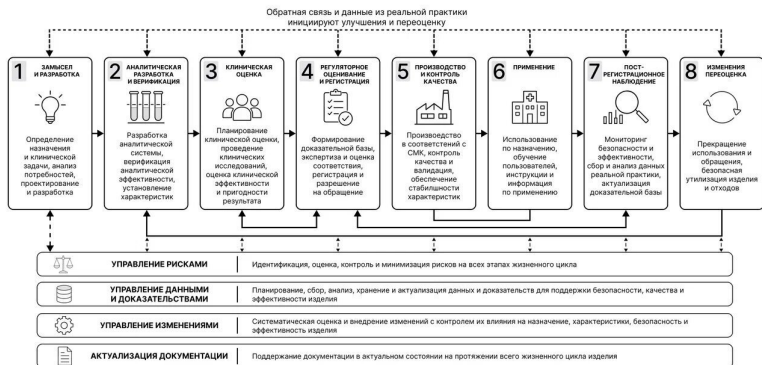


Рисунок 5.1. Жизненный цикл МИ ИВД

На каждом этапе нужны свои доказательства.

На этапе разработки формируется техническая и аналитическая основа изделия. Производитель выбирает принцип метода, реагенты, калибраторы, контрольные материалы, приборную часть, программное обеспечение, правила интерпретации результата и критерии пригодности клинического материала. Уже здесь будущая доказательная база начинает приобретать форму. Например, выбор порога для качественного теста определит, какие образцы должны быть включены в оценку пороговой зоны. Выбор диапазона измерений для количественного теста определит, какие уровни аналита должны быть подтверждены. Выбор алгоритма обработки данных определит, какие входные данные и какие

версии программного компонента должны быть испытаны.

Далее жизненный цикл переходит к подтверждению научной обоснованности. На этом этапе устанавливается, почему выбранный аналит или объект исследования имеет клинический смысл. Для хорошо изученного аналита эта часть доказательной базы может опираться на литературу, клинические рекомендации, систематические обзоры и устоявшуюся практику. Например, клиническое значение глюкозы при нарушениях углеводного обмена или HBsAg при вирусном гепатите В не требует полного доказательства заново. Но даже в таких случаях необходимо показать, что изделие использует известную связь именно в том назначении, которое заявлено. Если же речь идет о новом маркере, новом алгоритмическом индексе или новом клиническом применении известного аналита, научная обоснованность может потребовать отдельного анализа и дополнительных данных.

После этого подтверждается аналитическая эффективность. Здесь оценивается, способно ли конкретное изделие надежно получать результат. Для одного изделия критически важным будет предел обнаружения, для другого — правильность и прослеживаемость измерений, для третьего — устойчивость классификации около порога, для четвертого — интерференция, стабильность или воспроизводимость между сериями. В жизненном цикле аналитическая эффективность не должна пониматься как формальное испытание один раз и навсегда. Если позднее меняются реагенты, ка-

либраторы, приборная часть, алгоритм обработки или условия применения, необходимо оценить, сохраняют ли прежние аналитические данные свою применимость.

Клиническая эффективность подтверждает пригодность результата для заявленного медицинского применения. На этом этапе результат изделия соотносится с клиническим статусом, сравнительным методом, референтной рамкой, клиническим маршрутом или иной основой интерпретации. Для диагностического теста это может быть подтверждение способности различать наличие и отсутствие состояния. Для скринингового изделия — способность правильно выделять группу, требующую дальнейшего подтверждения. Для мониторингового изделия — пригодность результата для оценки динамики. Для изделия, влияющего на выбор терапии, — способность результата поддерживать соответствующее лечебное решение.

Регистрация и ввод изделия в обращение не завершают жизненный цикл доказательств. Они фиксируют, что на момент оценки представленные данные признаны достаточными для заявленного применения. Но после внедрения изделие начинает работать в более разнообразной среде, чем та, которая была представлена в испытаниях. Меняются лаборатории, пользователи, структура поступающих образцов, частота пограничных результатов, клинические маршруты, методы сравнения и иногда само понимание клинической роли анализа. Поэтому пострегистрационное наблюдение не яв-

ляется только административной обязанностью. Оно позволяет проверять, сохраняется ли доказательная база в реальных условиях применения.

Особенно отчетливо это видно на примере изделий для скрининга. В рамках испытаний может быть показано, что тест надежно работает на выбранной выборке. Но после внедрения в массовую программу появляются реальные вариации: неодинаковое качество образцов, разные сроки транспортировки, неоднородность популяции, изменение распространенности состояния, ошибки пользователя, повторные исследования, пограничные результаты. Эти данные важны для контроля качества, а также для понимания того, насколько исходная доказательная конструкция сохраняется в практике.

Еще более чувствительным становится жизненный цикл программных и алгоритмических изделий. В традиционной реагентной системе значимое изменение часто связано с составом реагента, калибратором, прибором или инструкцией. В программной системе клинически значимое изменение может возникнуть при обновлении версии алгоритма, изменении фильтрации данных, обновлении базы вариантов, корректировке правила классификации, изменении порога или формата отчета. Лабораторная часть изделия может остаться прежней, но итоговый клинический вывод изменится. Поэтому версия программного обеспечения должна рассматриваться не как техническая мелочь, а как элемент доказатель-

ной идентичности изделия.

Доказательная идентичность изделия — совокупность характеристик медицинского изделия для диагностики *in vitro*, определяющих применимость полученных для него доказательств. К ним относятся назначение изделия, определяемый анализ или иной объект исследования, состав и конфигурация изделия, тип клинического материала, целевая популяция, предполагаемый пользователь, условия применения, правила формирования и интерпретации результата, а для изделий с программным компонентом — также версия программного обеспечения, алгоритма, базы знаний и иные элементы, изменение которых способно повлиять на клинически значимый результат.

Изменение элемента доказательной идентичности не означает автоматически необходимости проведения нового клинико-лабораторного испытания. Оно означает необходимость документированно оценить, сохраняется ли применимость ранее полученных доказательств к измененному изделию.



Рисунок 5.2. Доказательная идентичность изделия

Если один из этих элементов существенно изменяется, необходимо оценить, сохраняется ли связь между ранее полученными доказательствами и фактически обращающимся изделием.

Изменение изделия может изменить доказательства.

Изменения в жизненном цикле могут быть разными по значимости. Одни изменения не влияют на клинически значимый результат и могут быть подтверждены техническим обоснованием. Другие затрагивают аналитическую работу изделия или интерпретацию результата и требуют дополни-

тельной проверки. Третьи фактически расширяют назначение изделия и потому требуют новой доказательной базы. Например, добавление нового типа материала, переход от лабораторного применения к применению у постели пациента, расширение популяции на бессимптомных лиц или изменение алгоритма классификации не должны автоматически считаться техническими уточнениями. Каждое из таких изменений может изменить риск и доказательные требования.

При изменении назначения особенно важно не переносить прежние доказательства механически. Если изделие было подтверждено для диагностики у пациентов с симптомами, это не означает, что оно автоматически пригодно для скрининга бессимптомной популяции. Если тест был испытан на сыворотке, это не подтверждает без дополнительного анализа применение в сухом пятне крови или тканевом материале. Если алгоритм был обучен и проверен на данных одной лаборатории, его перенос в другую среду может потребовать внешней проверки. Жизненный цикл тем самым постоянно возвращает оценку к исходному вопросу: относятся ли имеющиеся доказательства к текущему применению изделия?

Научное понимание анализа также может меняться. Для ряда маркеров со временем уточняется клиническая роль, появляются новые рекомендации, меняются пороговые значения, выявляются новые интерференты, пересматриваются критерии классификации вариантов или обновляются пред-

ставления о значимости биомаркера для терапии. В таких случаях даже неизменное изделие может оказаться в новой доказательной ситуации. Оно продолжает технически работать так же, но клинический смысл результата или правила его интерпретации могут измениться. Поэтому жизненный цикл включает контроль физического изделия, а также наблюдение за изменением научного и клинического контекста.

В рамках системы менеджмента качества, управления рисками и регуляторного сопровождения изготовитель должен обеспечивать документированную оценку изменений, способных повлиять на безопасность, эффективность, назначение или доказательную базу изделия. При такой оценке необходимо определить, затрагивает ли изменение аналитическую или клиническую эффективность, тип материала, целевую популяцию, правила интерпретации, программный компонент либо иной элемент доказательной идентичности изделия и какие дополнительные действия необходимы для подтверждения сохранения ранее сделанных выводов.

Первичная оценка характера и последствий изменения относится к ответственности изготовителя и должна быть обоснована и документирована. При этом сделанный изготовителем вывод о сохранении применимости ранее полученных доказательств может подлежать регуляторной или экспертной оценке в случаях, предусмотренных применимой процедурой. Поэтому существенность изменения определя-

ется не только его техническим масштабом, но прежде всего возможным влиянием на заявленный результат, риск его ошибочной интерпретации и границы ранее подтвержденного применения.

Реальное применение уточняет доказательную базу.

Не менее важно различать поддержание доказательств и расширение доказательств. Поддержание доказательств направлено на подтверждение того, что изделие продолжает работать в пределах уже заявленного применения. Расширение доказательств требуется тогда, когда изделие начинает претендовать на новую область: другой материал, другую популяцию, новую клиническую роль, новый алгоритмический вывод или новую среду применения. Эти ситуации нельзя смешивать. Поддержание может опираться на пострегистрационное наблюдение, контроль стабильности, анализ рекламаций и сопоставимость версий. Расширение обычно требует более содержательного обоснования и, при необходимости, новых исследований.

Жизненный цикл тесно связан с управлением рисками. В МИ ИВД риск часто проявляется через ошибочный результат и последующее действие. Поэтому любое изменение должно оцениваться не только по техническому признаку, но и по вопросу: может ли оно изменить вероятность или последствия ложноположительного, ложноотрицательного, неверно количественного, неопределенного или непра-

вильно интерпретированного результата? Если ответ положительный, прежняя доказательная база должна быть пересмотрена или дополнена.

Такой подход особенно важен для высокорисковых изделий: тестов для безопасности донорской крови, молекулярных изделий для выбора онкологической терапии, систем неонатального скрининга, тестов для выявления опасных инфекций, изделий для мониторинга критически значимых состояний и программных систем, автоматически формирующих клинический вывод. Для них изменения версии, материала, порога или алгоритма могут иметь значимые последствия, даже если внешне изделие остается тем же.

В итоге жизненный цикл медицинского изделия для диагностики *in vitro* показывает, что доказательная база не является статичным архивом. Она должна оставаться согласованной с фактически обращающимся изделием, его назначением, версией, инструкцией, клинической практикой и уровнем научного знания. Если эта согласованность сохраняется, ранее полученные доказательства продолжают поддерживать изделие. Если она нарушается, возникает необходимость в дополнительном обосновании, ограничении назначения, обновлении документации или проведении новых исследований.

Итог главы

Жизненный цикл медицинского изделия для диагностики *in vitro* охватывает разработку, подтверждение научной

обоснованности, аналитической и клинической эффективности, регистрацию, применение в практике, пострегистрационное наблюдение и изменения изделия. Клинические доказательства должны поддерживаться на протяжении всего этого цикла. Если меняются назначение, материал, популяция, версия программного обеспечения, алгоритм, правила интерпретации или клинический контекст, необходимо оценить, сохраняется ли применимость ранее полученных доказательств. Поэтому жизненный цикл является и технической последовательностью этапов, и важной частью доказательной архитектуры изделия.

Глава 6. Источники доказательств: исследования, литература и данные реальной практики

Источники доказательств имеют разную роль.

Клинические доказательства для медицинского изделия для диагностики *in vitro* редко формируются из одного источника. Обычно они складываются из нескольких групп данных: опубликованной литературы, клинических рекомендаций, аналитических исследований, клинических и клинико-лабораторных испытаний, архивного материала, биобанков, специальных панелей, данных по аналогичным изделиям, данных реальной практики (real-world data, RWD), формируемых на их основе доказательств реальной практики (real-world evidence, RWE), а также данных пост-регистрационного наблюдения. Но важен не сам перечень источников, а то, какую доказательную функцию выполняет каждый из них.



Рисунок 6.1. Основные источники доказательств

В регуляторной и экспертной работе часто встречается ошибка: большой объем собранных материалов воспринимается как сильная доказательная база. На практике это не всегда так. Десятки публикаций, несколько аналитических отчетов и массив клинических данных могут оставаться слабыми доказательствами, если они не связаны с назначением изделия. И наоборот, более компактный набор данных может быть убедительным, если каждый источник занимает понятное место в доказательной модели: один подтверждает

научную обоснованность, другой — аналитическую эффективность, третий — клиническую пригодность результата в заявленном сценарии.

Поэтому источник данных становится доказательством только тогда, когда ясно, какой вопрос он закрывает. Публикация о связи маркера с заболеванием может подтверждать научную обоснованность, но не доказывает работу конкретного изделия. Аналитическое исследование может показывать воспроизводимость метода, но не подтверждает автоматически клиническую пригодность результата. Клинико-лабораторное испытание может быть выполнено качественно, но не поддерживать всё заявленное назначение, если оно проведено на другой популяции, другом материале или в более узком клиническом сценарии.

Одни данные объясняют смысл результата, другие подтверждают работу изделия.

Для медицинского изделия для диагностики *in vitro* необходимо понимать, подтверждает ли источник научную обоснованность аналита, аналитическую эффективность конкретного изделия, клиническую эффективность результата, применимость к определенному материалу или популяции, либо ограничения заявленного назначения.

Литература чаще всего используется для подтверждения научной обоснованности. Она позволяет показать, что аналит, маркер, микроорганизм, генетический вариант, клеточная характеристика или иной объект исследования действи-

тельно связан с клиническим состоянием, патологическим процессом, прогнозом, ответом на лечение или другой медицинской задачей. Для хорошо изученных аналитов литература, клинические рекомендации и профессиональный консенсус могут составлять основу доказательной базы по клиническому смыслу аналита. Но литература должна быть применима к назначению изделия: важно, чтобы совпадали и название маркера, и клиническая роль, материал, популяция и правило интерпретации.

Например, данные о клиническом значении маркера в диагностике заболевания не всегда могут использоваться для обоснования его применения в скрининге. Публикации о значении биомаркера при одной опухоли не обязательно подтверждают его значение при другой. Данные о результате в сыворотке не всегда переносятся на сухое пятно крови, тканевой материал или цифровое изображение. Поэтому литературный источник нужно оценивать не по факту упоминания нужного аналита, а по его применимости к заявленному применению изделия.

Клинические рекомендации и консенсусные документы занимают особое место. Они показывают, как тот или иной результат используется в медицинской практике, какое место он занимает в диагностическом или лечебном маршруте, какие решения могут приниматься на его основе. Для хорошо изученных аналитов это может быть сильным подтверждением клинической роли результата. Однако реко-

мендации не заменяют доказательств по конкретному изделию. Они могут объяснить, почему определение HBsAg, вирусной нагрузки, EGFR, HER2, BCR::ABL1, фенилаланина или TREC имеет медицинский смысл, но не доказывают, что конкретная тест-система надежно и клинически пригодно получает соответствующий результат.

Аналитические исследования выполняют другую функцию. Они показывают, насколько надежно конкретное изделие получает результат. Здесь оцениваются правильность, прецизионность, предел обнаружения, предел количественного определения, линейность, диапазон измерений, аналитическая специфичность, интерференция, стабильность, воспроизводимость, прослеживаемость и иные характеристики, если они применимы к данному изделию. Для качественных тестов особое значение имеет поведение около порога; для количественных — диапазон и воспроизводимость; для молекулярных — специфичность мишени и предел обнаружения; для программных — устойчивость обработки данных и версия алгоритма.

Но аналитические исследования не должны использоваться как замена клинической эффективности. Даже если изделие демонстрирует отличную прецизионность, низкий предел обнаружения и отсутствие значимых интерференций, этого недостаточно, если не показано, что результат пригоден для заявленной клинической задачи. Аналитическая надежность создает основание для доверия к результату как

лабораторному факту, но не отвечает полностью на вопрос о его клиническом применении.

Клинические и клинико-лабораторные испытания связывают результат изделия с клиническим статусом, сравнительным методом, референтной рамкой или клиническим маршрутом. Именно поэтому они являются одним из центральных источников клинической эффективности. Однако и здесь важны границы применимости. Испытание должно соответствовать назначению изделия: материал, популяция, условия применения, пользователь, клиническая роль результата и метод сравнения должны быть релевантными. Иначе даже корректно проведенное исследование будет подтверждать только ту ситуацию, которая фактически была изучена, а не всю область, заявленную в инструкции.

Архивный клинический материал и биобанки особенно полезны в ситуациях, где трудно получить достаточное число проспективных случаев. Это характерно для редких заболеваний, редких генетических вариантов, отдельных онкологических мутаций, неонатального скрининга, некоторых инфекционных сценариев и проверки алгоритмов на сложных случаях. Такие материалы могут существенно усилить доказательную базу, если известны происхождение образца, условия хранения, критерии пригодности, клинический статус и связь с референтной оценкой. Но архивные материалы не должны незаметно подменять реальную целевую популяцию, если назначение изделия предполагает применение в

текущем клиническом потоке.

Специальные панели и искусственно сформированные образцы применяются тогда, когда обычный клинический поток не позволяет проверить важные, но редкие или критические ситуации. Например, они могут использоваться для оценки низких концентраций аналита, редких вариантов, слабopоложительных результатов, интерференентов, смешанных инфекций или пограничных случаев. Их ценность состоит в том, что они позволяют целенаправленно проверить уязвимые участки метода. Однако такие материалы имеют ограничения: они не всегда воспроизводят всю сложность реального клинического образца и не заменяют данные по фактической популяции применения.

Косвенные источники требуют осторожного применения.

Данные по аналогичным изделиям также могут использоваться как поддерживающий источник. Они помогают показать, что определенный технологический подход, аналит или клиническая роль уже имеют доказательную историю. Но перенос таких данных требует осторожности. Если новое изделие имеет другой принцип метода, другой порог интерпретации, другой материал, другую программную обработку или другое назначение, данные по аналогу не могут автоматически считаться применимыми. В таких случаях необходимо объяснить, что именно сопоставимо, а где требуются собственные данные.

Данные реальной практики становятся все более важным источником доказательств. Они показывают, как изделие работает после внедрения: в разных лабораториях, у разных пользователей, в неоднородных потоках образцов, при реальных преаналитических условиях и клинических маршрутах. Такие данные могут выявлять частоту невалидных результатов, поведение пограничных случаев, различия между площадками, редкие ошибки, проблемы интерпретации и влияние изменений практики. Для программных и алгоритмических изделий данные реальной практики особенно важны, потому что они позволяют отслеживать устойчивость модели в среде, которая может отличаться от среды разработки и испытаний.

Данные реальной практики (real-world data, RWD) — это данные, возникающие при реальном применении изделия и оказании медицинской помощи вне специально организованного исследования либо наряду с ним. После надлежащего сбора, анализа и интерпретации такие данные могут формировать доказательства реальной практики (*real-world evidence, RWE*), используемые для подтверждения или уточнения выводов о работе изделия в условиях практического применения. *RWD* и *RWE* не являются синонимами: наличие массива данных реальной практики еще не означает наличия доказательств. Доказательное значение возникает только после оценки происхождения данных, их качества, полноты, релевантности, возможных систематиче-

ских смещений и применимости к конкретному вопросу оценки.

При этом данные реальной практики требуют критической оценки. Они часто менее контролируемы, чем специально спланированное исследование. В них возможны пропуски, неполнота, смещения, неодинаковые правила включения, различия в качестве материала, неоднородные методы сравнения и слабая прослеживаемость клинического статуса. Поэтому такие данные могут быть очень ценными, но только если понятно, как они собраны, насколько они полны, как связаны с конкретным изделием и какие ограничения имеют.

В схеме 6.1 приведена обобщенная характеристика основных источников доказательств. Ее задача не в том, чтобы заменить подробную оценку каждого источника, а в том, чтобы показать их различную доказательную роль.

Схема 6.1 — Основные источники доказательств и их доказательная роль

— **Опубликованная литература** — *Основная доказательная роль: Подтверждение научной обоснованности и клинической роли анализа; Основное ограничение: Не доказывает автоматически работу конкретного изделия.*

— **Клинические рекомендации и консенсусные документы** — *Основная доказательная роль: Подтверждение места результата в медицинской практике; Основное ограничение: Не заменяют аналитические и клинические*

данные по изделию.

— **Аналитические исследования** — **Основная доказательная роль:** Подтверждение надежности получения результата конкретным изделием; **Основное ограничение:** Не доказывают автоматически клиническую пригодность результата.

— **Клинические и клинико-лабораторные испытания** — **Основная доказательная роль:** Подтверждение клинической эффективности в заявленном сценарии; **Основное ограничение:** Ограничены материалом, популяцией, дизайном и референтным подходом.

— **Архивный материал и биобанки** — **Основная доказательная роль:** Поддержка доказательств при редких состояниях, вариантах и сложных случаях; **Основное ограничение:** Требуют подтверждения пригодности и прослеживаемости материала.

— **Специальные панели и искусственно сформированные образцы** — **Основная доказательная роль:** Проверка редких, низкоуровневых, пограничных или трудных ситуаций; **Основное ограничение:** Не всегда отражают реальный клинический поток.

— **Данные по аналогичным изделиям** — **Основная доказательная роль:** Поддерживающее обоснование технологии, анализа или подхода; **Основное ограничение:** Требуют доказательства сопоставимости.

— **Данные и доказательства реальной практики**

(RWD/RWE) — Основная доказательная роль: Оценка поведения изделия в условиях реального применения, подтверждение или уточнение ранее сделанных выводов; **Основное ограничение:** Неоднородность, неполнота, возможные систематические смещения и необходимость подтверждения качества и применимости исходных данных.

На практике сильная доказательная база формируется не за счет механического накопления источников, а за счет их правильного распределения по доказательной функции. Один источник может подтверждать научную обоснованность, другой — аналитическую эффективность, третий — клиническую эффективность, четвертый — устойчивость результата после внедрения. Если эти источники согласованы с назначением изделия и между собой, они образуют доказательную линию. Если же они относятся к разным материалам, разным популяциям, разным версиям изделия или разным клиническим ролям, их суммирование не создает убедительного вывода.

Важно прямо указывать ограничения каждого источника. Литература может быть устаревшей или относиться к другой популяции. Аналитическое исследование может быть выполнено на модельных образцах. Клиническое испытание может не включать достаточное число пограничных случаев. Архивный материал может не полностью воспроизводить условия текущего применения. Данные реальной практики могут содержать неполные сведения о клиническом статусе.

Такие ограничения не обязательно делают источник непригодным, но они должны учитываться при формулировании итогового вывода.

Источник должен быть связан с доказательным выводом.

Следовательно, работа с источниками доказательств требует сбора данных, а также их критического размещения в доказательной модели. Нужно понимать, что именно подтверждает каждый источник, к какому элементу назначения он относится, какие ограничения имеет и можно ли на его основании поддержать заявленный вывод. Только при таком подходе набор материалов превращается в клинические доказательства.

Итог главы

Клинические доказательства медицинского изделия для диагностики *in vitro* формируются из разных источников: литературы, рекомендаций, аналитических исследований, клинических и клинико-лабораторных испытаний, архивных материалов, специальных панелей, данных по аналогам и данных реальной практики. Однако доказательная сила источника определяется не его названием, а тем, какой вопрос он закрывает и насколько он применим к назначению изделия. Сильная доказательная база возникает тогда, когда разные источники согласованы между собой и образуют последовательную линию от клинического смысла аналита к аналитической надежности изделия и клинической пригодности.

сти резултата.

Глава 7. Критическая оценка и взвешивание клинических доказательств

Данные еще не являются доказательствами.

Наличие данных само по себе еще не означает наличие клинических доказательств. Для медицинского изделия для диагностики *in vitro* важно не только собрать публикации, результаты аналитических исследований, протоколы испытаний, отчеты, сведения о методах сравнения, данные пост-регистрационного применения или материалы по аналогичным изделиям. Необходимо показать, что эти данные действительно относятся к заявленному назначению изделия, получены надежным способом, достаточны по объему, согласованы между собой и позволяют сделать именно тот вывод, который заявляется в документации.

Клинические доказательства не оцениваются по принципу простого накопления материалов. Большой объем документов может не подтверждать изделие, если данные относятся к другой популяции, другому типу биологического материала, другому клиническому сценарию, другой роли результата или иному способу применения. Напротив, меньший по объему, но хорошо спланированный и релевантный набор данных может иметь больший доказательный вес, ес-

ли он прямо отвечает на вопрос о пригодности изделия для заявленного применения.

Данные становятся доказательствами только тогда, когда понятно, какое именно утверждение они подтверждают. Например, публикации о связи маркера с заболеванием могут поддерживать научную обоснованность определения этого маркера. Однако они не доказывают автоматически, что конкретное изделие надежно определяет маркер, что результат пригоден для клинической классификации пациента или что изделие работает в заявленном типе биологического материала. Аналитические испытания могут подтвердить правильность, прецизионность, предел обнаружения, линейность, стабильность или устойчивость к интерференции, но эти данные также должны быть связаны с клиническим назначением изделия.

Для качественного теста важно не только показать, что изделие дает положительный и отрицательный результат. Нужно подтвердить, что такое разделение соответствует клиническому статусу, методу сравнения или иной обоснованной референтной рамке. Для количественного теста недостаточно показать воспроизводимость числового результата; необходимо понимать, используется ли результат для диагностики, мониторинга, выбора терапии, прогноза или иной медицинской задачи. Для молекулярного теста важно различать обнаружение последовательности, выявление клинически значимого варианта и интерпретацию результата для

конкретного клинического решения.

Критическая оценка доказательств нужна именно для того, чтобы отделить данные, которые действительно поддерживают заявленное назначение, от данных, которые имеют только косвенное, вспомогательное или ограниченное значение. В этой оценке учитывают не только итоговый результат исследования, но и его источник, дизайн, качество выполнения, состав выборки, метод сравнения, статистическую устойчивость, применимость к реальной практике и ограничения переноса вывода.

Сильное доказательство должно попадать в назначение.

Главный критерий доказательной ценности данных — их соответствие заявленному назначению изделия. Доказательство имеет вес только в пределах того назначения, к которому оно относится. Если данные получены в другой популяции, другой матрице, другом клиническом сценарии, у другого пользователя или для другой роли результата, их доказательная сила ограничена.

При оценке релевантности необходимо сопоставить данные с основными элементами назначения: определяемым анализом или объектом исследования, клиническим состоянием, целевой популяцией, типом биологического материала, предполагаемым пользователем, условиями применения, клинической ролью результата и способом интерпретации. Если хотя бы один из этих элементов существенно отличается

ся, требуется отдельное обоснование применимости данных. Например, если изделие заявлено для диагностики острой инфекции у симптомных пациентов, данные, полученные только на архивных образцах с высокой концентрацией возбудителя, будут иметь ограниченную доказательную силу. Если изделие заявлено для самотестирования, результаты, полученные только профессиональными пользователями в лабораторных условиях, не могут полностью подтверждать заявленное применение. Если набор реагентов заявлен для нескольких типов биологического материала, данные только по одному типу материала не должны автоматически переноситься на остальные.

Один и тот же аналит может иметь разное доказательное значение в разных клинических ролях. Маркер может быть пригоден для мониторинга, но недостаточен для первичной диагностики. Он может быть полезен для стратификации риска, но не для подтверждения заболевания. Он может иметь значение в одной популяции, но не быть проверенным в другой. Поэтому при критической оценке важно определить, какие элементы назначения подтверждены прямо, какие подтверждены косвенно, а какие остаются неподтвержденными.

Если доказательства относятся только к части назначения, итоговый вывод должен быть ограничен этой частью. Нельзя расширять вывод только потому, что технология изделия выглядит сходной, аналит тот же, а результаты в дру-

гой области применения были благоприятными. Для медицинских изделий для диагностики *in vitro* границы назначения одновременно являются границами доказательного вывода.

Надежность данных важна не меньше их количества.

Надежность источника доказательств определяется тем, насколько можно доверять полученным данным. Она зависит от дизайна исследования, качества его выполнения, контроля ошибок, выбора метода сравнения, полноты регистрации результатов, правил обработки проблемных данных, статистической устойчивости и прозрачности представления выводов.

Для клинических и клинико-лабораторных испытаний надежность повышается, если исследование заранее спланировано, имеет четкие критерии включения и исключения, использует обоснованный метод сравнения, предусматривает правила учета сомнительных, недействительных и отсутствующих результатов, включает достаточный объем данных и анализирует результаты в соответствии с заранее определенными критериями приемлемости.

Надежность снижается, если данные получены на удобной выборке, если проблемные результаты исключены без объяснения, если оператор знал результаты метода сравнения и мог повлиять на интерпретацию, если использован слабый или неподходящий референтный подход, если отсутствует

анализ подгрупп или если результаты объединены так, что скрывают слабые участки применения изделия.

Для литературных данных надежность зависит от качества публикаций, дизайна включенных исследований, прозрачности методов, сопоставимости популяций и отсутствия очевидных систематических ошибок. Для данных по аналогичным изделиям важно подтвердить, что аналогия действительно обоснована и не подменяет испытания конкретного изделия. Для пострегистрационных данных имеет значение полнота сбора, качество источника, возможность проверки и отсутствие выборочного представления только благоприятных результатов.

При оценке опубликованных исследований могут использоваться структурированные инструменты оценки риска систематических ошибок и применимости. Например, для исследований диагностической точности широко применяется QUADAS-2, позволяющий систематически оценивать отбор пациентов, индексный тест, референтный стандарт, поток участников и сроки исследования. Использование такого инструмента не заменяет профессионального суждения, но делает критическую оценку источников более последовательной и прозрачной.

Полнота доказательной базы означает, что представленные данные покрывают основные элементы заявленного назначения и ключевые риски неправильного результата. Недостаточно показать хороший результат в наиболее бла-

гоприятных условиях. Нужно оценить, подтверждены ли те условия, в которых изделие будет реально применяться: целевая популяция, тип биологического материала, клинически значимый диапазон значений, область около порога принятия решения, заявленные мишени, заявленные пользователи, условия применения и подгруппы, где возможны отличия работы изделия.

Особое значение полнота доказательной базы имеет для изделий с несколькими заявленными назначениями, несколькими мишенями, несколькими матрицами или несколькими группами пользователей. Высокая общая чувствительность мультиплексного теста не доказывает одинаковую эффективность по каждой мишени. Высокое общее согласие может скрывать низкое положительное согласие. Хороший результат у профессиональных пользователей не подтверждает автоматически применение изделия непрофессиональными пользователями.

Если доказательная база неполна, это не всегда означает невозможность положительного вывода. Но такой вывод должен быть ограничен. Необходимо четко указать, какие элементы назначения подтверждены, какие подтверждены частично, а какие требуют дополнительных данных.

Доказательства должны быть согласованы между собой.

Клинические доказательства часто формируются из нескольких источников: литературных данных, аналитиче-

ских испытаний, клинико-лабораторных испытаний, клинических данных, данных по методу сравнения, пострегистрационного опыта и сведений по аналогичным изделиям. Эти источники должны не просто существовать рядом, а согласованно поддерживать один доказательный вывод.

Согласованность означает, что разные источники данных не противоречат друг другу или что выявленные расхождения объяснены. Например, литературные данные могут подтверждать клиническую значимость аналита, аналитические испытания — способность изделия надежно определять этот аналит, а клинико-лабораторные испытания — пригодность результата для заявленной клинической задачи. В такой ситуации источники выполняют разные функции, но работают на общий вывод.

Если же литература подтверждает только общую биологическую связь маркера с заболеванием, аналитические данные получены на искусственных образцах, а клинико-лабораторные испытания проведены в другой популяции, доказательная база становится фрагментированной. Отдельные материалы могут быть полезны, но они не образуют убедительной линии доказательств для конкретного назначения.

Противоречия между источниками нельзя игнорировать. Если аналитические характеристики изделия выглядят убедительно, но клинико-лабораторные данные показывают низкую чувствительность в реальной выборке, итоговый вывод должен учитывать именно это ограничение. Если дан-

ные по аналогичному изделию благоприятны, но результаты испытаний конкретного изделия слабее, приоритет должен быть отдан прямым данным по оцениваемому изделию. Если пострегистрационный опыт выявляет проблемы, которые не были обнаружены в регистрационных испытаниях, доказательная база должна быть пересмотрена с учетом этих сведений.

Ограничения доказательств должны быть описаны явно. Ограничение не является недостатком само по себе, если оно корректно признано и учтено в выводе. Проблема возникает тогда, когда данные ограничены одной матрицей, одной популяцией, одной стадией заболевания, одним уровнем концентрации, одной площадкой или одним типом пользователя, а вывод формулируется так, будто подтверждено все заявленное назначение.

Для медицинских изделий для диагностики *in vitro* особенно опасен перенос вывода шире фактически изученных данных. Нельзя автоматически переносить результаты с плазмы на сыворотку, с венозной крови на капиллярную, с симптомных пациентов на скрининговую популяцию, с профессионального пользователя на непрофессионального, с лабораторных условий на внелабораторные, с одной мишени мультиплексного теста на все мишени панели. Такой перенос возможен только при наличии обоснования и подтверждающих данных.

Вес доказательств определяет силу вывода.

При взвешивании доказательств полезно различать прямые и косвенные данные. Прямые доказательства получены для оцениваемого изделия, в заявленном материале, целевой популяции и соответствующем клиническом сценарии. Косвенные доказательства относятся к аналогичному изделию, иной популяции, другому материалу, иной версии изделия или смежному клиническому применению и требуют отдельного обоснования переноса.

Косвенный характер данных не делает их автоматически неприемлемыми. Однако чем больше элементов назначения отличаются от условий, в которых данные были получены, тем более существенным становится вопрос их применимости и тем осторожнее должен формулироваться итоговый вывод.

Взвешивание клинических доказательств означает оценку относительной силы каждого источника в общей доказательной модели. Разные источники не имеют одинакового веса. Прямые данные испытаний конкретного изделия в заявленном применении обычно имеют больший вес, чем косвенные данные. Данные по аналогичным изделиям могут быть полезны, но требуют обоснования сопоставимости. Литературные данные могут поддерживать научную обоснованность, но не заменяют подтверждение работы конкретного изделия. Пострегистрационные данные могут уточнять доказательную базу, но их значение зависит от полноты, проверяемости и системности сбора.

Вес доказательств зависит от того, насколько источник отвечает на ключевой вопрос оценки. Если нужно подтвердить научную обоснованность анализа, важны публикации, клинические рекомендации, консенсусные документы и данные о связи анализа с медицинским состоянием. Если нужно подтвердить аналитическую эффективность, решающее значение имеют испытания конкретного изделия. Если нужно подтвердить клиническую эффективность, особенно важны данные, полученные на релевантном клиническом материале с обоснованной референтной рамкой и заранее определенными критериями оценки.

При взвешивании доказательств необходимо учитывать и силу отдельных источников, и их совокупность. Один сильный источник может закрывать ключевой вопрос, но не все элементы назначения. Несколько слабых источников не всегда создают сильную доказательную базу. Большое количество косвенных материалов не заменяет прямого подтверждения критически важного свойства изделия. Поэтому итоговый вывод должен основываться не на количестве документов, а на том, насколько доказательная база закрывает заявленные утверждения.

Схема 7.1 — Типовые критерии критической оценки клинических доказательств

— **Релевантность** — *Что оценивается: Соответствие данных назначению изделия, популяции, материалу, пользователю и клинической роли результата; Значение для*

доказательного вывода: Определяет, можно ли использовать данные для подтверждения заявленного применения.

— **Надежность** — **Что оценивается:** Дизайн исследования, метод сравнения, контроль ошибок, полнота регистрации и прозрачность анализа; **Значение для доказательного вывода:** Определяет, насколько можно доверять полученным данным.

— **Полнота** — **Что оценивается:** Покрытие всех существенных элементов назначения, матриц, мишеней, подгрупп и диапазонов; **Значение для доказательного вывода:** Определяет, не остается ли неподтвержденной часть заявления.

— **Согласованность** — **Что оценивается:** Отсутствие противоречий между литературными, аналитическими, клиническими и иными данными; **Значение для доказательного вывода:** Определяет, формируют ли источники единую доказательную линию.

— **Ограничения** — **Что оценивается:** Границы применимости данных и условия, в которых вывод может быть ослаблен; **Значение для доказательного вывода:** Определяет, насколько широко можно формулировать итоговый вывод.

— **Доказательный вес** — **Что оценивается:** Относительная сила источника в общей доказательной модели; **Значение для доказательного вывода:** Определяет, какие данные должны иметь приоритет при формировании вывода.

да.

Типовая ошибка критической оценки состоит в подмене доказательности объемом документации. Наличие большого числа публикаций, отчетов, таблиц и приложений не означает, что изделие подтверждено. Другая распространенная ошибка — перенос данных без обоснования: результаты по одной матрице, одной популяции или одному сценарию применения используются для подтверждения более широкого назначения. Часто встречается и подмена клинической эффективности аналитической: изделие хорошо измеряет сигнал, но не показано, что этот сигнал пригоден для заявленного клинического решения.

Еще одна ошибка связана с использованием общих показателей без анализа критических подгрупп. Общая чувствительность, общая специфичность или общее согласие могут выглядеть высокими, но при этом скрывать слабую работу изделия в отдельных мишенях, диапазонах концентраций, типах материала, стадиях заболевания или группах пользователей. Для доказательной оценки важен как общий результат, так и то, как изделие работает там, где ошибка наиболее значима.

Итоговый доказательный вывод должен быть сформулирован с учетом силы и ограничений всей доказательной базы. Он должен отвечать на несколько вопросов: что именно подтверждено, какими данными это подтверждено, насколько эти данные надежны, к какой области применения они

относятся, какие ограничения сохраняются и какие элементы назначения требуют дополнительных подтверждений. Если доказательства достаточны только для части назначения, вывод должен быть ограничен этой частью. Если доказательства неполны, это должно быть отражено в оценке и в дальнейшей программе получения данных.

Итог главы

Критическая оценка и взвешивание клинических доказательств позволяют превратить набор разнородных материалов в обоснованный доказательный вывод. Для медицинского изделия для диагностики *in vitro* важно и наличие данных, и их связь с назначением изделия, надежность, полнота, согласованность и применимость к заявленному клиническому сценарию. Доказательный вес источника определяется тем, насколько прямо и убедительно он подтверждает соответствующий элемент назначения. Итоговый вывод не должен быть шире полученных данных и должен ясно отражать как подтвержденные элементы, так и сохраняющиеся ограничения.

Глава 8. Документирование, демонстрация и генерация доказательств

Три вида доказательной работы.

При формировании клинических доказательств для медицинского изделия для диагностики *in vitro* важно различать три вида доказательной работы: документирование, демонстрацию и генерацию доказательств. Эти действия часто смешивают, хотя они решают разные задачи. Документирование отвечает на вопрос, какие данные уже существуют. Демонстрация показывает, почему эти данные применимы именно к заявленному назначению изделия. Генерация означает получение новых данных, если имеющихся сведений недостаточно для обоснованного вывода.

Разделение на документирование, демонстрацию и генерацию доказательств используется в настоящем издании как практическая модель организации доказательной работы. Оно не означает существования трех формально установленных регуляторных стадий. Его задача — помочь определить, какие сведения уже имеются, для каких утверждений их применимость должна быть специально обоснована и какие доказательные пробелы требуют получения новых данных.

Такое различие имеет практическое значение. Не каждое

изделие требует одинакового объема новых исследований. Для хорошо изученного аналита часть доказательной базы может быть уже сформирована литературой, клиническими рекомендациями и устойчивой медицинской практикой. В этом случае задача состоит не в том, чтобы заново доказывать очевидную клиническую связь, а в том, чтобы корректно ее документировать и показать применимость к конкретному изделию. Но если изделие использует новый аналит, новый материал, новую клиническую роль или новый алгоритм интерпретации, одного документирования обычно недостаточно. Тогда необходимо получать дополнительные данные.

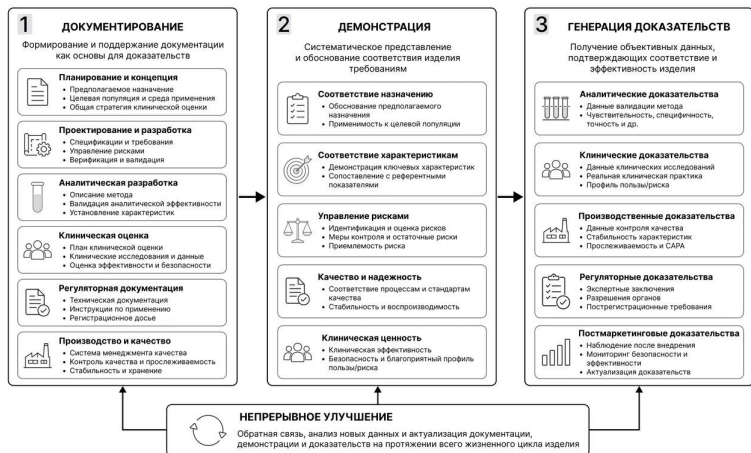


Рисунок 8.1. Документирование – демонстрация – генерация доказательств

Документирование не создает новых доказательств.

Документирование доказательств начинается с анализа уже доступных источников. Это могут быть опубликованные исследования, систематические обзоры, клинические рекомендации, консенсусные документы, данные по аналогичным изделиям, ранее выполненные аналитические исследования, архивные материалы или сведения реальной практики. Однако документирование не означает механическое перечисление источников. Каждый источник должен быть описан с точки зрения того, какой вопрос доказательной модели он закрывает и какие ограничения имеет.

Например, если изделие предназначено для определения хорошо известного маркера, литература может подтверждать связь этого маркера с заболеванием. Но в доказательной документации нужно показать не просто наличие публикаций, а их отношение к назначению изделия. Если изделие заявлено для диагностики, литература должна поддерживать диагностическую роль. Если для мониторинга — мониторинговую. Если для выбора терапии — предиктивную или терапевтически значимую роль. Без такого уточнения документирование остается неполным.

Демонстрация связывает данные с конкретным из-

делием.

Демонстрация доказательств является следующим шагом. Она отвечает на вопрос, можно ли использовать уже имеющиеся данные для конкретного изделия. Даже надежные публикации, качественные рекомендации или данные по аналогичным изделиям не всегда применимы напрямую. Нужно проверить, совпадают ли аналит, клиническая роль, тип материала, популяция, пользователь, условия применения, метод интерпретации и, при необходимости, версия программного обеспечения. Если эти элементы различаются, требуется обоснование переноса данных или получение новых доказательств.

Например, данные о клиническом значении аналита в сыворотке не всегда автоматически применимы к сухому пятну крови. Данные по пациентам с выраженной клинической картиной не всегда применимы к скринингу бессимптомных лиц. Публикации по одному методу выявления генетического варианта не всегда подтверждают изделие с другим алгоритмом фильтрации и интерпретации. Поэтому демонстрация доказательств — это не формальность, а критический этап, на котором определяется, могут ли уже имеющиеся сведения действительно поддерживать заявленное применение.

Генерация доказательств закрывает пробелы.

Генерация доказательств требуется тогда, когда существующих данных недостаточно. Она может включать анали-

тические исследования, клинические и клинико-лабораторные испытания, исследования на архивном материале, биобанках, специальных панелях, проверку программного компонента, внешнюю оценку алгоритма или сбор данных реальной практики. Генерация доказательств должна быть направлена не на получение “любых дополнительных данных”, а на закрытие конкретного доказательного пробела.

Различие между этими действиями особенно хорошо видно на примере хорошо изученных аналитов. Если изделие предназначено для количественного определения сердечного тропонина в диагностическом алгоритме острого повреждения миокарда, клиническое значение самого биомаркера может быть документировано на основании накопленной научной и клинической базы. Однако применимость этих знаний к конкретному изделию и конкретным порогам принятия решения требует отдельного обоснования, а аналитические и клинические характеристики конкретной системы должны подтверждаться собственными данными.

Иная ситуация возникает при новом биомаркере. Если маркер только начинает изучаться и его связь с клиническим состоянием подтверждена ограниченным числом исследований, документирование уже имеющихся публикаций будет лишь начальным этапом. Демонстрация их применимости может выявить серьезные ограничения: малые выборки, отсутствие независимого подтверждения, неоднородные популяции, разные пороги интерпретации. В таком случае для

подтверждения научной обоснованности и клинической эффективности может потребоваться генерация новых данных.

Та же логика действует при расширении назначения изделия. Если тест был подтвержден для диагностики у пациентов с симптомами, а затем заявляется для скрининга бессимптомной популяции, прежние данные можно документировать, но их применимость к новой роли должна быть специально продемонстрирована. Если такой перенос неубедителен, нужны новые исследования. Иначе возникает риск, что изделие будет заявлено шире, чем фактически подтверждено доказательствами.

Для программных и алгоритмических изделий разграничение этих действий имеет особое значение. Можно документировать клиническое значение входных параметров алгоритма. Можно продемонстрировать, что они относятся к заявленной медицинской задаче. Если клинически значимый результат формируется новой моделью, индексом, правилом классификации или иным алгоритмическим преобразованием, доказательная оценка должна относиться к входным параметрам, а также к итоговому результату алгоритма в пределах его заявленной клинической роли. В зависимости от назначения это может потребовать независимой проверки модели, оценки работы на внешних данных, анализа пограничных случаев, устойчивости к вариации входных данных и контроля версии алгоритма.

Схема 8.1 — Примеры различия между докумен-

тированием, демонстрацией и генерацией доказательств

— **Хорошо изученный анализ в стандартной роли**

— **Что можно документировать:** Литературу, рекомендации, устоявшуюся клиническую роль; **Что нужно продемонстрировать:** Применимость к назначению изделия; **Что может потребовать генерации данных:** Аналитическую и клиническую эффективность конкретного изделия.

— **Хорошо изученный анализ в новой роли** — **Что можно документировать:** Данные по прежнему применению; **Что нужно продемонстрировать:** Обоснованность переноса на новую роль; **Что может потребовать генерации данных:** Данные по новой клинической функции.

— **Новый анализ** — **Что можно документировать:** Предварительные публикации и биологическую правдоподобность; **Что нужно продемонстрировать:** Надежность и применимость имеющихся данных; **Что может потребовать генерации данных:** Данные о связи с клиническим состоянием и клинической эффективности.

— **Новый тип материала** — **Что можно документировать:** Данные по анализу в другом материале; **Что нужно продемонстрировать:** Обоснованность переноса на новый материал; **Что может потребовать генерации данных:** Данные о матриксных эффектах, стабильности и применимости результата.

— **Алгоритмический показатель** — **Что можно до-**

кументировать: Клиническое значение входных параметров; **Что нужно продемонстрировать:** Связь входных данных с заявленной задачей; **Что может потребовать генерации данных:** Данные по итоговому алгоритмическому результату.

— **Расширение назначения изделия** — **Что можно документировать:** Доказательства по исходному назначению; **Что нужно продемонстрировать:** Сопоставимость новой области применения; **Что может потребовать генерации данных:** Данные по новой популяции, материалу или клиническому сценарию.

Правильное разграничение этих действий помогает рационально планировать программу клинико-лабораторных испытаний. Если часть вопросов уже надежно закрыта документированными и применимыми данными, испытание не должно искусственно повторять их без необходимости. Но если остаются незакрытые вопросы, программа должна быть направлена именно на них. Например, если научная обоснованность анализа хорошо установлена, основное внимание может быть перенесено на аналитическую и клиническую эффективность конкретного изделия. Если же научная обоснованность сомнительна, начинать только с технической проверки метода будет методологически недостаточно.

Это различие важно и для регистрационного досье. Досье должно показывать и набор источников, и их функцию. В

нем должно быть понятно, какие сведения были документированы, какие данные были признаны применимыми после демонстрации, какие новые данные были получены и какие вопросы остаются ограничениями. Если эти роли не разделены, эксперт видит массив материалов, но не видит доказательной линии.

На практике слабое досье часто возникает не из-за полного отсутствия данных, а из-за неправильного распределения их доказательной роли. Литературу используют как замену испытаниям конкретного изделия. Аналитические данные подают как подтверждение клинической эффективности. Данные по одному материалу переносят на другой без объяснения. Результаты по старой версии алгоритма распространяют на новую версию без анализа изменений. Все эти ошибки связаны с тем, что документирование, демонстрация и генерация доказательств не были разделены.

Сильная доказательная стратегия строится иначе. Сначала формулируется назначение изделия. Затем анализируется, какие знания уже существуют и могут быть документированы. После этого оценивается, насколько они применимы к конкретному изделию. Далее выявляются пробелы, которые требуют новых данных. Только после этого разрабатывается программа испытаний или иного исследования. Такой порядок позволяет избежать двух крайностей: избыточных исследований там, где вопрос уже надежно закрыт, и недостаточной доказательности там, где литература или аналогичные

данные не могут заменить собственные данные по изделию.

В итоге документирование, демонстрация и генерация доказательств являются не отдельными бюрократическими этапами, а логикой работы с доказательной базой. Они позволяют понять, что уже известно, что применимо к данному изделию, а что еще нужно получить. Для медицинских изделий для диагностики *in vitro* это особенно важно, потому что доказательства почти всегда формируются из разных источников и должны быть связаны с конкретным назначением, материалом, популяцией, клинической ролью и риском применения.

Итог главы

Документирование, демонстрация и генерация доказательств помогают организовать доказательную работу при оценке медицинского изделия для диагностики *in vitro*. Документирование показывает, какие данные уже существуют. Демонстрация обосновывает их применимость к конкретному изделию и его назначению. Генерация обеспечивает получение новых данных, если имеющихся сведений недостаточно. Правильное разграничение этих действий позволяет рационально планировать клинико-лабораторные испытания, не повторять уже доказанное и не подменять данные по конкретному изделию общими литературными сведениями.

Итоги Части I

Часть I задала общую методологическую основу доказательной оценки медицинских изделий для диагностики *in*

vitro. Главная идея этой части состоит в том, что такие изделия оцениваются не только как лабораторные технологии, а как источники результата, способного повлиять на медицинское, лабораторное или организационное решение. Поэтому доказательная оценка должна начинаться с назначения изделия и клинической роли результата.

В части показано, что доказательная оценка может быть представлена через три взаимосвязанных элемента: научную обоснованность, аналитическую эффективность и клиническую эффективность. Научная обоснованность объясняет медицинское значение определяемого аналита или иного объекта исследования. Аналитическая эффективность подтверждает способность конкретного изделия надежно получать заявленный результат. Клиническая эффективность характеризует связь результата конкретного изделия с заявленным клиническим состоянием или иной медицински значимой характеристикой и его способность выполнять предусмотренную назначением клиническую функцию. Эти элементы не заменяют друг друга и должны быть согласованы между собой и с назначением изделия.

Отдельное значение имеет жизненный цикл изделия. Доказательства не заканчиваются регистрацией: они должны поддерживаться при изменении назначения, материала, популяции, версии программного обеспечения, алгоритма, правил интерпретации или клинической практики. Поэтому доказательная база должна оставаться связанной с фактиче-

ски обращающимся изделием на всех этапах его существования.

Часть I также показала, что источники данных сами по себе еще не являются доказательствами. Литература, рекомендации, аналитические исследования, клинико-лабораторные испытания, архивный материал, специальные панели и данные реальной практики приобретают доказательную силу только тогда, когда понятно, какой вопрос они закрывают и насколько они применимы к назначению изделия.

В дальнейшем эта методологическая рамка будет последовательно развернута в следующих частях издания: при рассмотрении научной обоснованности анализа, аналитической эффективности изделия, статистической обработки данных, клинической эффективности, особенностей отдельных типов МИ ИВД, клинических областей, программных и алгоритмических технологий, а также практики подготовки программы испытаний и регистрационного досье.

Часть II. Научная обоснованность медицинских изделий для диагностики *in vitro*

Аннотация части

Часть II посвящена научной обоснованности анализа как первому содержательному элементу клинических доказательств медицинского изделия для диагностики *in vitro*. Прежде чем оценивать правильность, прецизионность, предел обнаружения, диагностическую чувствительность или специфичность конкретного изделия, необходимо ответить на более ранний вопрос: почему вообще имеет медицинский смысл определять данный анализит или иной объект исследования.

Для изделий диагностики *in vitro* этот вопрос имеет принципиальное значение. Можно создать технически совершенный тест, который будет точно и воспроизводимо измерять показатель, не имеющий достаточного клинического значения для заявленной задачи. Такой тест может быть успешным как лабораторная технология, но слабым как медицинское изделие. Поэтому научная обоснованность не является второстепенным литературным приложением к аналитическим и клиническим исследованиям. Она задает смысловую основу всей последующей доказательной оценки.

В этой части рассматривается, что означает научная обоснованность, как устанавливается связь анализа с клиническим состоянием, физиологическим или патологическим процессом, чем отличается хорошо изученный анализ от нового, какие источники данных могут использоваться для подтверждения научной обоснованности и как такие данные должны критически оцениваться. Отдельное внимание уделяется литературному поиску, систематическим обзорам и документированию научной обоснованности в составе доказательной документации.

Особое значение имеет связь научной обоснованности с назначением изделия. Один и тот же анализ может быть хорошо обоснован для диагностики, но недостаточно обоснован для скрининга; применим для мониторинга, но не для выбора терапии; иметь значение в одной популяции, но не в другой. Поэтому научная обоснованность должна оцениваться не “вообще”, а применительно к конкретной клинической роли результата, типу материала, популяции и условиям применения изделия.

Цель части

Цель настоящей части состоит в том, чтобы сформировать у читателя понимание научной обоснованности как первого обязательного уровня доказательной оценки медицинских изделий для диагностики *in vitro*.

После изучения этой части читатель должен понимать, почему доказательная работа не может начинаться только с

аналитических характеристик изделия; чем клинически значимая связь анализа отличается от простой биологической правдоподобности или единичной статистической ассоциации; почему хорошо изученные и новые анализы требуют разных доказательных стратегий; какие источники данных могут подтверждать научную обоснованность; и как вывод о научной обоснованности должен быть связан с назначением изделия и последующей оценкой аналитической и клинической эффективности.

Глава 9. Понятие научной обоснованности

Научная обоснованность объясняет медицинский смысл результата.

Научная обоснованность является первым содержательным вопросом доказательной оценки медицинского изделия для диагностики *in vitro*. Она появляется раньше, чем обсуждение чувствительности метода, воспроизводимости результатов, диагностической эффективности или дизайна клинико-лабораторного испытания. До того как оценивать, насколько хорошо изделие выявляет или измеряет аналит, необходимо понять, почему этот аналит вообще имеет медицинское значение.

В международной терминологии этому понятию соответствует термин *scientific validity*. в настоящем издании далее используется русский термин — научная обоснованность.

Научная обоснованность — это доказанная связь между определяемым аналитом или иным объектом исследования и соответствующим клиническим состоянием, физиологическим или патологическим процессом либо медицинской задачей, для которой предназначено изделие. Она отвечает не на вопрос о том, насколько хорошо работает метод, а на вопрос о том, имеет ли медицинский смысл опре-

деление данного объекта в рамках заявленного назначения изделия.



Рисунок 9.1. Научная обоснованность как первый уровень доказательной оценки

Точный тест еще не всегда является клинически обоснованным.

Для диагностики *in vitro* это различие особенно важно. Лабораторная технология может быть точной, удобной, быстрой, автоматизированной и воспроизводимой, но если

измеряемый показатель не имеет доказанного клинического смысла для заявленной задачи, такая технология не образует полноценной доказательной базы медицинского изделия. И наоборот, если анализит хорошо известен и клинически значим, это еще не доказывает, что любое изделие, определяющее этот анализит, пригодно для медицинского применения. Научная обоснованность объясняет смысл выбора анализита, но не заменяет подтверждение работы конкретного изделия.

Именно здесь проходит одна из принципиальных границ доказательной оценки. Разработчик изделия может начать с технологии: имеется антитело, праймеры, сенсор, алгоритм, изображение или возможность измерить определенный сигнал. Но регуляторная логика начинается не с вопроса “можно ли это измерить?”, а с вопроса “зачем это измерять в медицинских целях?”. Если на этот вопрос нет убедительного ответа, последующие аналитические и клинические исследования теряют прочную основу. Изделие может демонстрировать хорошие лабораторные характеристики, но его результат будет недостаточно обоснован как медицински значимый.

Например, связь глюкозы с нарушениями углеводного обмена, HBsAg с вирусным гепатитом В или BCR::ABL1 с хроническим миелолейкозом не требует каждый раз доказывать заново с нуля. Эти анализиты имеют устоявшуюся клиническую основу. Но если изготовитель заявляет изделие для конкретной клинической роли, он должен показать, что ис-

пользует уже известную связь именно в тех пределах, в которых она подтверждена. Определение аналита для диагностики, мониторинга, прогноза, скрининга или выбора терапии — это не всегда одна и та же доказательная ситуация.

Для глюкозы крови клиническая связь с нарушениями углеводного обмена очевидна и давно закреплена в медицинской практике. Однако изделие для лабораторного количественного определения глюкозы, глюкометр для самостоятельного контроля и сенсор непрерывного мониторинга глюкозы опираются на одну и ту же базовую научную обоснованность, но реализуют ее в разных клинических и пользовательских сценариях. Поэтому научная обоснованность аналита может быть общей, а доказательная оценка конкретного изделия — различной. В одном случае основной вопрос будет связан с лабораторной точностью, в другом — с безопасностью самостоятельной интерпретации результата, в третьем — с пригодностью динамического сигнала для контроля состояния во времени.

Похожая ситуация возникает в инфекционной диагностике. Наличие нуклеиновой кислоты возбудителя может быть научно обоснованным маркером инфекции, но клинический смысл результата зависит от фазы заболевания, материала и назначения изделия. Один и тот же положительный результат может подтверждать активную инфекцию, отражать остаточное присутствие нуклеиновой кислоты, указывать на носительство или требовать сопоставления с клинической кар-

тиной. Поэтому научная обоснованность прямого теста на возбудителя должна учитывать не только сам факт связи мишени с инфекционным агентом, но и то, как результат будет интерпретироваться в заявленном клиническом сценарии.

В онкологии научная обоснованность особенно тесно связана с клинической ролью биомаркера. Один и тот же молекулярный признак может иметь диагностическое, прогностическое или предиктивное значение. Диагностическая роль помогает установить или уточнить заболевание. Прогностическая роль связана с вероятным течением заболевания. Предиктивная роль указывает на вероятность ответа на конкретное лечение. Если эти роли смешать, доказательная база становится методологически слабой. Данные о том, что маркер связан с неблагоприятным прогнозом, не всегда подтверждают его пригодность для выбора определенного препарата. Данные о наличии мутации в опухоли не всегда означают, что тест может использоваться как сопутствующая диагностика без дополнительных доказательств.

Важно установить, какая именно связь подтверждена

Научная обоснованность не сводится к тому, что анализит “имеет отношение” к заболеванию. В доказательной оценке важно установить, какая именно связь подтверждена. Анализит может быть связан с наличием заболевания, тяжестью процесса, риском развития состояния, прогнозом, ответом на терапию, фазой инфекции, активностью патологическо-

го процесса или принадлежностью к определенной биологической подгруппе. Каждая из этих ролей имеет собственное доказательное значение и не должна автоматически переноситься на другие роли.

Например, маркер может быть полезен для мониторинга течения заболевания, но не подходит для первичной диагностики. Другой маркер может иметь прогностическое значение, но не быть основанием для выбора конкретного лекарственного препарата. Третий может быть важен в уже установленной клинической группе, но не иметь достаточной ценности для массового скрининга. Поэтому научная обоснованность всегда должна рассматриваться вместе с назначением изделия.

Биологическая правдоподобность и ассоциация не равны научной обоснованности.

Особенно часто научную обоснованность смешивают с биологической правдоподобностью. Биологическая правдоподобность показывает, что связь анализа с состоянием возможна с точки зрения механизма. Это важный элемент рассуждения, но не окончательное доказательство. Для медицинского изделия необходимо показать, что связь и возможна, и подтверждена данными, воспроизводима, клинически значима и применима к заявленному использованию.

Не менее важно отличать научную обоснованность от единичной статистической ассоциации. Исследование может показать, что определенный показатель статистически

отличается у пациентов с заболеванием и без него. Но такой результат еще не означает, что показатель пригоден для диагностики, скрининга или прогноза. Нужно учитывать величину эффекта, качество исследования, независимое подтверждение, клинический контекст, популяцию, материал, пороги интерпретации и воспроизводимость результатов. Статистическая значимость не всегда означает медицинскую значимость.

На схеме 9.1 приведены основные понятия, с которыми не следует смешивать научную обоснованность.

Схема 9.1 — Что не является научной обоснованностью само по себе

— **Биологическая правдоподобность** — *Почему этого недостаточно: Объясняет возможный механизм связи, но не доказывает ее клиническую пригодность.*

— **Единичная статистическая ассоциация** — *Почему этого недостаточно: Может быть случайной, невозпроизводимой или клинически незначимой.*

— **Технологическая возможность измерения** — *Почему этого недостаточно: Показывает, что аналит можно измерить, но не объясняет, зачем это делать в медицинских целях.*

— **Высокая аналитическая чувствительность метода** — *Почему этого недостаточно: Подтверждает способность метода выявлять низкий уровень аналита, но не*

доказывает клинический смысл результата.

— **Упоминание маркера в публикациях** — *Почему этого недостаточно: Не показывает автоматически, что маркер пригоден для заявленного назначения изделия.*

— **Клиническая известность анализа в другой роли** — *Почему этого недостаточно: Не подтверждает без обоснования новую роль, материал, популяцию или правило интерпретации.*

Границу между биологической правдоподобностью и достаточной научной обоснованностью не следует понимать как объективно вычислимый порог. На практике это континуум, и решение о достаточности данных остаётся предметом экспертного суждения, зависящего от класса риска изделия и клинических последствий ошибочного результата: для изделий с серьёзными последствиями результата порог доказанности обычно выше, чем для изделий ориентировочного или вспомогательного назначения.

**Хорошо изученные и новые анализы требуют раз-
ной доказательной стратегии.**

Для хорошо изученных анализов научная обоснованность чаще всего подтверждается через документирование и демонстрацию применимости уже имеющихся данных. Это означает, что изготовитель или разработчик доказательной документации не обязан заново проводить фундаментальные исследования о связи глюкозы с сахарным диабетом или HBsAg с вирусным гепатитом В. Но он должен показать, что

заявленное изделие не выходит за пределы признанного клинического применения или что такое расширение дополнительно обосновано.

Здесь важно не спутать два разных уровня. Первый уровень — общая клиническая значимость аналита. Второй уровень — применимость этой значимости к конкретному назначению изделия. Если изделие определяет HBsAg для диагностики вирусного гепатита В, научная обоснованность может быть в значительной степени документирована. Но если заявляется иная роль результата, иной тип материала, иная популяция или иной клинический маршрут, одного общего знания о HBsAg уже недостаточно. Требуется показать, что известная клиническая связь сохраняет значение именно в этом применении.

Для новых аналитов ситуация другая. Если клиническое значение маркера только формируется, раздел о научной обоснованности становится более сложным. В нем необходимо анализировать не только наличие публикаций, но и качество данных, устойчивость связи, независимость подтверждения, применимость к целевой популяции и клиническую значимость результата. В некоторых случаях может потребоваться генерация новых данных, потому что литература не закрывает вопрос о связи аналита с заявленной медицинской задачей.

Новые аналиты особенно часто появляются в областях, где быстро развивается биомаркерная наука: онкология, ге-

нетика, метаболомика, протеомика, иммунология, цифровая патология и алгоритмическая диагностика. В этих направлениях ранние публикации могут выглядеть многообещающе, но они не всегда достаточны для регуляторного вывода. Исследования могут быть выполнены на малых выборках, в одной клинической группе, с неоднородными критериями заболевания, без независимого подтверждения или без ясного понимания того, как результат будет использоваться в медицинском маршруте. Поэтому новый анализ требует более осторожной доказательной стратегии.

Отдельного внимания требуют изделия, в которых результат формируется не одним анализом, а совокупностью признаков. Это могут быть мультиплексные панели, алгоритмические индексы, результаты секвенирования, цифровая морфология или системы искусственного интеллекта. В таких случаях научная обоснованность должна относиться к отдельным входным параметрам, а также к итоговому результату. Если программа объединяет несколько показателей и выдает категорию риска, нужно показать клинический смысл именно этой категории, а не только каждого показателя по отдельности.

Например, если алгоритм использует возраст пациента, несколько лабораторных показателей и цифровой признак изображения для расчета вероятности состояния, нельзя считать его научно обоснованным только потому, что каждый отдельный показатель когда-то изучался в литературе.

Итоговый индекс является самостоятельным объектом доказательной оценки. Нужно показать, что именно такая комбинация признаков и такое правило интерпретации связаны с заявленной медицинской задачей и дают клинически значимый результат.

Научная обоснованность выполняет еще одну важную функцию: она ограничивает последующие выводы. Если связь анализа с клиническим состоянием доказана только для определенной популяции, типа материала или клинической роли, то аналитическая и клиническая оценка изделия не должны незаметно выходить за эти пределы. Например, данные о маркере в опухолевой ткани не всегда применимы к циркулирующей опухолевой ДНК в плазме. Данные по взрослым пациентам не всегда применимы к новорожденным. Данные по симптомным пациентам не всегда применимы к скринингу бессимптомной популяции.

Поэтому раздел о научной обоснованности в доказательной документации должен завершаться не общим утверждением “аналит клинически значим”, а более точным выводом. В нем должно быть указано, с каким состоянием или медицинской задачей связан анализ, какая клиническая роль подтверждена, на каких данных основан вывод, к какой популяции и материалу эти данные применимы и какие ограничения сохраняются. Такой вывод становится основанием для перехода к оценке аналитической и клинической эффективности.

Итог главы

Научная обоснованность показывает, почему определяемый аналит или иной объект исследования имеет медицинский смысл для заявленного назначения изделия. Она не равна биологической правдоподобности, технологической возможности измерения, аналитической чувствительности метода или единичной статистической ассоциации. Для хорошо изученных аналитов основная задача состоит в корректном документировании и демонстрации применимости уже известных данных. Для новых аналитов требуется более развернутое подтверждение клинической связи и, при необходимости, получение дополнительных данных. Научная обоснованность также задает границы дальнейшей доказательной оценки: аналитическая и клиническая эффективность не должны выходить за пределы той клинической роли, материала и популяции, для которых связь аналита с медицинской задачей действительно подтверждена.

Глава 10. Связь анализа с клиническим состоянием, физиологическим или патологическим процессом

Связь анализа с состоянием должна быть конкретной.

Связь анализа с клиническим состоянием, физиологическим или патологическим процессом является содержательным ядром научной обоснованности. Именно эта связь объясняет, почему результат медицинского изделия для диагностики *in vitro* может иметь медицинское значение. Если такая связь не установлена, определение анализа остается только технической возможностью: его можно измерить, зарегистрировать, классифицировать или вычислить, но еще нельзя считать доказанным, что такой результат нужен для медицинского решения.

При этом выражение “аналит связан с заболеванием” слишком общее для доказательной оценки. Оно не показывает, какая именно связь имеется в виду, насколько она устойчива, в какой популяции подтверждена, в каком материале проявляется и для какой клинической роли может использоваться. Важно каждый раз уточнять: анализ помога-

ет подтвердить заболевание, исключить состояние, выделить группу риска, контролировать динамику процесса, оценить прогноз, выбрать терапию или стратифицировать пациента для дальнейшего медицинского маршрута.

В изделиях для диагностики *in vitro* под аналитом часто понимают не только измеряемое химическое вещество. Объектом исследования может быть белок, антиген, антитело, нуклеиновая кислота, генетический вариант, микроорганизм, токсин, клеточная популяция, морфологический признак, цифровое изображение, сигнал прибора или вычисляемый алгоритмом показатель. Поэтому при оценке научной обоснованности важно смотреть не на форму объекта, а на его клинический смысл. Вопрос всегда один: почему именно этот объект имеет значение для заявленной медицинской задачи?



Рисунок 10.1. Связь аналита с клинической задачей

Между аналитом и выводом есть несколько обязательных переходов.

Эта схема показывает главное: научная обоснованность не заканчивается на установлении факта, что аналит “имеет отношение” к заболеванию. Между аналитом и доказательным выводом есть несколько обязательных переходов. Нужно понять, какую клиническую роль выполняет результат, в каком материале он определяется, у какой популяции применяется и по какому правилу интерпретируется. Если хотя бы одно

из этих звеньев не определено, связь анализа с медицинской задачей остается неполной.

Связь анализа с состоянием может быть прямой или косвенной. В инфекционной диагностике выявление нуклеиновой кислоты возбудителя обычно ближе к прямому обнаружению инфекционного агента, но даже здесь клинический смысл результата зависит от контекста. Положительный результат может означать активную инфекцию, остаточное присутствие генетического материала, носительство, колонизацию или контаминацию, если речь идет о нестерильном локусе. Поэтому научная обоснованность такого теста должна учитывать факт существования мишени, а также клиническую интерпретацию ее обнаружения в конкретном материале и конкретной ситуации.

Серологические тесты показывают другую природу связи. Антитела не всегда свидетельствуют о текущем наличии инфекционного агента. Они могут отражать перенесенную инфекцию, вакцинацию, позднюю фазу заболевания, иммунный ответ или хронический процесс. Поэтому тест на антитела не может оцениваться по той же логике, что тест на РНК или ДНК возбудителя. Научная обоснованность серологического результата должна объяснять, какой именно клинический вопрос он закрывает: подтверждает ли факт контакта, помогает ли установить фазу инфекции, используется ли для скрининга, оценки иммунного статуса или иной медицинской задачи.

В онкологии связь анализа с клиническим состоянием часто еще сложнее. Биомаркер может иметь диагностическое значение, если помогает установить тип опухоли или молекулярный подтип. Он может иметь прогностическое значение, если связан с вероятным течением заболевания независимо от выбранного лечения. Он может иметь предиктивное значение, если указывает на вероятность ответа на конкретную терапию. Он может использоваться для мониторинга, если позволяет отслеживать остаточную болезнь, рецидив или появление резистентности. Эти роли нельзя смешивать. Данные о прогностическом значении биомаркера не доказывают автоматически его пригодность для выбора терапии, а данные о терапевтической значимости в одной нозологии не всегда применимы к другой.

Особенно показателен пример HER2, EGFR. В каждом случае клинический смысл результата зависит от заболевания, материала, метода, порога интерпретации и клинического действия. HER2 при раке молочной железы, HER2 при раке желудка, экспрессия HER2 по иммуногистохимии и амплификация гена по гибридизационному методу — это не полностью взаимозаменяемые доказательные ситуации. EGFR в опухолевой ткани при немелкоклеточном раке легкого и EGFR в циркулирующей опухолевой ДНК в плазме также требуют разных рассуждений о материале, чувствительности, клиническом маршруте и ограничениях результата.

Для количественных и пороговых тестов важен не только сам анализ.

Для количественных тестов связь анализа с состоянием часто выражается как через сам факт наличия анализа, так и через уровень, диапазон, динамику или порог. Вирусная нагрузка при хронической инфекции имеет значение не как бинарный ответ “есть/нет”, а как количественный показатель, используемый для мониторинга активности процесса и ответа на терапию. Концентрация фенилаланина в неонатальном скрининге имеет значение как первичный сигнал риска, который должен привести к повторному или подтверждающему этапу. Концентрация гормона, метаболита или лекарственного вещества может быть клинически значимой только в определенном интервале, при определенных условиях отбора материала и с учетом времени исследования.

Пороговые значения требуют отдельного внимания. Нельзя считать научно обоснованным только сам анализ, если клиническое решение принимается по порогу. Нужно понимать, почему выбран именно этот порог, для какой популяции он применим, как он соотносится с чувствительностью и специфичностью, как ведут себя результаты в пограничной зоне и что происходит с пациентом после положительного или отрицательного результата. В скрининге порог может быть выбран так, чтобы минимизировать пропуск случаев, но это увеличит число ложноположительных результатов. В подтверждающей диагностике логика может быть иной. Сле-

довательно, научная обоснованность включает и связь анализа с состоянием, и клинический смысл правила интерпретации.

Эта схема особенно важна для качественных и пороговых тестов. В таких изделиях клинический вывод часто формируется не самим числовым сигналом, а правилом его перевода в категорию. Поэтому научная обоснованность должна включать обоснование и анализа, и того, почему выбранное правило интерпретации соответствует заявленной медицинской задаче.

Схема 10.1 — Основные типы связи анализа с клинической задачей

— **Диагностическая** — **Что означает:** Анализ помогает подтвердить или исключить наличие состояния ; **Пример доказательного вопроса:** Позволяет ли результат различать пациентов с заболеванием и без него? .

— **Скрининговая** — **Что означает:** Анализ помогает выделить лиц с повышенной вероятностью состояния ; **Пример доказательного вопроса:** Пригоден ли результат для первичного отбора группы, требующей подтверждения? .

— **Мониторинговая** — **Что означает:** Анализ отражает изменение состояния во времени ; **Пример доказательного вопроса:** Можно ли использовать результат для оценки динамики заболевания или ответа на лечение? .

— **Прогностическая** — **Что означает:** Анализ связан с вероятным течением заболевания ; **Пример доказатель-**

ного вопроса: Связан ли результат с риском прогрессирования, рецидива или неблагоприятного исхода?.

— **Предиктивная** — **Что означает:** Аналит связан с вероятностью ответа на конкретное вмешательство; **Пример доказательного вопроса:** Может ли результат поддерживать выбор определенной терапии?.

— **Стратификационная** — **Что означает:** Аналит разделяет пациентов на клинически значимые подгруппы; **Пример доказательного вопроса:** Помогает ли результат выбрать маршрут обследования, наблюдения или тактику ведения?.

Эта схема показывает, почему один и тот же аналит не должен автоматически считаться обоснованным для всех возможных применений. Связь с заболеванием должна быть уточнена до конкретной роли. Если изделие заявлено для диагностики, требуется одно доказательное обоснование. Если для скрининга — другое. Если для выбора терапии — третье. Иногда эти роли могут поддерживаться одним и тем же массивом данных, но чаще требуется отдельная аргументация.

Приведённая типология связи аналита с клинической задачей — рабочая систематизация настоящего издания для целей изложения, а не терминология, дословно закреплённая в каком-либо международном документе. В зарубежной литературе и регуляторных источниках близкие по смыслу роли биомаркера чаще описываются набором diagnostic /

prognostic / predictive / companion, а стратификация риска обычно рассматривается как производная комбинация диагностической и прогностической функции, а не как самостоятельная категория.

Популяция, материал и время отбора меняют смысл результата.

Не менее важна популяция, в которой связь подтверждена. Аналит может хорошо работать у пациентов с выраженной клинической картиной, но быть слабым в популяции с низкой распространенностью состояния. В группе пациентов с симптомами предтестовая вероятность заболевания выше, и положительный результат может иметь один смысл. В массовом скрининге бессимптомных лиц тот же показатель может давать значительную ложноположительную нагрузку. Поэтому данные, полученные у пациентов с уже установленным или вероятным заболеванием, не всегда можно переносить на скрининговую популяцию.

Тип клинического материала также влияет на научную обоснованность. Аналит может иметь доказанное значение в одном материале, но не быть столь же интерпретируемым в другом. Например, молекулярный признак в опухолевой ткани и тот же признак в плазме как циркулирующая опухолевая ДНК имеют разную биологическую и аналитическую природу. Выявление микроорганизма в крови и выявление того же микроорганизма в материале из нестерильного локуса имеют разный клинический смысл. Определение метабо-

лита в сыворотке и в сухом пятне крови требует учета различий матрицы, стабильности, преаналитики и правил интерпретации.

Время получения образца также может менять значение результата. При инфекциях прямые маркеры возбудителя, антигены и антитела появляются и исчезают в разные периоды. При мониторинге лечения клиническое значение имеет как абсолютный уровень аналита, так и изменение относительно предыдущих результатов. В неонатальном скрининге время взятия сухого пятна крови после рождения может существенно влиять на концентрацию некоторых маркеров. Поэтому связь аналита с состоянием часто нельзя рассматривать вне временного контекста.

Для генетических тестов связь между вариантом и клиническим состоянием требует особенно аккуратного анализа. Обнаружение варианта само по себе не всегда означает заболевание или риск. Нужно учитывать патогенность варианта, тип наследования, пенетрантность, экспрессивность, популяционную частоту, клиническую интерпретацию и правила классификации. Для соматических вариантов в онкологии дополнительно важно, связан ли вариант с диагнозом, прогнозом, чувствительностью или резистентностью к конкретной терапии. Поэтому научная обоснованность генетического теста должна опираться на факт выявления варианта, а также на признанную клиническую интерпретацию этого варианта.

Для мультиплексных панелей и высокопроизводительного секвенирования проблема становится еще шире. Панель может включать десятки или сотни генов, вариантов, микроорганизмов или маркеров. Научная обоснованность такой панели не должна ограничиваться общим утверждением, что “гены связаны с заболеванием” или “микроорганизмы являются возбудителями”. Для каждого заявленного результата или группы результатов должно быть понятно, какую клиническую роль они выполняют. Иначе панель может содержать аналитически обнаруживаемые элементы, клиническая интерпретация которых для заявленного назначения остается неясной.

Алгоритмические показатели требуют отдельного подхода. Если изделие выдает вычисляемый индекс, категорию риска или автоматическую интерпретацию изображения, научная обоснованность должна относиться к итоговому результату алгоритма. Недостаточно показать, что каждый входной параметр имеет некоторую клиническую связь. Нужно доказать, что их сочетание, вес, правило классификации и итоговая категория действительно связаны с заявленной медицинской задачей. В противном случае алгоритм может выглядеть убедительно технически, но его медицинский смысл останется недоказанным.

Описать связь недостаточно — ее нужно оценить.

Связь аналита с клиническим состоянием должна быть описана, а также критически оценена. Для этого важно

учитывать качество источников, дизайн исследований, объем выборки, независимость подтверждения, репрезентативность популяции, определение клинического статуса, используемый материал и метод сравнения. Наличие нескольких слабых или однотипных исследований не всегда создаст сильную доказательную базу. Напротив, один хорошо выполненный систематический обзор или несколько независимых исследований в релевантной популяции могут дать более убедительное основание для вывода.

Также необходимо учитывать противоречивые данные. Если часть публикаций подтверждает связь анализа с состоянием, а часть показывает слабую или нестабильную связь, это нельзя игнорировать. В научной обоснованности важны не только положительные источники. Нужно объяснить, почему данные различаются: из-за популяции, метода, материала, порога, дизайна исследования или клинической роли результата. Если противоречия существенны и не разрешены, итоговый вывод должен быть осторожным.

В доказательной документации связь анализа с клиническим состоянием должна завершаться ясным выводом. Такой вывод не должен звучать слишком общо. Недостаточно написать, что “аналит связан с заболеванием”. Нужно указать, с каким состоянием или процессом связана мишень, какая клиническая роль подтверждена, в какой популяции и материале, какими источниками это подтверждается и какие ограничения сохраняются. Именно этот вывод затем опре-

деляет, какие аналитические и клинические исследования необходимы для конкретного изделия.

Следовательно, научная обоснованность строится не вокруг формального упоминания аналита в литературе, а вокруг клинически интерпретируемой связи. Чем точнее описана эта связь, тем устойчивее вся последующая доказательная оценка. Если связь определена расплывчато, клиническая эффективность изделия также будет трудно интерпретируема. Если же связь ясно привязана к назначению, материалу, популяции и клиническому действию, она становится надежной основой для дальнейшего подтверждения аналитической и клинической эффективности.

Итог главы

Связь аналита с клиническим состоянием, физиологическим или патологическим процессом является основой научной обоснованности. Для доказательной оценки недостаточно указать, что аналит “связан с заболеванием”. Необходимо определить характер этой связи: диагностический, скрининговый, мониторинговый, прогностический, предиктивный или стратификационный. Нужно учитывать популяцию, тип клинического материала, фазу процесса, время отбора образца, пороги интерпретации, пользователя, условия применения и клинический маршрут. Только такая конкретизированная связь позволяет обоснованно переходить к оценке аналитической и клинической эффективности медицинского изделия для диагностики *in vitro*.

Глава 11. Хорошо изученный и новый анализ: различия в доказательной стратегии

Известность анализа не равна доказанности его роли.

При подтверждении научной обоснованности важно различать хорошо изученные и новые анализы. Это различие не является формальной классификацией по “возрасту” маркера или времени его первого описания в литературе. Оно отражает степень доказанности клинической связи анализа с конкретной медицинской задачей. Чем надежнее и устойчивее эта связь подтверждена, тем больше оснований использовать уже накопленные данные. Чем новее или менее воспроизводима связь, тем осторожнее должна быть доказательная стратегия.

Хорошо изученный анализ — это не просто известный показатель. Это анализ или иной объект исследования, клиническая роль которого подтверждена достаточным объемом надежных данных и закреплена в медицинской практике, клинических рекомендациях, систематических обзорах или профессиональном консенсусе. Для такого анализа обычно не требуется заново доказывать сам факт его связи с соответствующим состоянием. Но это не означает, что изделие,

определяющее такой анализ, автоматически считается доказанным. Изготовитель должен показать, что использует известную клиническую связь именно в тех пределах, в которых она подтверждена.

Новый анализ — это анализ или объект исследования, клиническое значение которого для заявленной медицинской задачи еще не является устоявшимся. Он может быть биологически правдоподобным, перспективным, описанным в нескольких публикациях или связанным с новой технологией, но этого недостаточно для уверенного регуляторного вывода. Для нового анализа требуется более развернутая оценка качества данных, воспроизводимости связи, независимого подтверждения и применимости к целевой популяции.

Хорошо изученный анализ — это анализ или иной объект исследования, связь которого с определенной клинической задачей подтверждена надежными, воспроизводимыми и применимыми данными.

Новый анализ — это анализ или объект исследования, для которого такая связь еще не является достаточно установленной и требует дополнительного доказательного подтверждения.



Рисунок 11.1. Выбор доказательной стратегии для хорошо изученного аналита и нового аналита.

Один аналит может быть доказан в одной роли и не доказан в другой.

На практике один и тот же аналит может быть хорошо изученным в одной клинической роли и недостаточно обоснованным в другой. Например, маркер может быть давно признанным диагностическим признаком заболевания, но его использование для прогноза, мониторинга, скрининга или выбора терапии может требовать отдельного доказательного обоснования. Поэтому вопрос всегда должен ставиться точ-

нее: не “известен ли аналит?”, а “известна ли и доказана ли его роль именно в том назначении, которое заявлено для изделия?”.

Это особенно важно для изделий, которые выходят за пределы традиционного применения аналита. Если тест определяет известный инфекционный маркер у пациента с симптомами, доказательная стратегия будет одной. Если тот же маркер используется для скрининга бессимптомной популяции, доказательная нагрузка меняется. Если известный онкологический биомаркер применяется не для уточнения диагноза, а для выбора конкретной терапии, требуется подтверждение именно предиктивной роли. Если количественный показатель используется не для разовой оценки, а для мониторинга динамики, нужно обосновать значение изменений результата во времени.

Для хорошо изученного аналита нужно показать применимость известных данных.

Для хорошо изученного аналита основная задача состоит в документировании и демонстрации применимости уже имеющихся данных. Документирование означает, что в доказательной документации должны быть представлены источники, подтверждающие клиническую роль аналита: клинические рекомендации, систематические обзоры, ключевые исследования, консенсусные документы и устоявшаяся практика. Демонстрация применимости означает, что эти источники должны быть связаны с конкретным назначением

изделия. Недостаточно просто сослаться на известность маркера. Нужно показать, что совпадают клиническая роль, материал, популяция, правило интерпретации и условия применения.

Например, HBsAg является хорошо изученным маркером вирусного гепатита В. Но доказательная задача будет различаться в зависимости от того, заявлено ли изделие для диагностики у пациента с клиническим подозрением, для скрининга доноров, для подтверждающего этапа или для использования в определенной лабораторной среде. Базовая научная обоснованность маркера сохраняется, но объем данных, необходимых для конкретного изделия и конкретного назначения, будет различаться. То же относится к глюкозе, вирусной нагрузке, BCR::ABL1, HER2, EGFR, фенилаланину, и другим хорошо известным анализам.

Новый анализ требует осторожной доказательной стратегии.

Для нового анализа доказательная стратегия более требовательна. Здесь нельзя ограничиться ссылкой на механизм или ранние публикации. Нужно оценить, насколько связь анализа с состоянием устойчива, подтверждена ли она в независимых исследованиях, насколько релевантны выборки, как установлен клинический статус, какова величина эффекта, имеются ли противоречивые данные и можно ли использовать результат в заявленном клиническом маршруте. Если эти вопросы остаются открытыми, научная обосно-

ванность должна формулироваться осторожно, а назначение изделия — не расширяться за пределы имеющихся доказательств.

Новые аналиты часто появляются в быстро развивающихся областях: молекулярной онкологии, геномике, протеомике, метаболомике, иммунодиагностике, цифровой патологии и алгоритмической диагностике. В этих областях первые данные могут быть многообещающими, но не всегда достаточно зрелыми. Исследования могут быть ретроспективными, выполненными в одном центре, основанными на малых выборках, включать отобранные случаи или использовать разные правила интерпретации. Поэтому для нового аналита важно и наличие публикаций, и их доказательная надежность.

Схема 11.1 — Различия доказательной стратегии для хорошо изученного и нового аналита

— **Состояние доказательной базы** — *Хорошо изученный аналит*: Клиническая роль подтверждена надежными и применимыми данными; *Новый или недостаточно изученный аналит*: Клиническая роль требует дополнительного подтверждения.

— **Основная задача** — *Хорошо изученный аналит*: Документировать известные данные и показать их применимость к назначению изделия; *Новый или недостаточно изученный аналит*: Подтвердить связь с клиническим состоянием или медицинской задачей.

— **Основные источники** — *Хорошо изученный анализ*: Клинические рекомендации, систематические обзоры, консенсусные документы, устоявшаяся практика; *Новый или недостаточно изученный анализ*: Первичные исследования, независимые подтверждения, собственные данные при необходимости.

— **Основной риск** — *Хорошо изученный анализ*: Необоснованное расширение назначения за пределы известных данных; *Новый или недостаточно изученный анализ*: Преждевременное признание клинической значимости.

— **Типичная ошибка** — *Хорошо изученный анализ*: Перенос доказательств с одной роли, популяции или материала на другую без обоснования; *Новый или недостаточно изученный анализ*: Опора на биологическую правдоподобность или единичные публикации.

— **Потребность в новых данных** — *Хорошо изученный анализ*: Обычно ниже для научной обоснованности, но сохраняется для конкретного изделия; *Новый или недостаточно изученный анализ*: Обычно выше, особенно при новой клинической роли или недостаточной литературной базе.

Расширение применения требует дополнительного обоснования.

Даже когда анализ хорошо изучен, необходимо внимательно оценивать границы применения известных данных. Если изделие заявляет новый тип материала, прежняя науч-

ная обоснованность может оказаться неполной. Например, клиническое значение маркера в сыворотке не всегда переносится на сухое пятно крови, мочу, кал, ткань, мазок или цифровое изображение. При смене материала меняются концентрационный диапазон, стабильность, преаналитические факторы, матриксные эффекты и правила интерпретации. Поэтому известность аналита не отменяет необходимости обоснования материала.

То же относится к популяции. Данные, полученные у взрослых пациентов с выраженной клинической картиной, не всегда применимы к детям, новорожденным, бессимптомным лицам, донорам, пациентам на терапии или лицам из группы риска. В разных популяциях меняются распространенность состояния, спектр клинических случаев, вероятность ложных результатов и последствия ошибки. Поэтому научная обоснованность должна быть соотнесена с аналитом, а также с той группой, для которой заявлено изделие.

Особое внимание нужно уделять новой клинической роли известного аналита. Например, если маркер используется для мониторинга, его научная обоснованность должна включать значение изменения результата во времени. Если маркер используется для прогноза, необходимо показать связь с исходами. Если результат используется для выбора терапии, нужна предиктивная связь с ответом на конкретное вмешательство. Нельзя автоматически переносить доказательства с диагностической роли на прогностическую или предиктив-

ную.

Для алгоритмических и многопараметрических изделий понятие “новый анализ” приобретает дополнительный смысл. Итоговый индекс, категория риска или алгоритмический вывод могут рассматриваться как новый объект доказательной оценки, даже если входные параметры хорошо известны. Если программа объединяет несколько лабораторных показателей, данные изображения и клинические признаки, то научная обоснованность должна относиться к итоговому результату. Нужно показать, что именно эта комбинация признаков, весов и правил интерпретации имеет клиническое значение.

Например, отдельные показатели метаболического профиля в неонатальном скрининге могут быть давно известны. Но если программный модуль использует их сочетание для автоматической классификации образца как первично положительного или отрицательного, доказательной оценке подлежит уже не только каждый показатель, но и сама логика классификации. Аналогично в цифровой патологии известность маркера PD-L1 или HER2 не доказывает автоматически клиническую обоснованность конкретного алгоритма, который сегментирует изображение, выбирает клетки и рассчитывает итоговую категорию.

Схема 11.2 — Когда хорошо изученный анализ требует дополнительного обоснования
— Новая клиническая роль — Почему требуется до-

полнительное обоснование: Данные по диагностике не подтверждают автоматически прогноз, мониторинг или выбор терапии.

— **Новый тип клинического материала** — **Почему требуется дополнительное обоснование:** Меняются матриксные эффекты, стабильность, концентрации и интерпретация результата.

— **Новая целевая популяция** — **Почему требуется дополнительное обоснование:** Меняются распространенность состояния, спектр случаев и вероятность ошибочного результата.

— **Новый порог интерпретации** — **Почему требуется дополнительное обоснование:** Нужно показать клинический смысл выбранного порога и поведение результата около него.

— **Новый алгоритм обработки** — **Почему требуется дополнительное обоснование:** Итоговый результат может иметь иной клинический смысл, чем отдельные входные данные.

— **Расширение назначения изделия** — **Почему требуется дополнительное обоснование:** Доказательства должны покрывать всю заявленную область применения.

Для новых анализов особенно важно не завышать выводы. Если имеющиеся публикации показывают только предварительную связь, раздел научной обоснованности должен так и

отражать: данные являются многообещающими, но ограниченными; требуется независимое подтверждение; клиническая роль пока установлена не полностью; назначение.

Новый анализ может быть принят в доказательную модель только тогда, когда его связь с медицинской задачей описана достаточно ясно. Необходимо понимать, какой результат ожидается, как он будет интерпретироваться, что изменится в клиническом маршруте и какие риски связаны с ошибкой. Если эти вопросы не решены, дальнейшие аналитические исследования могут подтвердить только техническую возможность измерения, но не медицинскую значимость результата.

Практически различие между хорошо изученным и новым анализом должно отражаться в структуре доказательной документации. Для хорошо изученного анализа раздел научной обоснованности может быть более компактным, но он должен ясно показывать применимость известных данных. Для нового анализа раздел должен быть более развернутым: с описанием литературного поиска, критической оценкой источников, анализом противоречий, обсуждением ограничений и, при необходимости, обоснованием дополнительных исследований.

В итоге различие между хорошо изученным и новым анализом помогает рационально распределить доказательные усилия. Не нужно заново доказывать то, что уже надежно установлено. Но нельзя прикрывать общими знаниями новое или расширенное применение. Задача состоит в том, чтобы

определить, какие данные уже можно использовать, где требуется демонстрация применимости, а где необходимо получить новые доказательства.

Итог главы

Различие между хорошо изученным и новым анализом определяет характер подтверждения научной обоснованности. Для хорошо изученного анализа обычно достаточно документировать известные данные и показать их применимость к назначению изделия. Для нового или недостаточно изученного анализа требуется более развернутая оценка связи с клиническим состоянием, качества источников, воспроизводимости данных и ограничений. При этом один и тот же анализ может быть хорошо изученным в одной клинической роли и недостаточно обоснованным в другой. Поэтому доказательная стратегия всегда определяется как самим анализом, так и назначением изделия, типом материала, популяцией, правилом интерпретации и клиническим риском.

Глава 12. Источники данных для установления научной обоснованности

**Источник должен объяснять медицинский смысл
аналита.**

Научная обоснованность устанавливается не одним универсальным источником, а совокупностью данных, которые показывают связь анализа или иного объекта исследования с клиническим состоянием, физиологическим или патологическим процессом либо медицинской задачей. При этом важно понимать, что источники данных для научной обоснованности отличаются по своей функции от источников, подтверждающих аналитическую или клиническую эффективность конкретного изделия. На этом этапе главный вопрос состоит не в том, насколько хорошо работает тест-система, прибор или программный модуль, а в том, почему выбранный анализ имеет медицинский смысл.

Для хорошо изученных анализов такими источниками обычно являются клинические рекомендации, систематические обзоры, руководства профессиональных обществ, крупные клинические исследования, консенсусные документы и устоявшаяся практика применения результата. Для новых анализов доказательная база чаще складывается из

первичных исследований, исследований ассоциации с клиническим состоянием, данных о биологическом механизме, независимых подтверждений, работ по клинической интерпретации результата и, при необходимости, собственных исследований разработчика. Но в обоих случаях источник должен оцениваться через один и тот же вопрос: подтверждает ли он связь аналита с той клинической ролью, которая заявлена для изделия.

Нельзя считать достаточным любой источник, в котором просто упоминается нужный аналит. Публикация может показывать, что показатель отличается у пациентов и здоровых лиц, но не объяснять, пригоден ли он для диагностики. Исследование может описывать механизм заболевания, но не показывать клиническую применимость результата. Рекомендация может упоминать маркер в одном клиническом маршруте, но изделие может быть заявлено для другой роли. Поэтому источник данных для научной обоснованности должен быть тематически связан с аналитом, а также доказательно применим к назначению изделия.

Клинические рекомендации показывают место результата в практике.

Одним из наиболее сильных источников для хорошо изученных аналитов являются клинические рекомендации. Их значение состоит в том, что они отражают не только наличие научных данных, но и место результата в медицинской практике. Если аналит включен в диагностический, скри-

нинговый, мониторинговый или терапевтический маршрут, это обычно говорит о том, что его клиническая роль признана профессиональным сообществом. Например, рекомендации могут показывать, когда следует определять определенный инфекционный маркер, какой результат считается основанием для подтверждающего этапа, как интерпретируется молекулярная находка в онкологии или какой показатель используется для мониторинга ответа на лечение.

Однако клинические рекомендации тоже нужно читать критически. Они могут относиться к другому маршруту пациента или другому методу исследования. Иногда рекомендация подтверждает клинический смысл анализа, но не касается конкретного типа материала или технологии. Поэтому рекомендации редко являются единственным источником научной обоснованности. Они особенно сильны тогда, когда согласуются с систематическими обзорами, клиническими исследованиями и практикой применения результата.

Обзоры и первичные исследования показывают устойчивость связи.

Систематические обзоры и метаанализы также занимают важное место. Их ценность состоит в том, что они обобщают несколько исследований и позволяют оценить устойчивость связи анализа с клинической задачей. Если разные исследования, выполненные в разных группах, показывают сходную связь маркера с состоянием, это усиливает научную обоснованность. Если же обзор выявляет выраженную неоднород-

ность, противоречия, слабые дизайны или нестабильность результатов, это также важно: такие данные не всегда позволяют сделать уверенный вывод и могут требовать более осторожного назначения изделия.

Для новых анализов особое значение имеют первичные клинические исследования. Они могут показывать связь аналита с заболеванием, стадией процесса, прогнозом, ответом на терапию или другой медицинской задачей. Но такие исследования нужно оценивать по качеству. Важно, как сформирована выборка, как установлен клинический статус, был ли независимый контроль, насколько достаточен объем данных, каковы критерии включения и исключения, как определялись пороги, проводилась ли внешняя проверка и воспроизводится ли результат в других исследованиях. Чем менее зрелый анализ, тем больше значение имеет критическая оценка таких деталей.

Механизм поддерживает обоснование, но не заменяет клинические данные.

Данные о биологическом механизме могут поддерживать научную обоснованность, но обычно не заменяют клинические данные. Механистическое объяснение показывает, почему связь аналита с процессом возможна. Например, экспрессия определенного белка может быть связана с сигнальным путем опухоли, мутация — с механизмом резистентности, метаболит — с нарушением ферментного пути, антитела — с иммунным ответом. Но биологическая правдопо-

добность сама по себе не доказывает, что результат пригоден для клинического применения. Она должна быть связана с данными, показывающими воспроизводимую и клинически значимую связь.

Особую группу составляют консенсусные документы и экспертные мнения. Они часто используются в онкологии, генетике, инфекционной диагностике, редких заболеваниях и лабораторной медицине. Их значение особенно велико там, где прямые клинические исследования ограничены, а интерпретация результата требует профессионального соглашения. Например, классификация генетических вариантов, уровни клинической значимости соматических мутаций или профессиональные рекомендации по интерпретации морфологических признаков могут играть важную роль в научной обоснованности. Но и здесь необходимо учитывать, насколько документ актуален, на каких данных основан консенсус и применим ли он к заявленному назначению изделия.



Рисунок 12.1. Как источник данных поддерживает научную обоснованность.

Для некоторых изделий источником научной обоснованности могут быть данные о клиническом маршруте. Это особенно важно для скрининга, мониторинга и изделий, результат которых запускает последующее действие. В таких случаях нужно показать и связь анализа с состоянием, и место результата в цепочке медицинских решений. Например, первично положительный результат в неонатальном скрининге не является окончательным диагнозом, но научно обоснован как сигнал для повторного или подтверждающего этапа. Такая роль должна подтверждаться и биологической связью

маркера с заболеванием, и данными о том, что раннее выявление и последующее действие имеют медицинский смысл.

На рисунке 12.1 показано, что один и тот же источник может иметь разный доказательный вес в разных контекстах. Например, систематический обзор по диагностической точности маркера может быть сильным источником для диагностического назначения, но не закрывать вопрос о прогнозе или выборе терапии. Клиническая рекомендация может подтверждать место результата в стандартном маршруте, но не обосновывать новый тип материала. Первичное исследование может быть убедительным для одной популяции и неприменимым для другой. Поэтому источник данных всегда нужно “привязывать” к конкретному вопросу научной обоснованности.

Важным источником могут быть данные по аналогичным диагностическим подходам. Если определенный анализ уже используется в клинической практике, данные по существующим изделиям, методам или лабораторным подходам могут поддерживать научную обоснованность. Однако их роль нужно ограничивать. Они помогают подтвердить клинический смысл анализа или подхода, но не доказывают автоматически эффективность нового изделия. Особенно осторожно следует использовать такие данные, если новое изделие имеет другой тип материала, иной принцип обнаружения, иной порог, другую программную интерпретацию или заявляет более широкую клиническую роль.

Для алгоритмического результата нужно обосновывать итоговый вывод.

Для алгоритмических и программных изделий источники научной обоснованности имеют свои особенности. Если итоговый результат формируется алгоритмом, источники должны подтверждать не только клинический смысл входных параметров, но и клиническое значение итогового вывода. Например, если модель объединяет несколько лабораторных показателей и выдает категорию риска, литература по каждому отдельному показателю не является достаточной. Нужны данные, показывающие, что именно комбинация признаков и правило классификации связаны с медицинской задачей. Это может быть опубликованная разработка модели, независимая проверка алгоритма, клиническое исследование или внешняя валидация на релевантной выборке.

В редких заболеваниях, неонатальном скрининге и некоторых онкологических ситуациях прямые данные могут быть ограничены. В таких случаях научная обоснованность часто формируется из нескольких источников: данных о патогенезе, клинических описаний, регистров, консенсусов, биобанковых материалов, архивных подтвержденных случаев и практики специализированных центров. Такая доказательная база может быть приемлемой, но она должна быть честно описана. Необходимо указать, какие данные являются прямыми, какие поддерживающими, какие имеют ограничения и какие вопросы остаются открытыми.

Схема 12.1 — Основные источники данных для установления научной обоснованности

— **Клинические рекомендации** — *Что может подтвердить:* Признанную клиническую роль аналита и место результата в медицинском маршруте; *Основные ограничения:* Могут не относиться к конкретному материалу, популяции или новому назначению.

— **Систематические обзоры и метаанализы** — *Что может подтвердить:* Устойчивость связи аналита с состоянием или клинической задачей; *Основные ограничения:* Зависят от качества и однородности включенных исследований.

— **Первичные клинические исследования** — *Что может подтвердить:* Прямую связь аналита с состоянием, прогнозом, ответом на терапию или иной задачей; *Основные ограничения:* Могут быть ограничены дизайном, выборкой, смещениями и отсутствием внешнего подтверждения.

— **Консенсусные документы и экспертные классификации** — *Что может подтвердить:* Признанную интерпретацию результата или клиническую значимость признака; *Основные ограничения:* Требуют оценки актуальности и основания консенсуса.

— **Данные о биологическом механизме** — *Что может подтвердить:* Биологическую правдоподобность

связи анализа с процессом; **Основные ограничения:** Не заменяют клиническое подтверждение.

— **Данные по аналогичным подходам** — **Что может подтвердить:** Поддержку клинического смысла анализа или метода; **Основные ограничения:** Не доказывают автоматически применимость к новому изделию.

— **Регистры и данные специализированных центров** — **Что может подтвердить:** Сведения по редким состояниям, вариантам и клиническим сценариям; **Основные ограничения:** Возможны неполнота, селекция случаев и ограниченная переносимость.

— **Данные реальной практики** — **Что может подтвердить:** Фактическое использование результата в клиническом маршруте; **Основные ограничения:** Требуют оценки качества, прослеживаемости и полноты данных.

— **Данные по алгоритмическим моделям** — **Что может подтвердить:** Клиническое значение итогового индекса или категории риска; **Основные ограничения:** Нужна независимая проверка и оценка применимости к целевой популяции.

Наличие источника из этой схемы не означает автоматической достаточности научной обоснованности. Например, клиническая рекомендация может быть сильным источником, если она прямо относится к заявленному применению изделия. Но если изделие заявлено для новой роли результата, такая рекомендация может выполнять только поддер-

живающую функцию. Первичное исследование может быть важным, если оно релевантно и качественно выполнено, но слабым, если основано на малой выборке и не подтверждено независимо. Данные реальной практики могут быть ценными, если они прослеживаемы и относятся к фактическому клиническому маршруту, но недостаточными, если в них неясно определен клинический статус.

Оценка источников для научной обоснованности должна учитывать актуальность. В быстро развивающихся областях данные могут устаревать. Это особенно характерно для молекулярной онкологии, генетики, интерпретации вариантов, инфекционной диагностики и алгоритмических систем. Маркер, который ранее рассматривался как перспективный, может утратить значение при появлении новых данных. Напротив, показатель, который раньше имел ограниченную роль, может получить более сильное обоснование после появления новых клинических исследований или рекомендаций. Поэтому научная обоснованность должна опираться на актуальный массив данных, а не только на исторические публикации.

Еще одна важная задача — различать источники, подтверждающие саму связь анализа с состоянием, и источники, подтверждающие правило интерпретации результата. Иногда анализ как таковой хорошо обоснован, но порог, категория или алгоритм интерпретации требуют отдельного подтверждения. Например, уровень показателя может быть

связан с заболеванием, но конкретный диагностический порог должен быть обоснован отдельно. Мутация может иметь клиническое значение, но правило включения ее в отчет должно учитывать классификацию варианта и клинический контекст. Алгоритм может использовать известные параметры, но итоговая категория риска требует отдельной проверки.

При подготовке доказательной документации источники данных для научной обоснованности нужно не просто перечислять, а группировать по доказательной функции. Одни источники подтверждают биологическую основу связи, другие — клиническую роль результата, третьи — применимость к конкретной популяции, четвертые — значение порога или правила интерпретации, пятые — место результата в маршруте пациента. Такая группировка делает вывод более прозрачным и показывает, какие элементы научной обоснованности подтверждены надежно, а какие остаются ограниченными.

Важно также не скрывать слабые и противоречивые данные. Если в литературе имеются исследования, которые не подтверждают связь аналита с заявленной задачей, они должны быть рассмотрены. Игнорирование таких данных делает обоснование селективным. Корректнее объяснить, почему результаты различаются: возможно, исследования включали разные популяции, материалы, методы, пороги или клинические роли. Если противоречия невозможно объ-

яснить, это должно ограничивать силу вывода о научной обоснованности.

В итоге выбор источников для установления научной обоснованности зависит от зрелости анализа и заявленного назначения изделия. Для хорошо изученного анализа достаточно надежного документирования и демонстрации применимости признанных источников. Для нового анализа требуется более глубокий анализ первичных данных и, возможно, генерация новых доказательств. Для алгоритмического результата нужно подтверждать не только входные параметры, но и итоговый вычисляемый вывод. Во всех случаях источник данных должен отвечать на конкретный вопрос: почему результат изделия имеет медицинский смысл именно в заявленном применении.

Итог главы

Источники данных для установления научной обоснованности должны подтверждать связь анализа или иного объекта исследования с клиническим состоянием, физиологическим или патологическим процессом либо медицинской задачей. К таким источникам относятся клинические рекомендации, систематические обзоры, первичные исследования, консенсусные документы, данные о биологическом механизме, данные по аналогичным подходам, регистры, данные реальной практики и материалы по алгоритмическим моделям. Их доказательная сила определяется не названием источника, а его применимостью к назначению изделия,

клинической роли результата, типу материала, популяции и правилу интерпретации.

Глава 13. Литературный поиск при установлении научной обоснованности

Литературный поиск должен отвечать на конкретный доказательный вопрос.

Литературный поиск при установлении научной обоснованности отличается от общего поиска клинических доказательств. Его задача состоит не в том, чтобы собрать все возможные сведения об изделии или методе, а в том, чтобы ответить на более конкретный вопрос: подтверждена ли связь анализа или иного объекта исследования с клиническим состоянием, физиологическим или патологическим процессом либо медицинской задачей, заявленной для изделия. Поэтому такой поиск должен строиться не вокруг технологии как таковой, а вокруг клинического смысла определяемого объекта.

На практике это различие часто теряется. Разработчик может искать публикации по названию анализа, гена, белка, микроорганизма или метода и получать большой массив источников. Но значительная часть таких публикаций может не отвечать на вопрос научной обоснованности. Одни работы могут быть посвящены аналитической методике, другие — фундаментальному механизму, третьи — экспери-

ментальным моделям, четвертые — другой популяции или другому материалу. Поэтому результативность поиска определяется не количеством найденных публикаций, а тем, насколько точно сформулирован вопрос.

Вопрос поиска нельзя строить только вокруг названия анализа.

Для научной обоснованности вопрос поиска должен включать не только анализ, но и клиническую роль результата. Если анализ используется для диагностики, нужно искать данные о его связи с наличием или отсутствием состояния. Если он используется для скрининга, нужны данные о возможности выделить группу, требующую дальнейшего подтверждения. Если результат предназначен для мониторинга, важна связь с динамикой процесса или ответом на лечение. Если изделие заявлено для выбора терапии, ключевыми становятся данные о предиктивной роли анализа. Один и тот же анализ в этих разных ролях требует разных поисковых стратегий.

Например, поиск по молекулярному маркеру в онкологии может дать публикации о его частоте в опухолях, связи с прогнозом, роли в выборе терапии, механизме резистентности, методах обнаружения и данных по разным типам материала. Все эти источники могут быть полезны, но не одинаково. Если изделие заявлено для выбора терапии, главными будут данные, связывающие результат с ответом на конкретное лечение или включением пациента в определенный терапевти-

ческий маршрут. Если изделие заявлено для диагностики, приоритет будет у данных, подтверждающих связь маркера с заболеванием или молекулярным подтипом.

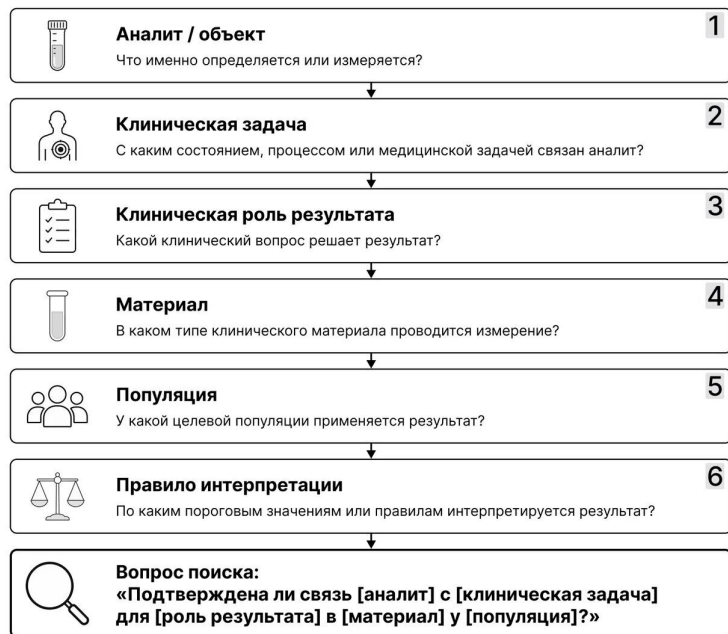


Рисунок 13.1. Формирование запроса литературного поиска для научной обоснованности

Если из вопроса исключены материал, популяция и

клиническая роль, найденные публикации могут оказаться слишком разнородными. Такой массив будет трудно использовать для итогового вывода, потому что он будет содержать данные о разных медицинских ситуациях, которые нельзя механически объединять.

Особое значение имеет тип клинического материала. При установлении научной обоснованности важно понимать, где именно определяется аналит и как это влияет на клинический смысл результата. Данные по сыворотке не всегда применимы к плазме, цельной крови, моче, калу, сухому пятну крови, тканевому материалу или цифровому изображению. В молекулярной онкологии данные по опухолевой ткани не всегда можно переносить на циркулирующую опухолевую ДНК в плазме. В микробиологии обнаружение микроорганизма в стерильном и нестерильном локусе имеет разный клинический смысл. Поэтому материал должен быть учтен уже на этапе формирования поискового запроса.

Целевая популяция также влияет на поиск. Если изделие предназначено для пациентов с симптомами, публикации по уже диагностированной или клинически подозрительной группе могут быть релевантными. Если изделие заявлено для скрининга, нужно искать данные по популяциям с низкой распространенностью состояния, по первичному отбору и подтверждающему маршруту. Если речь идет о новорожденных, детях, беременных, иммунокомпрометированных пациентах, донорах или пациентах на терапии, поиск

должен учитывать именно эту группу. Иначе научная обоснованность будет опираться на данные, которые не полностью соответствуют назначению изделия.

Поиск должен быть воспроизводимым

Поиск должен учитывать и правило интерпретации результата. Для количественных показателей важно искать как связь аналита с состоянием, так и данные о порогах, диапазонах, динамике и клиническом значении изменений. Для качественных тестов важны данные о том, как положительный, отрицательный, сомнительный или пограничный результат связан с клиническим выводом. Для алгоритмических изделий нужно искать данные и по отдельным входным параметрам, и по итоговому индексу, категории риска или правилу классификации.

Схема 13.1 — Элементы поисковой стратегии при установлении научной обоснованности

— **Аналит или объект исследования** — *Что следует учитывать:* Название, синонимы, устаревшие обозначения, международные обозначения; *Пример уточнения:* Белок, ген, вариант, микроорганизм, антиген, антитело, индекс.

— **Клиническое состояние или процесс** — *Что следует учитывать:* Заболевание, синдром, физиологическое состояние, патологический процесс; *Пример уточнения:* Инфекция, опухоль, метаболическое нарушение, иммунный ответ.

— **Клиническая роль результата** — **Что следует учитывать:** Диагностика, скрининг, мониторинг, прогноз, выбор терапии, стратификация; **Пример уточнения:** Диагностика, скрининг, мониторинг, прогноз..

— **Тип материала** — **Что следует учитывать:** Матрица и условия получения образца; **Пример уточнения:** Сыворотка, плазма, ткань, сухое пятно крови, мазок, изображение.

— **Популяция** — **Что следует учитывать:** Возраст, клиническая группа, стадия, статус лечения, риск-группа; **Пример уточнения:** Новорожденные, доноры, симптомные пациенты, пациенты на терапии.

— **Правило интерпретации** — **Что следует учитывать:** Порог, категория, индекс, динамика, алгоритм; **Пример уточнения:** Cut-off, risk score, variant classification, fold change.

— **Источники данных** — **Что следует учитывать:** Базы данных, рекомендации, обзоры, первичные исследования, консенсусы; **Пример уточнения:** PubMed, рекомендации обществ, базы вариантов, регистры.

Выбор источников зависит от зрелости аналита. Для хорошо изученных аналитов важное место занимают клинические рекомендации, систематические обзоры и консенсусные документы. Поиск в этом случае должен выявить, как именно результат используется в медицинской практике и не выходит ли заявленное назначение изделия за пределы

признанной роли анализа. Для нового анализа большее значение приобретают первичные исследования, независимые подтверждения, данные о механизме, ранние клинические работы и исследования, показывающие связь результата с медицинской задачей.

При поиске по новым анализам особенно важно не ограничиваться положительными публикациями. Ранние исследования часто демонстрируют многообещающие результаты, но позднее эффект может оказаться слабее, нестабильнее или ограниченным определенной подгруппой. Поэтому поиск должен быть построен так, чтобы выявлять не только подтверждающие, но и противоречивые или отрицательные данные. Для научной обоснованности важно понимать весь массив сведений, а не только наиболее благоприятные публикации.

Отдельный вопрос — использование специализированных баз данных. В генетике и онкологии большое значение могут иметь базы клинической интерпретации вариантов, базы терапевтической значимости мутаций, руководства профессиональных сообществ и экспертные классификации. В инфекционной диагностике могут использоваться документы по диагностическим алгоритмам и клиническим рекомендациям. В редких заболеваниях — регистры, экспертные сети и данные специализированных центров. Такие источники не заменяют критической оценки, но помогают уточнить, насколько признана клиническая роль анализа.

При формировании поисковой стратегии необходимо учитывать синонимы и изменения терминологии. Один и тот же ген, белок, микроорганизм, синдром или метод может обозначаться по-разному в старых и новых публикациях. В онкологии могут меняться названия маркеров и классификации опухолей. В инфекционной диагностике могут обновляться названия микроорганизмов и таксономия. В генетике могут использоваться разные форматы описания вариантов. Если эти особенности не учесть, поиск может быть неполным.

Важным элементом является период поиска. Для устоявшихся аналитов иногда достаточно сосредоточиться на современных рекомендациях и актуальных обзорах, дополняя их ключевыми первичными исследованиями. Для новых аналитов может потребоваться более полный поиск с момента появления первых данных, чтобы проследить развитие доказательной базы. В быстро меняющихся областях важно учитывать актуальность источников: данные десятилетней давности могут быть недостаточны, если изменились методы, классификации, терапевтические подходы или понимание клинической роли маркера.

Критерии включения и исключения публикаций должны соответствовать вопросу научной обоснованности. Если изделие заявлено для определенной клинической роли, публикации по другой роли могут быть включены только как поддерживающие, но не как основные. Если заявлен опреде-

ленный материал, данные по другому материалу требуют отдельного обоснования. Если целевая популяция узкая, публикации по широкой или иной популяции нужно использовать осторожно. Такой подход помогает избежать ситуации, когда раздел научной обоснованности выглядит объемным, но фактически опирается на нерелевантные источники.

Отбор источников должен быть прозрачным.

Результаты поиска должны быть документированы. В доказательной документации нужно указать, какие базы и источники использовались, какие ключевые слова применялись, какой период охвачен, какие критерии включения и исключения использованы, сколько источников было найдено и какие из них включены в анализ. Но еще важнее не сама техническая фиксация поиска, а итоговая интерпретация: какие данные действительно подтверждают связь анализата с заявленной медицинской задачей, какие имеют только поддерживающее значение, а какие ограничивают вывод.

Литературный поиск при установлении научной обоснованности должен завершаться не списком публикаций, а аргументированным выводом. В этом выводе необходимо показать, подтверждена ли связь анализата с клиническим состоянием, какой тип связи установлен, в какой популяции и материале, насколько данные воспроизводимы и применимы к назначению изделия. Если данные неоднородны, ограничены или противоречивы, это должно быть отражено прямо. Такой вывод затем определяет, можно ли переходить к под-

тверждению аналитической и клинической эффективности в заявленном объеме или требуется дополнительное обоснование.

Итог поиска должен показывать границы имеющихся данных.

В связи с этим, литературный поиск для научной обоснованности — это не библиографическая работа “по теме”, а способ проверить клинический смысл анализа. Он должен быть достаточно широким, чтобы выявить весь значимый массив данных, и достаточно точным, чтобы не смешивать разные клинические роли, материалы и популяции. Его результатом должна стать понятная доказательная позиция: что известно о связи анализа с медицинской задачей, насколько это применимо к изделию и какие ограничения должны быть учтены в дальнейшем.

Итог главы

Литературный поиск при установлении научной обоснованности направлен на подтверждение связи анализа или иного объекта исследования с заявленной медицинской задачей. Такой поиск должен строиться и вокруг названия анализа, и вокруг клинической роли результата, типа материала, целевой популяции и правила интерпретации. Его задача — выявить надежные, актуальные и применимые источники, критически оценить их и сформулировать вывод о том, достаточно ли данных для подтверждения научной обоснованности. Сильный литературный поиск не просто собирает

публикации, а показывает, какой клинический смысл имеет результат изделия и где проходят границы этого вывода.

Глава 14. Систематические обзоры в подтверждении научной обоснованности

Систематический обзор нужен не всегда, но иногда он принципиален.

Систематический обзор занимает особое место в подтверждении научной обоснованности медицинского изделия для диагностики *in vitro*. Его задача состоит не просто в том, чтобы собрать публикации по аналиту, а в том, чтобы показать, насколько устойчива, воспроизводима и применима связь аналита или иного объекта исследования с заявленной медицинской задачей. Если обычный литературный обзор может дать общее представление о теме, то систематический обзор должен дать проверяемый ответ на конкретный доказательный вопрос.

Для научной обоснованности это особенно важно. Связь аналита с заболеванием, прогнозом, ответом на терапию, скрининговой задачей или иной клинической ролью часто формируется не одной публикацией, а совокупностью исследований. Отдельная работа может показать статистически значимую ассоциацию, но этого недостаточно для уверенного вывода. Необходимо понять, воспроизводится ли эта связь в других исследованиях, сохраняется ли она в реле-

вантной популяции, относится ли к заявленному материалу и не зависит ли от случайного выбора порога, метода или клинической группы.

Систематический обзор особенно полезен тогда, когда анализ находится между двумя состояниями: он уже не является полностью новым, но его клиническая роль еще требует аккуратного обоснования. Например, по маркеру может существовать несколько исследований, часть из которых показывает связь с заболеванием, а часть дает менее убедительные результаты. В такой ситуации простой набор ссылок легко становится селективным: можно выбрать только благоприятные публикации и создать впечатление доказанности. Систематический обзор снижает этот риск, потому что требует заранее заданного вопроса, прозрачной стратегии поиска, критериев включения и исключения, оценки качества источников и обобщения всех релевантных данных.



Рисунок 14.1. Логика систематического обзора при подтверждении научной обоснованности

Схема показывает, что систематический обзор должен начинаться не с поиска “всего по аналиту”, а с назначения изделия. Если изделие заявлено для диагностики, обзор должен отвечать на вопрос о диагностической связи. Если для скрининга — о пригодности анализата для первичного отбора группы, требующей подтверждения. Если для мониторинга

— о связи результата с динамикой состояния или ответом на лечение. Если для выбора терапии — о предиктивной роли результата. Без такой привязки обзор может быть формально обширным, но доказательно размытым.

Не каждый раздел научной обоснованности требует полноценного систематического обзора. Для хорошо изученного анализа в стандартной клинической роли часто достаточно актуальных клинических рекомендаций, признанных систематических обзоров, консенсусных документов и ключевых источников, подтверждающих устоявшееся применение. Например, в большинстве случаев не нужно заново проводить систематический обзор связи глюкозы с нарушениями углеводного обмена или HbA_{1c} с инфекцией вирусом гепатита В. Но если заявляется новая роль результата, новый материал, новая популяция или новый алгоритм интерпретации, необходимость более строгого обзора возрастает.

Для новых и спорных анализов систематический обзор становится особенно важным. Он позволяет отделить устойчивую клиническую связь от предварительных или случайных находок. В биомаркерных исследованиях нередко встречаются ранние публикации с многообещающими результатами, которые затем не подтверждаются в независимых выборках. Это может быть связано с малым объемом исходного исследования, селекцией пациентов, ретроспективным дизайном, неоптимальным выбором группы сравнения, различиями в материале или публикационным смещением. Си-

стематический обзор помогает увидеть как положительные данные, так и их ограничения.

Схема 14.1 — Когда достаточно литературного анализа, а когда желателен систематический обзор
— Хорошо изученный анализ в стандартной роли — *Обычный литературный анализ может быть достаточен: Если роль подтверждена рекомендациями, консенсусом и устоявшейся практикой; Систематический обзор особенно желателен: Если есть противоречия в современных данных или изменилось назначение.*

— Хорошо изученный анализ в новой роли — *Обычный литературный анализ может быть достаточен: Только как предварительное обоснование; Систематический обзор особенно желателен: Нужен для оценки применимости данных к новой роли.*

— Новый анализ — *Обычный литературный анализ может быть достаточен: Обычно недостаточен; Систематический обзор особенно желателен: Желателен для оценки устойчивости и качества данных.*

— Новый тип материала — *Обычный литературный анализ может быть достаточен: Может описать общую связь анализа с состоянием; Систематический обзор особенно желателен: Нужен для оценки переноса данных на новый материал.*

— Алгоритмический индекс или модель — *Обычный литературный анализ может быть достаточен:*

Может описать клинический смысл входных параметров; Систематический обзор особенно желателен: Нужен для оценки итогового алгоритмического результата.

— Противоречивые публикации — Обычный литературный анализ может быть достаточен: Недостаточен без анализа причин расхождений; Систематический обзор особенно желателен: Нужен для системной оценки неоднородности и риска смещения.

Вопрос обзора должен быть уже, чем тема анализа.

При подготовке систематического обзора важно заранее сформулировать вопрос. В отличие от общего обзора темы, вопрос должен быть достаточно конкретным. Он должен включать анализ или объект исследования, клиническое состояние или процесс, клиническую роль результата, тип материала и целевую популяцию. Если обзор строится слишком широко, в него попадают исследования с разными задачами: диагностика, прогноз, мониторинг, выбор терапии, разные материалы, разные методы и разные пороги. Такой массив трудно интерпретировать, а вывод о научной обоснованности становится неопределенным.

Систематический поиск должен выявлять не только удобные публикации.

Следующий этап — систематический поиск. Он должен быть воспроизводимым: должны быть указаны базы данных и источники, период поиска, поисковые запросы, синонимы, критерии включения и исключения. При необходимости ис-

пользуются клинические рекомендации, документы профессиональных обществ, специализированные базы данных, регистры, базы генетических вариантов или базы клинической значимости молекулярных изменений. Важно, чтобы поиск выявлял благоприятные публикации, а также данные, которые ограничивают или ставят под сомнение связь анализа с заявленной задачей.

Критическая оценка отличает обзор от перечня литературы.

После отбора источников проводится критическая оценка. Именно здесь систематический обзор отличается от простого перечисления литературы. Для каждого включенного исследования нужно оценить риск смещения, применимость к назначению изделия, качество определения клинического статуса, тип материала, метод сравнения, выборку, статистическую обработку и полноту представления результатов. Исследования, формально относящиеся к одному анализу, могут сильно отличаться по доказательной пригодности. Одни будут прямо подтверждать заявленную роль, другие — иметь только поддерживающее значение, третьи — быть неприменимыми из-за иной популяции или материала.

При такой оценке целесообразно использовать признанные методические инструменты. Для исследований диагностической точности может применяться QUADAS-2, позволяющий структурировать оценку риска смещения и применимости данных. Если сравниваются диагностические тесты

между собой, может быть полезен QUADAS-C. Для прогноз-
стических или предиктивных моделей применим PROBAST.
Подходы GRADE могут использоваться для оценки уверен-
ности в совокупности доказательств, когда это уместно для
поставленного вопроса. PRISMA, в свою очередь, служит
ориентиром для прозрачного представления систематиче-
ского обзора, включая описание поиска, отбора источников
и причин исключения.

Важно и обратное: для чисто ассоциативных, когортных
или механистических исследований — а именно они часто
составляют основную часть литературы о новом анализе на
раннем этапе, до появления полноценных диагностических
исследований, — домены QUADAS-2 не рассчитаны и не
должны применяться формально, поскольку в таких работах
отсутствует «индексный тест» в том смысле, в котором его
предполагает этот инструмент.



Рисунок 14.2. Выбор инструмента оценки качества и применимости доказательств

Эти инструменты не должны превращать раздел научной обоснованности в формальное упражнение. Они помогают сделать оценку более прозрачной, но не заменяют содержательного вывода. Главный вопрос остается прежним: подтверждают ли включенные данные связь аналита с заявленной медицинской задачей именно для того материала, популяции и клинической роли, которые указаны в назначении изделия? Если инструмент оценки качества показывает низкий риск смещения, но исследование выполнено в другой по-

пуляции или для другой роли результата, его применимость к данному изделию все равно может быть ограниченной.

Особое внимание в систематическом обзоре уделяется неоднородности данных. Исследования могут различаться по популяции, стадии заболевания, материалу, методу определения аналита, порогам, референтной рамке и конечным точкам. Такая неоднородность не всегда является недостатком: иногда она показывает, что связь аналита сохраняется в разных условиях. Но если результаты существенно расходятся, необходимо объяснить причины. Возможно, аналит имеет значение только в определенной подгруппе, только при определенном пороге, только в одном материале или только для одной клинической роли.

Метаанализ может быть полезным, но он не является обязательной частью каждого систематического обзора. В МИ ИВД исследования часто неоднородны по дизайну, материалу, порогам и референтным подходам. Если такие данные механически объединить, итоговая числовая оценка может выглядеть убедительно, но фактически скрывать методологические различия. Поэтому решение о метаанализе должно быть обосновано. Иногда качественное обобщение с подробным анализом различий между исследованиями будет более корректным, чем формальное объединение показателей.

Для научной обоснованности важен средний эффект, а также клиническая интерпретируемость связи. Если исследования показывают статистически значимую связь аналита

с заболеванием, но величина эффекта мала, пороги нестабильны, популяции различаются, а клинический маршрут не определен, вывод должен быть осторожным. Напротив, если связь подтверждена в нескольких независимых работах, соответствует биологическому механизму, отражена в рекомендациях и применима к заявленному материалу и популяции, научная обоснованность будет значительно сильнее.

Отдельный вопрос возникает при оценке алгоритмических показателей и прогностических моделей. В таких случаях систематический обзор должен быть направлен не только на входные параметры, но и на итоговую модель или индекс. Если модель объединяет несколько признаков и выдает категорию риска, обзор должен оценивать, подтверждена ли клиническая значимость именно такой категории. Здесь особенно важны внешняя проверка, риск переобучения, применимость выборки разработки к реальной популяции, прозрачность модели и устойчивость результата. Именно поэтому для таких исследований полезны инструменты, ориентированные на оценку моделей, например PROBAST.

В систематическом обзоре необходимо учитывать и публикационное смещение. Положительные результаты чаще публикуются, чем отрицательные или нейтральные. Для новых биомаркеров это особенно опасно: ранний массив литературы может выглядеть более убедительно, чем он является на самом деле. Если в обзоре не рассматриваются неопубликованные данные, отрицательные исследования, малые вы-

борки и отсутствие независимого подтверждения, вывод о научной обоснованности может быть завышен.

Заподозрить смещение публикаций можно визуальными и формальными методами (например, асимметрия funnel plot, тест Egger), применимыми при наличии достаточного числа независимых исследований для метаанализа; расчётные аспекты этих методов не рассматриваются в настоящей части и относятся к области статистического анализа совокупности доказательств.

Обзор должен завершаться выводом о научной обоснованности.

Завершаться систематический обзор должен не формальным перечислением найденных исследований, а содержательным выводом. В этом выводе необходимо указать, подтверждена ли связь аналита с клиническим состоянием или медицинской задачей, насколько она устойчива, в каких популяциях и материалах показана, какая клиническая роль результата подтверждена, каковы основные ограничения и можно ли использовать эти данные для заявленного назначения изделия. Если данные достаточны только для части назначения, это должно быть прямо указано.

Соответственно, систематический обзор в подтверждении научной обоснованности выполняет функцию фильтра и синтеза. Он помогает отделить надежные данные от предварительных, устойчивую связь от случайной ассоциации, применимые источники от нерелевантных. Его ценность не в

количестве включенных публикаций, а в том, насколько ясно он показывает: имеет ли результат изделия медицинский смысл в заявленном применении и насколько уверенно это можно утверждать.

Итог главы

Систематический обзор является важным инструментом подтверждения научной обоснованности, особенно для новых, спорных или расширенно применяемых аналитов. Он позволяет оценить устойчивость связи аналита с клинической задачей, применимость данных к материалу и популяции, риск смещения, неоднородность исследований и ограничения вывода. Для прозрачного представления обзора могут использоваться ориентиры PRISMA, а для критической оценки источников — QUADAS-2, QUADAS-C, PROBAST и подходы GRADE. При этом методические инструменты не заменяют главного экспертного вопроса: подтверждает ли совокупность данных научную обоснованность аналита именно для заявленного назначения медицинского изделия для диагностики *in vitro*.

Глава 15. Критическая оценка и анализ данных о научной обоснованности

Критическая оценка показывает, каким данным можно доверять.

После того как источники данных найдены и отобраны, начинается один из самых важных этапов подтверждения научной обоснованности — критическая оценка. Ее задача состоит не в том, чтобы просто разделить публикации на “подходящие” и “неподходящие”, а в том, чтобы понять, насколько полученные данные действительно подтверждают связь анализатора или иного объекта исследования с заявленной медицинской задачей. Для медицинского изделия для диагностики *in vitro* важно не только наличие публикаций, но и качество, применимость и интерпретируемость этих публикаций.

На практике раздел научной обоснованности часто ослабляется именно на этом этапе. В доказательную документацию включают большое количество источников, но не объясняют, какие из них являются ключевыми, какие имеют поддерживающее значение, какие ограничены по применимости, а какие вообще не должны использоваться для заявленного вывода. В результате раздел выглядит объемным, но

неубедительным. Критическая оценка нужна для того, чтобы превратить массив литературы и иных данных в аргументированную доказательную позицию.

Даже качественное исследование может быть нерелевантным.

Первый вопрос критической оценки — относится ли источник к заявленному назначению изделия. Даже качественное исследование может быть слабым доказательством, если оно выполнено в другой популяции, на другом материале, для другой клинической роли или с другим правилом интерпретации результата. Например, исследование маркера у пациентов с уже подтвержденным заболеванием может быть полезно для понимания биологии процесса, но не обязательно подходит для обоснования скрининга бессимптомной популяции. Данные по опухолевой ткани не всегда применимы к плазме, а данные по диагностической роли биомаркера не всегда подтверждают его предиктивное значение.

Второй вопрос — насколько надежно само исследование. Здесь оцениваются дизайн, размер и структура выборки, критерии включения и исключения, способ определения клинического статуса, наличие подходящего референтного стандарта или сравнительного подхода, качество статистического анализа, полнота представления результатов и риск систематической ошибки. Исследование может быть релевантным по теме, но слабым по методологии. В таком случае оно не должно нести основную доказательную нагрузку.

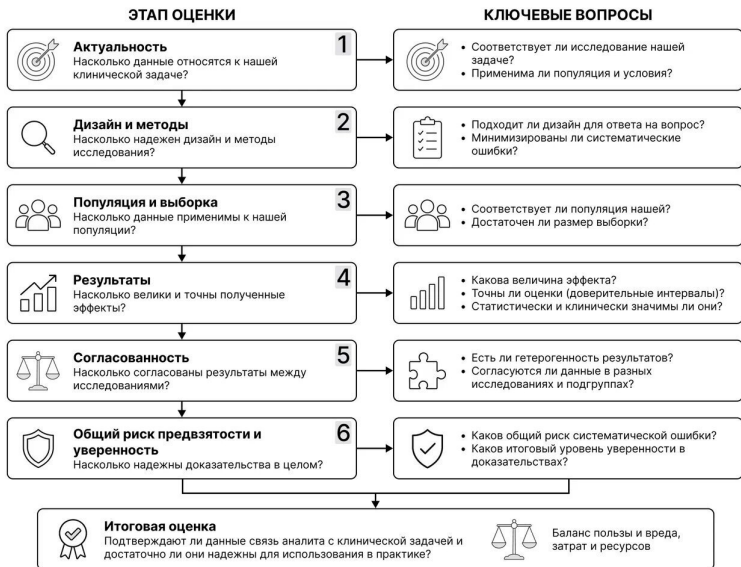


Рисунок 15.1. Логика критической оценки данных о научной обоснованности

Рисунок 15.1 показывает, что критическая оценка должна идти в двух направлениях одновременно. С одной стороны, нужно оценить внутреннее качество исследования. С другой — его применимость к конкретному изделию. Высококачественное исследование, выполненное в нерелевантной популяции, не может автоматически подтверждать заявлен-

ное назначение. И наоборот, исследование, близкое к нужной клинической задаче, но методологически слабое, может иметь только поддерживающий вес.

Риск смещения и применимость нужно оценивать отдельно.

Одним из центральных понятий критической оценки является риск смещения. Смещение возникает тогда, когда дизайн или выполнение исследования систематически искажает результат. Например, если в исследование включены только очевидно положительные и очевидно отрицательные случаи, оценка связи анализа с состоянием может выглядеть более убедительной, чем в реальной клинической популяции. Если клинический статус установлен не у всех участников одинаково, возникает риск неправильной классификации. Если порог интерпретации выбран после просмотра данных, результат может быть переоценен. Если исследование выполнено только в одном центре и не проверено независимо, его переносимость может быть ограничена.

Для исследований диагностической точности особенно важно, как определен референтный стандарт. Если референтный подход сам неточен, неполон или зависит от результата оцениваемого теста, вывод о связи анализа с клиническим состоянием становится менее надежным. В МИ ИВД это встречается часто: не для всех состояний существует идеальный “золотой стандарт”. В микробиологии, онкологии, генетике, редких заболеваниях и скрининговых про-

граммах референтная рамка может быть составной, включать клиническое наблюдение, экспертное заключение, подтверждающий метод или комбинацию данных. Это допустимо, но должно быть описано и оценено.

Особое значение имеет применимость данных. Исследование может быть методологически сильным, но относиться к иной медицинской ситуации. Например, данные о диагностической точности теста у госпитализированных пациентов с высокой вероятностью заболевания не всегда применимы к амбулаторной популяции. Данные по взрослым не всегда применимы к новорожденным. Данные по свежему материалу не всегда применимы к архивному. Данные по одной версии алгоритма не всегда применимы к другой. Поэтому применимость должна оцениваться отдельно от методологического качества.

Для структурированной оценки качества могут использоваться признанные инструменты. Они не заменяют анализа, но помогают сделать его прозрачным и воспроизводимым. В исследованиях диагностической точности часто применяется QUADAS-2. Если сравниваются два или более диагностических теста, может быть полезен QUADAS-C. Для прогностических и предиктивных моделей применяется PROBAST. Для оценки уверенности в совокупности доказательств могут использоваться подходы GRADE. Эти инструменты особенно полезны, когда раздел научной обоснованности опирается на несколько исследований разного ка-

чества.

Схема 15.1 — Инструменты критической оценки данных

— **QUADAS-2** — *Для чего применяется:* Оценка исследований диагностической точности; *Как помогает при оценке научной обоснованности:* Помогает оценить риск смещения и применимость данных по пациентам, индексному тесту, референтному стандарту и потоку исследования.

— **QUADAS-C** — *Для чего применяется:* Оценка сравнительных исследований диагностических тестов; *Как помогает при оценке научной обоснованности:* Помогает оценить данные, где важно сравнить один диагностический подход с другим.

— **PROBAST** — *Для чего применяется:* Оценка прогностических и предиктивных моделей; *Как помогает при оценке научной обоснованности:* Полезен для алгоритмов, индексов риска, прогностических моделей и предиктивных инструментов.

— **GRADE** — *Для чего применяется:* Оценка уверенности в совокупности доказательств; *Как помогает при оценке научной обоснованности:* Помогает оценить общую надежность вывода с учетом риска смещения, непрямоты, неоднородности, неточности и публикационного смещения.

— **PRISMA** — *Для чего применяется:* Прозрачное

*представление систематического обзора ; **Как помогает при оценке научной обоснованности:** Помогает описать поиск, отбор, исключение источников и структуру обзора, но не является инструментом оценки качества .*

Важно понимать границы применения этих инструментов. QUADAS-2 не отвечает на вопрос, является ли анализ клинически значимым сам по себе; он помогает оценить качество исследований диагностической точности. PROBAST не доказывает медицинскую ценность модели; он помогает выявить риск смещения в разработке и проверке прогностической или предиктивной модели. GRADE не заменяет содержательный вывод эксперта; он помогает оценить уверенность в совокупности доказательств. Поэтому инструменты должны использоваться как способ структурировать критическую оценку, а не как механическая процедура.

Тип клинической роли меняет критерии оценки.

Критическая оценка должна учитывать тип клинической связи. Для диагностической роли важны популяция пациентов, референтный стандарт, спектр заболевания и контрольные группы. Для скрининговой роли — распространенность состояния, последствия ложноположительных и ложноотрицательных результатов, подтверждающий маршрут и пригодность данных для первичного отбора. Для мониторинговой роли — динамика показателя, воспроизводимость изменений и связь с клиническим течением. Для прогностической роли — связь с исходами и длительность наблюдения. Для

предиктивной роли — связь результата с ответом на конкретное вмешательство.

Например, исследование, показывающее различие уровня маркера между пациентами с заболеванием и здоровыми лицами, может быть полезным для первичного понимания диагностической связи. Но оно не доказывает прогностическую роль, если в нем не оценивались исходы. Оно не доказывает предиктивную роль, если не было данных о терапии и ответе на нее. Оно не доказывает мониторинговую роль, если не анализировалась динамика. Поэтому критическая оценка всегда должна сопоставлять дизайн исследования с той ролью результата, которая заявлена для изделия.

Для новых аналитов критическая оценка особенно важна. Ранние исследования часто завышают эффект. Это может происходить из-за малой выборки, ретроспективного отбора, отсутствия внешней проверки, публикационного смещения, использования удобных контрольных групп или выбора порога после анализа данных. В таких случаях связь аналита с состоянием может выглядеть убедительно, но при независимой проверке оказывается слабее. Поэтому для новых маркеров следует особенно внимательно оценивать независимость подтверждения и воспроизводимость результатов.

Для хорошо изученных аналитов критическая оценка тоже необходима, но ее акцент иной. Здесь обычно не нужно заново доказывать базовую связь аналита с состоянием. Важнее оценить, не выходит ли заявленное назначение изде-

лия за пределы признанного применения. Если изделие использует хорошо известный аналит в новой популяции, новом материале, новом клиническом сценарии или с новым алгоритмом интерпретации, даже хорошо установленная научная обоснованность требует дополнительной оценки применимости.

Противоречия нельзя игнорировать

Противоречивые данные требуют отдельного анализа. Если одни исследования подтверждают связь аналита с клинической задачей, а другие нет, нельзя просто выбрать наиболее удобные источники. Нужно понять, чем они различаются. Возможно, в них использовались разные материалы, разные пороги, разные стадии заболевания, разные критерии клинического статуса или разные методы определения аналита. Иногда противоречия показывают, что аналит имеет значение только в определенной подгруппе. Иногда — что связь недостаточно устойчива для заявленного применения.

Неоднородность данных сама по себе не всегда является недостатком. Если разные исследования в разных условиях дают сходные результаты, это может усиливать уверенность в научной обоснованности. Но если результаты расходятся, а причины неоднородности неясны, вывод должен быть осторожным. Особенно опасно объединять неоднородные данные в один общий вывод без объяснения различий. Для МИ ИВД это критично, потому что материал, популяция и клиническая роль результата непосредственно влияют на интер-

претацию.

Схема 15.2 — Типовые ограничения данных о научной обоснованности

— **Нерелевантная популяция** — *Как проявляется:* Данные получены у другой группы пациентов или обследуемых лиц; *Как влияет на вывод:* Ограничивает переносимость вывода на заявленную популяцию.

— **Иной тип материала** — *Как проявляется:* Исследован другой биологический материал или формат данных; *Как влияет на вывод:* Требуется обоснования переноса научной обоснованности.

— **Другая клиническая роль** — *Как проявляется:* Данные относятся к диагностике, а изделие заявлено для прогноза, мониторинга или выбора терапии; *Как влияет на вывод:* Не подтверждает заявленную роль результата.

— **Малый объем выборки** — *Как проявляется:* Исследование включает небольшое число случаев; *Как влияет на вывод:* Повышает неопределенность и риск случайного эффекта.

— **Отсутствие независимого подтверждения** — *Как проявляется:* Связь показана только в одной группе или одном центре; *Как влияет на вывод:* Снижает уверенность в воспроизводимости.

— **Неясный референтный подход** — *Как проявляется:* Клинический статус установлен неполно или неодинаково; *Как влияет на вывод:* Ослабляет достоверность связи

аналита с состоянием.

— **Выбор порога после анализа данных** — *Как проявляется:* Порог определен ретроспективно по лучшему результату; *Как влияет на вывод:* Может завышать диагностическую или прогностическую ценность.

— **Публикационное смещение** — *Как проявляется:* Доступны преимущественно положительные исследования; *Как влияет на вывод:* Может создавать ложное впечатление устойчивой связи.

— **Источник финансирования и конфликт интересов** — *Как проявляется:* Исследование профинансировано производителем оцениваемого изделия или заинтересованной стороной, конфликт интересов авторов не раскрыт; *Как влияет на вывод:* Может создавать систематический сдвиг в сторону благоприятного результата.

— **Устаревшие данные** — *Как проявляется:* Изменились классификации, методы, клинические маршруты или терапия; *Как влияет на вывод:* Требуется проверки актуальности вывода.

Критическая оценка должна завершаться описанием ограничений, а также определением доказательного веса источников. Одни источники могут быть основными, потому что прямо относятся к назначению изделия и имеют хорошее качество. Другие могут быть поддерживающими: они помогают понять механизм, контекст или общую клиническую роль, но не закрывают ключевой вопрос. Третьи могут

быть ограниченными и использоваться только с оговорками. Некоторые источники после анализа должны быть признаны неприменимыми к заявленному назначению.

Такой подход особенно важен при подготовке итогового вывода о научной обоснованности. Вывод должен быть соразмерен силе данных. Если связь подтверждена надежными и применимыми источниками, можно формулировать уверенный вывод. Если данные ограничены, неоднородны или относятся только к части назначения, вывод должен быть ограниченным. Если данные предварительные, противоречивые или неприменимые, раздел научной обоснованности не должен создавать видимость доказанности.

Для алгоритмических и программных изделий критическая оценка должна дополнительно учитывать данные, на которых разрабатывался и проверялся алгоритм. Если итоговый результат формируется моделью, необходимо оценить, была ли выборка разработки репрезентативной, отделены ли данные обучения от данных проверки, была ли внешняя валидация, нет ли переобучения, как алгоритм работает в пограничных случаях и применим ли он к реальной популяции. Здесь особенно полезны инструменты оценки прогностических и предиктивных моделей, но итоговый вопрос остается тем же: имеет ли алгоритмический результат доказанный клинический смысл для заявленного применения.

Оценка должна завершаться определением доказательного веса.

Критическая оценка данных о научной обоснованности должна быть прозрачной. В доказательной документации желательно указать источники, а также показать, почему одни из них признаны ключевыми, другие — поддерживающими, а третьи — ограниченными. Такая прозрачность облегчает экспертную оценку и снижает риск спора о том, почему конкретные публикации были включены или исключены из доказательного вывода.

В конечном счете критическая оценка нужна для того, чтобы не завысить и не занижить научную обоснованность. Завышение возникает, когда предварительные или косвенные данные подаются как окончательное доказательство. Занижение — когда хорошо установленные данные необоснованно требуют повторного доказательства. Задача состоит в балансе: признать то, что уже надежно известно, показать применимость этих данных к назначению изделия и честно обозначить вопросы, которые требуют дополнительного подтверждения.

Итог главы

Критическая оценка данных о научной обоснованности позволяет определить, насколько надежно и применимо подтверждена связь анализа с заявленной медицинской задачей. Она включает оценку риска смещения, качества дизайна, релевантности популяции и материала, референтного подхода, независимости подтверждения, неоднородности и противоречий. Инструменты QUADAS-2, QUADAS-C, PROBAST,

GRADE и PRISMA помогают структурировать эту работу, но не заменяют экспертного вывода о применимости данных к назначению конкретного медицинского изделия для диагностики *in vitro*. Сильный раздел научной обоснованности должен не просто перечислять источники, а показывать их доказательный вес, ограничения и границы итогового вывода.

Глава 16. Документирование научной обоснованности в составе доказательной документации

Документирование должно показывать логику вывода.

Документирование научной обоснованности является завершающим этапом работы с первым элементом клинических доказательств медицинского изделия для диагностики *in vitro*. На предыдущих этапах устанавливается, что представляет собой аналит или иной объект исследования, с каким клиническим состоянием или медицинской задачей он связан, какие источники подтверждают эту связь, насколько они надежны и применимы. В доказательной документации все это должно быть представлено не как набор разрозненных ссылок, а как связная аргументация.

Главная задача раздела о научной обоснованности состоит в том, чтобы показать, почему результат изделия имеет медицинский смысл в пределах заявленного назначения. Поэтому такой раздел должен начинаться не с перечня публикаций и не с описания метода, а с ясного указания на назначение изделия и клиническую роль результата. Только после этого можно обоснованно объяснять, какие данные подтверждают связь аналита с соответствующим состоянием, про-

цессом или медицинской задачей.

На практике слабость раздела о научной обоснованности часто связана с неправильной последовательностью изложения. Сначала приводится обзор литературы, затем перечисляются характеристики метода, затем делается общий вывод о клинической значимости анализа. Но при этом остается неясным, относится ли эта литература к заявленному назначению изделия, тому же материалу, той же популяции и той же клинической роли. Такой раздел может выглядеть объемным, но не выполнять доказательную функцию.

Сначала назначение, затем анализ и клиническая роль.

Правильная логика обратная. Сначала нужно определить, что именно заявлено для изделия. Затем — какой анализ или объект исследования лежит в основе результата. Далее — какую клиническую роль выполняет этот результат. После этого — какие источники подтверждают соответствующую связь. И только затем — насколько эти источники надежны, применимы и достаточны для вывода.



Рисунок 16.1. Документирование научной обоснованности как доказательная линия

Рисунок 16.1 показывает, что документирование научной обоснованности должно быть построено как доказательная линия. Если из нее выпадает хотя бы одно звено, вывод становится уязвимым. Например, можно хорошо описать связь аналита с заболеванием, но не указать, что изделие заявлено для мониторинга, а не для диагностики. Можно привести ка-

чественные публикации по сыворотке, но не объяснить, почему они применимы к сухому пятну крови. Можно показать клиническое значение входных параметров алгоритма, но не обосновать итоговый индекс или категорию риска.

Раздел должен описывать не только аналит, но и его клиническую роль.

Раздел о научной обоснованности должен ясно описывать аналит или иной объект исследования. Для классических лабораторных тестов это может быть белок, гормон, метаболит, антиген, антитело, нуклеиновая кислота, микроорганизм или генетический вариант. Для морфологических и цифровых изделий объектом может быть клеточный признак, зона окрашивания, структура изображения или количественный морфометрический показатель. Для алгоритмических изделий объектом доказательной оценки может быть вычисляемый индекс, категория риска, автоматическая классификация или иной результат программной обработки.

Важно не ограничиваться названием аналита. Нужно объяснить, почему именно этот объект связан с медицинской задачей. Например, если изделие определяет инфекционный маркер, следует показать, что его обнаружение имеет смысл в данном материале и клиническом сценарии. Если изделие определяет онкологический биомаркер, необходимо указать, является ли его роль диагностической, прогностической, предиктивной или мониторинговой. Если изделие формирует алгоритмический вывод, нужно показать, что

именно итоговая категория результата имеет клиническое значение.

Следующий элемент — описание клинической роли результата. Это один из наиболее важных фрагментов раздела. Формулировка “аналит связан с заболеванием” недостаточна. Необходимо указать, используется ли результат для подтверждения или исключения состояния, первичного скрининга, мониторинга динамики, прогноза, выбора терапии, стратификации риска или маршрутизации пациента. Если клиническая роль не определена, невозможно оценить, достаточны ли приведенные источники.

Для хорошо изученного аналита раздел научной обоснованности может быть относительно компактным, но не формальным. В нем достаточно показать, что клиническая роль аналита устоялась, подтверждена рекомендациями, систематическими обзорами, консенсусными документами или длительной практикой. Однако даже в такой ситуации нужно обязательно указать, что заявленное назначение изделия не выходит за пределы этой устоявшейся роли. Если выходит, требуется дополнительное обоснование.

Для нового или недостаточно изученного аналита раздел должен быть более развернутым. В нем необходимо описать литературный поиск, источники данных, характер связи аналита с состоянием, качество исследований, воспроизводимость результатов, противоречия, ограничения и применимость к заявленному назначению. В таком случае просто-

го утверждения о перспективности маркера недостаточно. Нужно показать, что связь имеет достаточную доказательную опору для той роли, которую изделие должно выполнять.

Источники должны быть представлены как доказательные блоки

Особое внимание следует уделять источникам данных. В разделе научной обоснованности они должны быть представлены не как библиографический список, а как доказательные блоки. Одни источники могут подтверждать биологическую основу связи, другие — клиническую роль результата, третьи — значение порога или правила интерпретации, четвертые — применимость к конкретной популяции или материалу. Такая группировка делает текст понятным и позволяет увидеть, какие элементы научной обоснованности подтверждены сильнее, а какие слабее.

Например, для скринингового изделия часть источников может подтверждать связь маркера с заболеванием, другая часть — значение раннего выявления, третья — правила первичного отбора и подтверждающего этапа. Для изделия, связанного с выбором терапии, одни источники могут подтверждать биологическую роль маркера, другие — связь с ответом на препарат, третьи — клинические рекомендации по использованию результата. Для алгоритмического изделия нужно отдельно показать источники по входным параметрам и источники по итоговому алгоритмическому выво-

ду.

Применимость данных нужно показать явно.

Документирование научной обоснованности обязательно должно включать оценку применимости данных. Это означает, что для каждого ключевого блока источников нужно показать, насколько он относится к заявленному материалу, популяции, клинической роли и правилу интерпретации. Если данные полностью соответствуют назначению изделия, это следует указать. Если они отличаются, необходимо объяснить, почему перенос допустим или почему вывод должен быть ограничен.

Типичный пример — перенос данных между материалами. Публикации по сыворотке не всегда подтверждают применение в плазме, сухом пятне крови, моче, кале или тканевом материале. Данные по опухолевой ткани не всегда применимы к циркулирующей опухолевой ДНК. Данные по стерильному локусу не равны данным по нестерильному локусу. Если такие различия есть, они должны быть прямо отражены в разделе научной обоснованности.

То же касается популяции. Если данные получены у пациентов с выраженной клинической картиной, а изделие заявлено для скрининга бессимптомных лиц, применимость источников ограничена. Если данные получены у взрослых, а изделие заявлено для новорожденных или детей, нужно дополнительное обоснование. Если исследования выполнены у пациентов до лечения, а изделие используется для монито-

ринга терапии, это также требует отдельного анализа.

Для количественных и пороговых тестов необходимо документировать и связь аналита с состоянием, и клинический смысл порога или диапазона. Если результат интерпретируется как положительный, отрицательный, пограничный или требующий повторного исследования, раздел научной обоснованности должен объяснять, почему такое правило интерпретации имеет медицинский смысл. Порог не должен выглядеть как чисто техническая настройка, если именно от него зависит клиническое действие.

Для программных и алгоритмических изделий нужно фиксировать, какой именно результат является предметом научной обоснованности. Если программа только оформляет отчет или передает данные, научная обоснованность относится прежде всего к аналиту и медицинской роли результата. Но если программа классифицирует изображение, фильтрует варианты, рассчитывает индекс риска или автоматически формирует категорию результата, научная обоснованность должна относиться к итоговому выводу алгоритма. Нельзя ограничиться ссылками на клиническую значимость отдельных входных данных.

В разделе научной обоснованности также необходимо отражать ограничения. Ограничения не ослабляют документ автоматически; напротив, они делают вывод честным и экспертно устойчивым. Если данные применимы только к определенной популяции, это нужно указать. Если связь под-

тверждена только для одного материала, нельзя распространять ее на другие материалы без обоснования. Если роль анализа подтверждена для диагностики, но не для прогноза или выбора терапии, это должно быть прямо отражено. Если данные по новому анализу предварительные, вывод должен быть осторожным.

Схема 16.1 — Типовые ошибки документирования научной обоснованности

— **Перечисление публикаций без аргументации** — *Как проявляется:* Раздел выглядит как библиография без связи с назначением изделия; *Почему это критично:* Непонятно, какой вывод подтверждают источники.

— **Обоснование анализа “вообще”** — *Как проявляется:* Описана связь маркера с заболеванием без указания клинической роли; *Почему это критично:* Нельзя оценить применимость к диагностике, скринингу, мониторингу или выбору терапии.

— **Подмена научной обоснованности аналитическими характеристиками** — *Как проявляется:* Вместо клинического смысла анализа описаны предел обнаружения, точность или воспроизводимость; *Почему это критично:* Не закрыт вопрос, зачем анализ определяется.

— **Перенос данных на другой материал без обоснования** — *Как проявляется:* Данные по одному типу образца используются для другого; *Почему это критично:* Ме-

няется матрица, стабильность и клиническая интерпретация.

— **Перенос данных на другую популяцию без обоснования** — *Как проявляется:* Данные по симптомным пациентам используются для скрининга бессимптомных лиц; *Почему это критично:* Меняются распространенность, предтестовая вероятность и риск ошибки.

— **Смещение клинических ролей** — *Как проявляется:* Данные по диагностике используются для прогноза или выбора терапии; *Почему это критично:* Разные роли требуют разных доказательств.

— **Игнорирование порога или правила интерпретации** — *Как проявляется:* Обоснован анализ, но не обосновано клиническое значение категории результата; *Почему это критично:* Непонятно, почему результат должен приводить к конкретному действию.

— **Обоснование входных параметров вместо алгоритма** — *Как проявляется:* Описаны отдельные показатели, но не итоговый индекс или категория риска; *Почему это критично:* Клинический смысл алгоритмического результата остается недоказанным.

— **Отсутствие ограничений** — *Как проявляется:* Вывод сформулирован шире, чем позволяют данные; *Почему это критично:* Возникает разрыв между доказательствами и заявленным назначением.

Итоговый вывод не должен быть слишком общим.

Итоговый вывод о научной обоснованности должен быть сформулирован отдельно и ясно. Он не должен быть слишком общим. Недостаточно написать: “научная обоснованность подтверждена”. Нужно указать, для какого анализа или объекта исследования она подтверждена, с какой клинической задачей он связан, для какой роли результата, в каком материале и популяции, на основании каких источников и с какими ограничениями.

Например, корректный вывод может быть построен так: определяемый анализ имеет подтвержденную связь с соответствующим клиническим состоянием; эта связь относится к заявленной диагностической или иной роли результата; приведенные источники применимы к заявленному материалу и популяции; имеющиеся ограничения не препятствуют использованию данных в пределах заявленного назначения. Если же часть назначения не подтверждена, вывод должен прямо указывать, что для этой части требуются дополнительные данные или ограничение области применения.

Важно, чтобы вывод о научной обоснованности не подменял выводы об аналитической и клинической эффективности. Научная обоснованность показывает, что анализ имеет медицинский смысл. Она не доказывает, что конкретное изделие надежно его определяет, и не всегда доказывает, что результат изделия клинически эффективен. Поэтому после раздела научной обоснованности должны логически сле-

довать разделы об аналитической эффективности и клинической эффективности. Их нельзя заменить литературным описанием клинического значения анализа.

В регистрационном досье раздел научной обоснованности должен быть связан с другими документами. Если в назначении изделия указаны определенные материалы, популяции и клинические роли, раздел научной обоснованности должен покрывать именно их. Если в инструкции по применению заявлены ограничения, они должны быть согласованы с выводом о научной обоснованности. Если программа клинико-лабораторных испытаний строится на определенном клиническом сценарии, он должен соответствовать той роли анализа, которая обоснована в данном разделе.

Для хорошо подготовленного досье характерна внутренняя согласованность. Назначение изделия, раздел научной обоснованности, программа испытаний, аналитические данные, клинические данные и итоговый отчет по клиническим доказательствам должны говорить об одном и том же применении. Если научная обоснованность подтверждает одно, аналитические исследования выполнены для другого, а итоговый вывод сформулирован для третьего, досье становится методологически слабым, даже если каждый документ по отдельности выглядит качественно.

Практически раздел научной обоснованности можно строить в следующей логике: краткое описание назначения изделия; описание анализа или объекта исследования; ха-

рактеристика клинической роли результата; обзор источников данных; критическая оценка и применимость источников; анализ ограничений; итоговый вывод. Такая структура не должна восприниматься как жесткий шаблон, но она помогает избежать главной ошибки — превращения научной обоснованности в несвязанный литературный обзор.

В результате документирование научной обоснованности должно отвечать на простой, но принципиальный вопрос: почему результат данного изделия имеет медицинский смысл именно в заявленном применении? Если раздел убедительно отвечает на этот вопрос, он создает прочное основание для дальнейшей оценки аналитической и клинической эффективности. Если не отвечает, последующие данные будут восприниматься как недостаточно связанные с медицинской задачей изделия.

Итог главы

Документирование научной обоснованности должно представлять связь анализа или иного объекта исследования с заявленной медицинской задачей как связную доказательную линию. Такой раздел должен начинаться от назначения изделия, раскрывать клиническую роль результата, приводить релевантные источники, оценивать их качество и применимость, фиксировать ограничения и завершаться точным выводом. Научная обоснованность не должна подменяться аналитическими характеристиками, клинической эффективностью или простым перечнем публикаций. Ее за-

дача — показать, почему результат медицинского изделия для диагностики *in vitro* имеет медицинский смысл в пределах заявленного назначения.

Итоги Части II

Часть II была посвящена научной обоснованности как первому содержательному элементу клинических доказательств медицинского изделия для диагностики *in vitro*. В ней показано, что доказательная оценка не может начинаться только с технических характеристик изделия. Прежде чем оценивать правильность, прецизионность, предел обнаружения, диагностическую чувствительность или клиническую эффективность, необходимо установить, почему сам аналит или иной объект исследования имеет медицинское значение.

Научная обоснованность строится вокруг связи аналита с клиническим состоянием, физиологическим или патологическим процессом либо медицинской задачей. Такая связь должна быть конкретизирована: диагностическая, скрининговая, мониторинговая, прогностическая, предиктивная или стратификационная. Недостаточно указать, что аналит “связан с заболеванием”. Нужно понимать, для какой роли результата, в какой популяции, в каком материале и по какому правилу интерпретации эта связь подтверждена.

В части рассмотрено различие между хорошо изученными и новыми аналитами. Для хорошо изученного аналита обычно не требуется заново доказывать базовую клиническую связь, но необходимо показать применимость извест-

ных данных к конкретному назначению изделия. Для нового или недостаточно изученного анализата требуется более развернутая доказательная стратегия, включая критическую оценку литературы, анализ воспроизводимости, применимости и, при необходимости, получение новых данных.

Источники научной обоснованности могут включать клинические рекомендации, систематические обзоры, первичные исследования, консенсусные документы, данные о биологическом механизме, регистры, данные реальной практики, материалы по аналогичным подходам и исследования алгоритмических моделей. Их доказательная сила определяется не названием источника, а тем, насколько они подтверждают заявленную клиническую роль результата и применимы к материалу, популяции и условиям использования изделия.

Отдельное значение имеют литературный поиск, систематические обзоры и критическая оценка данных. Поиск должен строиться и вокруг названия анализата, и вокруг клинической задачи, материала, популяции и правила интерпретации. Систематический обзор помогает оценить устойчивость связи анализата с медицинской задачей. Критическая оценка позволяет определить риск смещения, применимость источников, неоднородность, противоречия и ограничения вывода. При необходимости могут использоваться инструменты PRISMA, QUADAS-2, QUADAS-C, PROBAST и GRADE.

Главный практический вывод Части II состоит в том, что

научная обоснованность должна быть документирована как доказательная линия, а не как набор литературных ссылок. Она должна завершаться ясным выводом: какой аналит или объект исследования имеет медицинский смысл, для какой клинической роли, в каком материале, для какой популяции и с какими ограничениями. Только после этого можно обоснованно переходить к оценке аналитической эффективности конкретного изделия и его клинической эффективности.

Часть III . Аналитическая эффективность медицинских изделий для диагностики in vitro

Аннотация части

Часть III посвящена аналитической эффективности медицинского изделия для диагностики in vitro. Если научная обоснованность отвечает на вопрос, почему определяемый аналит или иной объект исследования имеет медицинский смысл, то аналитическая эффективность отвечает на следующий вопрос: способно ли конкретное изделие надежно выявлять, измерять, классифицировать или характеризовать этот объект в условиях, заявленных изготовителем.

Для изделий диагностики in vitro этот уровень доказательств имеет самостоятельное значение. Даже если аналит хорошо изучен и его клиническая роль не вызывает сомнений, нельзя автоматически считать, что любое изделие, предназначенное для его определения, будет давать надежный результат. Конкретная тест-система, набор реагентов, анализатор, калибратор, контрольный материал, программный модуль или алгоритм должны подтвердить свои аналитические характеристики применительно к заявленному материалу, диапазону, условиям применения, пользователю и правилам интерпретации результата.

В этой части рассматриваются общие принципы подтверждения аналитической эффективности, правильность, прецизионность, линейность, диапазон измерений, предел обнаружения, предел количественного определения, аналитическая специфичность, интерференция, стабильность, воспроизводимость, прослеживаемость и калибровка. Эти характеристики не должны восприниматься как формальный перечень испытаний. Их выбор, глубина оценки и критерии приемлемости зависят от назначения изделия, типа результата, клинической роли, материала и риска ошибочного решения.

Особое внимание уделяется тому, что аналитическая эффективность не равна клинической эффективности. Хорошая аналитическая работа изделия создает необходимую основу для доверия к результату, но сама по себе не всегда доказывает его пригодность для медицинского решения. В то же время клиническая эффективность не может быть убедительной, если аналитические характеристики изделия не подтверждены. Поэтому аналитическая эффективность занимает центральное место между научной обоснованностью аналита и клинической пригодностью результата.

Цель части

Цель настоящей части состоит в том, чтобы сформировать у читателя понимание аналитической эффективности как самостоятельного и обязательного элемента доказательной оценки медицинского изделия для диагностики *in vitro*.

После изучения этой части читатель должен понимать,

почему аналитические характеристики подтверждаются для конкретного изделия, а не для метода “вообще”; как назначение изделия влияет на выбор аналитических исследований; чем правильность отличается от прецизионности; за чем оцениваются линейность, диапазон измерений, предел обнаружения и предел количественного определения; как аналитическая специфичность и интерференция связаны с риском ложного результата; почему стабильность, воспроизводимость, прослеживаемость и калибровка имеют значение для надежности результата; и почему аналитическая эффективность должна быть согласована с типом материала, пользователем, условиями применения и клинической ролью результата.

Глава 17. Общие принципы подтверждения аналитической эффективности

Аналитическая эффективность является вторым крупным элементом доказательной оценки медицинского изделия для диагностики *in vitro*. Если научная обоснованность показывает, почему аналит или иной объект исследования имеет медицинский смысл, то аналитическая эффективность показывает, насколько надежно конкретное изделие получает результат по этому объекту. Это принципиально разные вопросы. Научная обоснованность относится к клиническому смыслу аналита, а аналитическая эффективность — к работе изделия как измерительной, детекционной, классификационной или вычислительной системы.

Для диагностики *in vitro* эта граница особенно важна. Хорошо известный аналит не делает изделие автоматически надежным. Глюкоза, HBsAg, вирусная нагрузка, EGFR, HER2, BCR::ABL1, фенилаланин или TREC могут иметь устойчивую клиническую роль, но конкретное изделие должно показать, что оно способно определять соответствующий объект в заявленном материале, диапазоне, условиях применения и формате результата. Поэтому аналитическая эффективность всегда подтверждается для конкретного изделия, его вер-

сии, состава реагентов, калибраторов, приборной платформы, программного компонента и правил интерпретации.

В этом смысле аналитическая эффективность является связующим звеном между научной обоснованностью и клинической эффективностью. Если аналитическая основа ненадежна, клинический результат не может быть устойчивым. Даже хорошо спланированное клинико-лабораторное испытание будет трудно интерпретировать, если неизвестно, как изделие ведет себя около порога, каков предел обнаружения, насколько воспроизводимы результаты, есть ли интерференция, как обеспечена прослеживаемость или как меняется результат при хранении образца. Клиническая эффективность предполагает, что аналитический результат уже имеет достаточную надежность.

Аналитическая эффективность — это способность медицинского изделия для диагностики *in vitro* надежно выявлять, измерять, классифицировать или характеризовать заявленный аналит либо иной объект исследования в установленных условиях применения. Она подтверждает не клинический смысл аналита, а пригодность конкретного изделия получать достоверный и воспроизводимый результат в пределах заявленного назначения.

Назначение изделия как исходная точка аналитической оценки

Подтверждение аналитической эффективности начинается с назначения изделия. Именно назначение определяет,

какие характеристики нужно оценивать и насколько глубокой должна быть эта оценка. Для количественного теста центральное значение имеют правильность, прецизионность, линейность, диапазон измерений, предел количественного определения и прослеживаемость. Для качественного теста важны предел обнаружения, аналитическая специфичность, устойчивость классификации, поведение слабopоложительных и пограничных образцов. Для молекулярного теста особое значение имеют аналитическая чувствительность, специфичность мишени, матрица, ингибиторы, контаминация и варианты последовательности. Для программного компонента — входные данные, версия алгоритма, правила обработки и устойчивость итоговой классификации.

Именно поэтому невозможно установить единый универсальный перечень аналитических исследований, одинаково достаточный для всех изделий. Одни характеристики обязательны для большинства изделий, другие зависят от типа результата, принципа метода, материала, пользователя и клинической роли. Линейность имеет принципиальное значение для количественного теста, но не является центральной характеристикой для простого качественного теста. Предел обнаружения критически важен для теста, который должен выявлять низкий уровень аналита, но может иметь иное значение для теста, работающего с высокими концентрациями. Интерференция имеет разный риск для биохимического анализа крови, молекулярного теста, иммунологического

метода или цифрового изображения.



Рисунок 17.1. Аналитическая эффективность как подтверждение надежности результата конкретного изделия для диагностики *in vitro*

Рисунок 17.1 показывает, что аналитическая оценка должна строиться не от абстрактного перечня показателей, а от реального применения изделия. Если изделие предназначено для мониторинга, особое значение имеет воспроизводимость результатов во времени и способность надежно отражать динамику. Если изделие используется для скрининга, важна устойчивость результата около порога и риск ложного отрицательного результата. Если результат влияет на выбор терапии, критично подтвердить, что изделие надежно выявляет именно те варианты, маркеры или категории, от которых зависит лечебное решение.

Материал, пользователь и условия применения

Аналитическая эффективность всегда связана с клиническим материалом. Один и тот же метод может работать по-разному в сыворотке, плазме, цельной крови, сухом пятне крови, тканевом материале, мазке, кале, моче или ликворе. Матрица может влиять на извлечение аналита, фон, ингибирование реакции, стабильность, интерференцию, порог обнаружения и воспроизводимость. Поэтому аналитические данные, полученные на одном типе материала, не должны автоматически распространяться на другой материал без обоснования. Если изделие заявляет несколько материалов, каждый из них должен быть подтвержден непосредственно или через убедительное обоснование сопоставимости.

Не меньшее значение имеют условия применения и пользователь. Изделие, применяемое в специализированной лаборатории обученным персоналом, и изделие для использования у постели пациента или самостоятельного применения пациентом работают в разных условиях риска. В первом случае можно ожидать соблюдения лабораторных процедур, контроля качества и квалифицированной интерпретации. Во втором возрастает значение устойчивости к ошибкам пользователя, простоты инструкции, стабильности результата в менее контролируемой среде и риска неправильного действия после получения результата. Поэтому аналитическая эффективность должна оцениваться в условиях, соответствующих предполагаемому применению.

Критерии приемлемости и клинически значимая ошибка

Критерии приемлемости аналитических характеристик не должны устанавливаться произвольно. Они должны быть связаны с назначением изделия и клинически значимой ошибкой. Для одних изделий небольшое аналитическое отклонение может не иметь существенного значения для медицинского решения. Для других даже небольшая ошибка около порога может изменить категорию результата и привести к другому медицинскому действию. Поэтому оценка аналитической эффективности должна учитывать статистические показатели, а также последствия аналитической ошибки.

Для количественного мониторингового теста важно понимать, какое изменение результата считается клинически значимым. Если аналитическая вариабельность сопоставима с ожидаемым клиническим изменением, результат трудно использовать для мониторинга. Для качественного теста важно оценить устойчивость классификации вблизи порога: слабopоложительный образец не должен случайно становиться отрицательным из-за малых колебаний метода. Для молекулярного теста низкий предел обнаружения имеет значение только тогда, когда он соответствует клинической роли результата и реальному уровню аналита в заявленном материале.

Схема 17.1 — Связь типа изделия с ключевыми ана-

литическими вопросами

— **Качественный тест** — *Ключевые аналитические вопросы:* Предел обнаружения, специфичность, устойчивость классификации, пограничные и слабоположительные образцы.

— **Количественный тест** — *Ключевые аналитические вопросы:* Правильность, прецизионность, линейность, диапазон измерений, предел количественного определения, прослеживаемость.

— **Полуколичественный тест** — *Ключевые аналитические вопросы:* Стабильность категорий, обоснованность порогов, воспроизводимость переходов между категориями.

— **Молекулярный тест** — *Ключевые аналитические вопросы:* Аналитическая чувствительность, специфичность мишени, ингибиторы, контаминация, варианты последовательности, матрица.

— **Иммунологический тест** — *Ключевые аналитические вопросы:* Перекрестная реактивность, интерференция, hook-эффект, матричные эффекты, стабильность реагентов.

— **Тест для мониторинга** — *Ключевые аналитические вопросы:* Воспроизводимость во времени, сопоставимость серий, аналитическая вариабельность относительно клинически значимого изменения.

— **Изделие с программным компонентом** — *Ключевые аналитические вопросы:* Версия алгоритма, вход-

ные данные, устойчивость обработки, воспроизводимость классификации, управление изменениями.

— **Самотестирование или применение вне лаборатории** — *Ключевые аналитические вопросы: Устойчивость к пользовательским ошибкам, условия хранения и применения, читаемость результата, пригодность инструкции.*

Валидация, верификация и рутинный контроль

При подтверждении аналитической эффективности важно различать валидацию, верификацию и рутинный контроль. Валидация показывает, что метод или изделие пригодны для заявленного применения и способны обеспечивать результат с требуемыми характеристиками. Верификация подтверждает, что заявленные характеристики воспроизводятся в конкретных условиях применения. Рутинный контроль поддерживает стабильность работы после внедрения. Эти действия связаны, но не заменяют друг друга.

Для регистрационной и доказательной оценки важно понимать, какие данные относятся к исходному подтверждению аналитической эффективности, а какие — к текущему контролю качества. Рутинный контроль не должен подменять первичную аналитическую оценку изделия. В то же время подтвержденные аналитические характеристики должны поддерживаться после выпуска изделия в обращение через контроль серий, калибровку, контрольные материалы, мониторинг стабильности и управление изменениями.

Серийность, стабильность и прослеживаемость

Аналитическая эффективность должна быть связана с серийностью и стабильностью изделия. Для наборов реагентов и контрольных материалов важно показать, что характеристики сохраняются между сериями, в течение срока годности и при заявленных условиях хранения и транспортирования. Для приборов необходимо учитывать техническое обслуживание, калибровку, программное обеспечение и условия эксплуатации. Для программных изделий важны версия алгоритма, контроль изменений и сопоставимость результатов после обновлений. Если эти элементы не контролируются, аналитическая надежность может быть доказана только для ограниченной конфигурации изделия.

Отдельное место занимает метрологическая прослеживаемость. Для количественных измерений результат должен быть связан с признанной референтной системой, если такая система существует. Это особенно важно для калибраторов, контрольных материалов и изделий, результаты которых используются для мониторинга или сравнения между лабораториями. Отсутствие полной прослеживаемости не всегда делает изделие непригодным, но требует ясного описания того, как обеспечивается сопоставимость и интерпретируемость результата.

Связь аналитической эффективности с клинической оценкой

Аналитическая эффективность не должна рассматривать-

ся в отрыве от последующей клинической оценки. Напротив, аналитические данные помогают определить, какие клинические результаты можно ожидать и какие риски нужно специально проверить. Если аналитическое исследование показывает слабую воспроизводимость около порога, клинико-лабораторное испытание должно учитывать пограничные случаи. Если выявлены потенциальные интерференты, необходимо понимать, встречаются ли они в целевой популяции. Если диапазон измерений ограничен, это влияет на заявленную область применения и интерпретацию результата.

При этом аналитическая эффективность не заменяет клиническую эффективность. Хорошая прецизионность, низкий предел обнаружения или отсутствие значимой интерференции показывают, что изделие надежно получает аналитический результат. Но они сами по себе не доказывают, что результат приводит к правильному медицинскому решению. Поэтому аналитические данные должны рассматриваться как необходимая основа клинической оценки, но не как ее замена.

Документирование аналитической эффективности

В доказательной документации данные об аналитической эффективности должны быть представлены как обоснование надежности результата. Недостаточно указать, что испытания проведены или что характеристики соответствуют требованиям. Нужно показать, какие характеристики оценены, почему они выбраны, на каком материале, в каких условиях,

с какими критериями приемлемости и как полученные результаты поддерживают назначение изделия. Такой подход позволяет связать аналитический раздел с научной обоснованностью и клинической эффективностью.

Типичная ошибка состоит в том, что аналитическую эффективность оформляют как независимый технический блок, не связанный с клинической ролью результата. В результате в документации могут быть представлены многочисленные аналитические показатели, но остается неясным, почему именно они достаточны для заявленного применения. Можно подробно показать прецизионность, но не оценить интерференцию, значимую для целевой популяции. Можно подтвердить предел обнаружения на модельных образцах, но не показать поведение изделия в заявленной матрице. Можно оценить линейность, но не обосновать диапазон, в котором результат имеет клинический смысл.

Следовательно, общие принципы подтверждения аналитической эффективности сводятся к нескольким взаимосвязанным положениям. Аналитические характеристики должны подтверждаться для конкретного изделия, в заявленном материале, с учетом типа результата, пользователя, условий применения и клинической роли. Критерии приемлемости должны быть связаны с риском аналитической ошибки и ее возможными последствиями. Данные должны быть достаточны не сами по себе, а в связи с назначением изделия и последующей клинической оценкой.

Итог главы

Аналитическая эффективность показывает, насколько надежно конкретное медицинское изделие для диагностики *in vitro* выявляет, измеряет, классифицирует или характеризует заявленный аналит либо иной объект исследования. Она не доказывает клинический смысл аналита и не заменяет клиническую эффективность, но является необходимым условием доверия к результату. Подтверждение аналитической эффективности должно строиться от назначения изделия, типа результата, клинического материала, условий применения, пользователя и риска ошибочного решения. Именно поэтому аналитическая оценка не является формальным перечнем испытаний, а представляет собой доказательство надежности результата в пределах заявленного применения.

Глава 18. Правильность и прецизионность

Правильность и прецизионность относятся к базовым характеристикам аналитической эффективности медицинского изделия для диагностики *in vitro*. Эти характеристики часто рассматриваются вместе, потому что обе связаны с надежностью результата. Однако они отвечают на разные вопросы и не могут заменять друг друга. Правильность показывает, насколько результат близок к принятому опорному значению. Прецизионность показывает, насколько близки повторные результаты между собой при повторном исследовании одного и того же объекта.

Для доказательной оценки медицинского изделия для диагностики *in vitro* это различие имеет принципиальное значение. Изделие может стабильно давать близкие друг к другу результаты, но все они будут смещены относительно опорного (истинного) значения. В таком случае метод будет прецизионным, но неправильным. Возможна и другая ситуация: средний результат серии измерений близок к опорному значению, но отдельные результаты сильно разбросаны. Тогда метод может быть приемлемым по среднему значению, но недостаточно прецизионным для практического применения.

И в том и в другом случае возникает риск для интерпретации результата. Систематически смещенный результат может устойчиво приводить к неправильной оценке концентрации аналита, неверной категории пациента или ошибочной динамике. Непрецизионный результат может менять интерпретацию при повторном исследовании: сегодня образец окажется выше порога, завтра ниже, хотя реальное состояние не изменилось. Поэтому правильность и прецизионность должны оцениваться совместно, особенно если результат используется около клинического порога, для мониторинга или для выбора медицинского действия.

Правильность — это близость среднего результата измерений к принятому опорному, референтному или аттестованному значению.

Прецизионность — это близость независимых результатов измерений между собой при повторном исследовании одного и того же объекта в установленных условиях.

Правильность преимущественно связана с систематической ошибкой, а прецизионность — со случайной вариабельностью результата.

Правильность и прецизионность удобно объяснять через образ мишени. Центр мишени соответствует опорному (истинному) значению. Если результаты сгруппированы близко к центру, метод одновременно правильный и прецизионный. Если результаты сгруппированы плотно, но далеко от центра, метод прецизионный, но имеет систематическое смещение.

ние. Если результаты разбросаны вокруг центра, среднее может быть близким к опорному значению, но отдельные измерения нестабильны. Если результаты и разбросаны, и смещены, метод имеет проблемы как с правильностью, так и с прецизионностью.



Примечание — Правильность (trueess) отражает отсутствие систематической ошибки (bias), прецизионность (precision) - степень воспроизводимости (повторяемости) результатов

Рисунок 18.1. Четыре типовые ситуации: правильность и прецизионность

Такая иллюстрация показывает, почему нельзя делать вывод о надежности изделия только по одному показателю. Хорошая прецизионность не гарантирует правильность, а приемлемое среднее значение не гарантирует устойчивость от-

дельных результатов.

Расчетные показатели правильности и прецизионности

Для количественных результатов правильность и прецизионность обычно оценивают с использованием нескольких базовых расчетов. Они не являются единственными возможными, но позволяют понять логику анализа данных.

Среднее значение серии измерений рассчитывают по формуле:

$$\bar{x} = (\sum(i=1..n) x_i) / (n),$$

где: $\bar{x}$ — среднее значение результатов измерений; x_i — отдельный результат измерения; n — число измерений.

Смещение результата относительно опорного значения рассчитывают как:

$$\text{Bias} = \bar{x} - x_{\text{ref}},$$

где: Bias — смещение; $\bar{x}$ — среднее значение результатов измерений; x_{ref} — опорное, референтное или аттестованное значение.

Смещение в процентах рассчитывают по формуле:

$$\text{Bias}(\%) = ((\bar{x} - x_{\text{ref}}) / x_{\text{ref}}) \times 100.$$

Прецизионность часто выражают через стандартное отклонение:

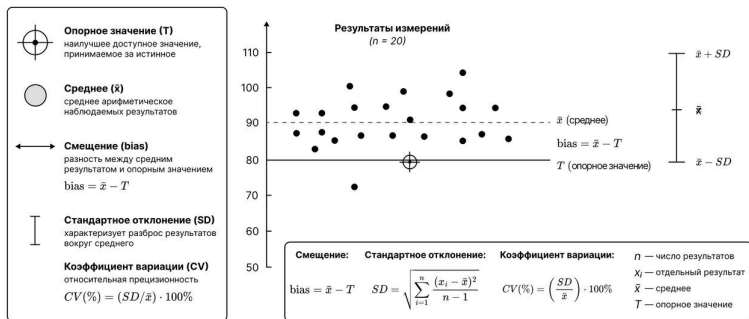
$$\text{SD} = \sqrt{(\sum(i=1..n) (x_i - \bar{x})^2) / (n - 1)},$$

где: SD — стандартное отклонение; x_i — отдельный результат измерения; $\bar{x}$ — среднее значение; n — число измерений.

Для сравнения разброса при разных уровнях аналита используют коэффициент вариации:

$$CV(\%) = (SD / \bar{x}) \times 100$$

Коэффициент вариации удобен тем, что выражает разброс относительно среднего значения. Однако его нужно интерпретировать осторожно, особенно при низких концентрациях аналита, где малое абсолютное изменение может давать высокий процентный разброс. Поэтому для нижних уровней диапазона иногда важнее оценивать не только CV, но и абсолютное стандартное отклонение, устойчивость классификации и клинический смысл отклонения.



Опорное значение (T)	Значение, наиболее близкое к истинному (например, сертифицированное референсное значение или консенсусное значение)
Среднее ($\bar{x}$)	Характеризует центральную тенденцию результатов
Смещение (bias)	Показывает, насколько результаты систематически выше (+) или ниже (-) опорного значения
SD и CV	Показывают разброс результатов: SD - в единицах измерения; CV - в процентах, для сравнения в разных диапазонах

Примечание — Хорошая правильность соответствует малому смещению ($bias \approx 0$); хорошая прецизионность - малому SD (и CV)

Рисунок 18.2. Опорное значение, среднее, смещение и

прецизионность (SD/CV)

Пример 18.1 — Оценка смещения и прецизионности на одном уровне анализа

Предположим, что контрольный материал имеет аттестованное значение 100 Ед/мл. При повторных измерениях получены результаты: 96; 98; 97; 99; 100 Ед/мл.

Среднее значение:

$$\bar{x} = (96 + 98 + 97 + 99 + 100)/(5) = 98,0 \text{ Ед/мл}$$

Смещение:

$$\text{Bias} = 98,0 - 100,0 = - 2,0 \text{ Ед/мл}$$

Смещение в процентах:

$$\text{Bias} (\%) = (- 2,0)/(100,0) \times 100 \% = - 2,0\%$$

В этом примере средний результат на 2 % ниже опорного значения. Если заранее установленный критерий приемлемости допускает смещение не более $\pm 5\%$, результат по правильности можно считать приемлемым. Если же для данного изделия и клинической задачи допустимое смещение составляет только $\pm 1\%$, результат будет неприемлемым, несмотря на то что измерения выглядят достаточно близкими друг к другу.

Для оценки прецизионности рассчитывают стандартное отклонение и коэффициент вариации. В данном примере результаты расположены близко к среднему значению, поэтому можно ожидать невысокий CV. Но окончательный вывод должен делаться не визуально, а на основании расче-

та и заранее установленного критерия приемлемости.

Этот пример показывает важный принцип: числовой результат сам по себе еще не говорит, приемлема характеристика или нет. Необходим заранее установленный критерий. Он должен быть связан с назначением изделия, клинически значимой ошибкой, требованиями к методу и риском неправильной интерпретации.

Подтверждение правильности

Подтверждение правильности требует сравнения результата изделия с некоторым опорным основанием. Таким основанием может быть стандартный образец, референтный материал, контрольный материал с установленным значением, референтная методика, признанный метод сравнения, межлабораторное консенсусное значение или панель охарактеризованных клинических образцов. Выбор основания зависит от аналита, доступности референтной системы, типа материала и назначения изделия.

Наиболее сильным вариантом является сравнение с материалом или методом, имеющим метрологически обоснованное значение. Если существует референтный материал или референтная методика, их использование повышает доказательную ценность оценки правильности. Если такой системы нет, допустимо использовать обоснованный сравнительный подход, но его статус должен быть объяснен. Нельзя считать правильность подтвержденной только потому, что результаты совпали с любым доступным коммерческим набором, ес-

ли не показано, почему этот набор может выполнять роль сравнительного основания.

Для медицинских изделий для диагностики *in vitro* особенно важно, чтобы правильность оценивалась на релевантном материале и в значимом диапазоне. Материалы на основе буфера, искусственно приготовленные образцы или контрольные материалы могут быть полезны, но они не всегда отражают поведение реального клинического образца. Матрица может влиять на извлечение аналита, связывание, фон, интерференцию, стабильность и сигнал. Поэтому, если изделие заявлено для сыворотки, плазмы, цельной крови, сухого пятна крови, мазка, ткани или другого материала, оценка правильности должна учитывать именно этот материал либо содержать обоснование переноса данных.

Правильность также должна оцениваться на нескольких уровнях аналита. Проверка только одного удобного уровня не показывает, как изделие работает в нижней части диапазона, в области клинического порога или на высоких значениях. Для количественного теста обычно важно включить низкий, средний и высокий уровни, а также уровни около клинически значимых решений. Для теста, используемого для мониторинга, необходимо учитывать диапазон ожидаемых изменений. Для скринингового теста особое значение имеет область порог решения (cut-off).

При сравнении с методом сравнения важно не только рассчитать среднее смещение. Нужно оценить, нет ли пропор-

ционального смещения, которое меняется в зависимости от уровня аналита. Например, на низких концентрациях метод может давать приемлемые результаты, а на высоких — систематически занижать значения. Или наоборот, смещение может быть малым в средней части диапазона, но критичным около нижней границы. Поэтому для оценки правильности часто используют не только расчет bias, но и графики сравнения методов, анализ различий по уровням, регрессионные подходы и оценку клинически значимых зон.

Пример 18.2 — Оценка правильности по контрольным материалам с установленными значениями

Для оценки правильности количественного изделия использованы три контрольных материала с установленными значениями аналита: низкий, средний и высокий уровень. Каждый материал измеряли 5 раз. Критерий приемлемости установлен заранее: смещение результата не должно превышать $\pm 5\%$ от установленного значения.

Схема 18.1 — Результаты оценки правильности по трем уровням аналита

— **Низкий** — Установленное значение, Ед/мл: 20,0;
Результаты измерений, Ед/мл: 18,8; 19,1; 18,9; 19,0; 19,2;
Среднее значение, Ед/мл: 19,0; Смещение, Ед/мл: -1,0;
Смещение, %: -5,0 %; Оценка: Приемлемо на границе.

— **Средний** — Установленное значение, Ед/мл: 100,0;
Результаты измерений, Ед/мл: 97,5; 98,0; 98,2; 97,8; 98,5;
Среднее значение, Ед/мл: 98,0; Смещение, Ед/мл: -2,0;

Смещение, %: -2,0 %; **Оценка:** Приемлемо.

— **Высокий** — **Установленное значение, Ед/мл:** 500,0;
Результаты измерений, Ед/мл: 465; 470; 468; 472; 469;
Среднее значение, Ед/мл: 468,8; **Смещение, Ед/мл:** -31,2;
Смещение, %: -6,2 %; **Оценка:** Неприемлемо.

Расчет для среднего уровня:

$$\bar{x} = (97,5 + 98,0 + 98,2 + 97,8 + 98,5)/(5) = 98,0 \text{ Ед/мл}$$

$$\text{Bias} = 98,0 - 100,0 = - 2,0 \text{ Ед/мл}$$

$$\text{Bias (\%)} = (- 2,0)/(100,0) \times 100\% = - 2,0\%$$

Так как смещение составляет -2,0 %, а заранее установленный критерий приемлемости равен ± 5 %, правильность на среднем уровне считается подтвержденной.

Для высокого уровня:

$$\bar{x} = (465 + 470 + 468 + 472 + 469)/(5) = 468,8 \text{ Ед/мл}$$

$$\text{Bias} = 468,8 - 500,0 = - 31,2 \text{ Ед/мл}$$

$$\text{Bias (\%)} = (- 31,2)/(500,0) \times 100 = - 6,2\%$$

Смещение составляет -6,2 %, что превышает установленный критерий ± 5 %. Следовательно, правильность на высоком уровне не подтверждена.

В этом примере изделие демонстрирует приемлемую правильность на низком и среднем уровнях, но имеет неприемлемое отрицательное смещение на высоком уровне. Такой результат нельзя описывать общей фразой «правильность подтверждена». Более корректный вывод должен быть ограниченным:

Правильность изделия подтверждена на низком и сред-

нем уровнях аналита в пределах установленного критерия приемлемости. На высоком уровне выявлено отрицательное смещение -6,2 %, превышающее критерий ± 5 %. Требуется анализ причин смещения, дополнительная проверка верхней части диапазона, возможное ограничение диапазона измерений или корректировка калибровки при наличии обоснования.

Этот пример показывает, что оценка правильности должна проводиться не только в одной точке. Если проверить только средний уровень, можно ошибочно сделать вывод о приемлемости изделия во всем диапазоне. Поэтому для количественных изделий важно оценивать несколько уровней, включая клинически значимые области, нижнюю и верхнюю часть диапазона, а также зоны принятия решений.

Подтверждение прецизионности

Подтверждение прецизионности требует повторных измерений одного и того же объекта в заданных условиях. Но условия могут быть разными, и от этого зависит смысл полученного результата. Если все измерения выполнены одним оператором, на одном приборе, одной серией реагентов и в короткий промежуток времени, оценивается *повторяемость*. Если измерения выполняются в разные дни, разными операторами, с разными сериями реагентов или калибровками внутри одной лаборатории, оценивается *внутрилабораторная прецизионность*. Если исследование охватывает разные лаборатории или площадки, речь идет о *воспроизво-*

димости.

Для подтверждения прецизионности важно заранее определить, какие источники вариабельности должны быть включены. Это зависит от назначения изделия. Если изделие будет использоваться в одной лаборатории на одной платформе, достаточно одного дизайна. Если изделие предназначено для применения в разных лабораториях, с разными операторами и партиями реагентов, необходимо учитывать более широкий набор факторов. Если изделие предназначено для самостоятельного применения, необходимо учитывать пользовательскую вариабельность и возможные ошибки выполнения процедуры.

Прецизионность также не следует подтверждать только на одном уровне аналита. Для количественных тестов обычно выбирают несколько уровней: низкий, близкий к клиническому порогу, средний и высокий. Для качественных тестов особенно важны отрицательные, слабopоложительные и пограничные образцы. Для мониторинговых тестов важно оценить уровни, на которых врач будет интерпретировать динамику результата. Для полуколичественных тестов нужно проверить устойчивость категорий и переходов между ними.

Недостаточная прецизионность имеет разный смысл в разных клинических сценариях. Если тест используется для грубой ориентировочной оценки, умеренная вариабельность может быть допустимой. Если результат используется для выбора терапии, контроля токсичности, мониторинга вирус-

ной нагрузки, оценки минимальной остаточной болезни или стратификации риска, даже небольшой разброс может быть значимым. Поэтому критерии приемлемости прецизионности должны быть связаны не только с техническими возможностями метода, но и с клиническим применением результата.

Схема 18.2 — Различие правильности, прецизионности и связанных характеристик

— **Правильность** — *Основной вопрос:* Насколько результат близок к опорному значению?; *Что оценивается:* Систематическое смещение результата; *Типичный риск при недостаточности:* Устойчиво неверная оценка уровня аналита.

— **Прецизионность** — *Основной вопрос:* Насколько близки повторные результаты между собой?; *Что оценивается:* Случайная вариабельность результата; *Типичный риск при недостаточности:* Нестабильная интерпретация результата.

— **Повторяемость** — *Основной вопрос:* Насколько близки результаты в одинаковых условиях?; *Что оценивается:* Минимальная вариабельность в одной серии или близких условиях; *Типичный риск при недостаточности:* Нестабильность даже при кратковременном повторе.

— **Внутрилабораторная прецизионность** — *Основной вопрос:* Насколько близки результаты при обычных из-

менениях внутри лаборатории?; **Что оценивается:** Влияние дней, операторов, серий, калибровок, приборов; **Типичный риск при недостаточности:** Нестабильность в реальной лабораторной практике.

— **Воспроизводимость** — **Основной вопрос:** Насколько близки результаты в разных лабораториях или условиях?; **Что оценивается:** Межлабораторные и межплощадочные различия; **Типичный риск при недостаточности:** Ограниченная переносимость результата между условиями применения.

Пример 18.3 — Оценка прецизионности по коэффициенту вариации

Для оценки прецизионности количественного изделия использован контрольный материал со средним уровнем анализа. Выполнено 10 повторных измерений в одной серии. Заранее установленный критерий приемлемости: коэффициент вариации не должен превышать 5 %.

Получены результаты, Ед/мл: 98; 101; 99; 100; 102; 97; 100; 99; 101; 100.

Среднее значение:

$$\bar{x} = (98 + 101 + 99 + 100 + 102 + 97 + 100 + 99 + 101 + 100) / (10) = 99,7 \text{ Ед/мл}$$

Предположим, что рассчитанное стандартное отклонение составляет:

$$SD = 1,49 \text{ Ед/мл}$$

Коэффициент вариации:

$$CV (\%) = (1,49)/(99,7) \times 100 = 1,49\%$$

Так как коэффициент вариации составляет 1,49 %, а критерий приемлемости установлен как не более 5 %, повторяемость на данном уровне считается подтвержденной.

Однако такой вывод ограничен только условиями проведенного исследования: один уровень аналита, одна серия, близкие условия выполнения. На основании этого примера нельзя автоматически утверждать, что прецизионность подтверждена для всех уровней аналита, всех дней, всех партий реагентов, всех операторов и всех условий применения. Для более широкого вывода требуется дизайн, включающий соответствующие источники вариабельности.

Значение для количественных, качественных и полуколичественных изделий

Для количественных изделий правильность и прецизионность обычно выражаются числовыми показателями. Здесь важно оценить как среднее значение и разброс, так и то, как они меняются по диапазону измерений. Если результат используется для диагностики по порогу, мониторинга динамики или коррекции терапии, аналитическая ошибка должна рассматриваться через возможное клиническое последствие.

Для качественных тестов правильность не всегда выражается как смещение относительно числового значения. Здесь

большее значение имеет согласие с обоснованным сравнительным подходом и устойчивость классификации. Если изделие выдает результат «обнаружено» или «не обнаружено», нужно показать, насколько стабильно оно классифицирует отрицательные, положительные, слабоположительные и пограничные образцы. Прецизионность качественного теста фактически проявляется как повторяемость категории результата.

Для полуколичественных тестов и шкал важно оценивать устойчивость категорий. Если результат выражается как «низкий», «средний», «высокий», как балльная шкала, индекс или категория риска, нужно понимать, как часто один и тот же образец при повторном исследовании попадает в ту же категорию. Если переход между категориями меняет медицинское действие, такая вариабельность становится клинически значимой.

Для мониторинговых изделий прецизионность имеет особое значение. Врач интерпретирует и единичное значение, и изменение во времени. Если аналитическая вариабельность сопоставима с ожидаемым клинически значимым изменением, результат трудно использовать для мониторинга. Поэтому в таких изделиях нужно оценивать, какое изменение можно считать реальным, а какое может объясняться разбросом метода.

Для изделий с программным или алгоритмическим компонентом правильность и прецизионность требуют адап-

тированной трактовки. Если программа рассчитывает числовой показатель, оценивается корректность вычислений и воспроизводимость результата при одинаковых входных данных. Если алгоритм классифицирует изображение, генетический вариант или иной объект, оценивается устойчивость классификации при допустимых вариациях входных данных, качества изображения, параметров обработки и версии программного обеспечения. Если обновление алгоритма меняет классификацию, прежние данные о прецизионности и правильности могут требовать пересмотра.

Критерии приемлемости

Критерии приемлемости для правильности и прецизионности должны быть установлены заранее. Их нельзя подбирать после получения результатов. Критерии должны быть обоснованы через назначение изделия, клинически значимую ошибку, состояние технологии, требования применимых стандартов и риск неправильного решения.

Для одних изделий допустимое смещение может быть относительно широким, если оно не меняет клинический вывод. Для других даже небольшое смещение около порога может быть неприемлемым. Аналогично, CV может быть технически низким, но клинически недостаточным, если результат используется для оценки малых изменений во времени. Поэтому критерий приемлемости должен отвечать как на вопрос «хорошо ли выглядит число», так и на вопрос «не приведет ли такая ошибка к неправильной интерпретации».

результата».

Особенно важно заранее определить критерии для уровней около клинического решения. Если порог отделяет отрицательный результат от положительного, стабильность именно этой зоны имеет большее значение, чем поведение явно положительных или явно отрицательных образцов. Если изделие используется для мониторинга, критерий должен учитывать минимальное изменение, которое имеет клиническое значение. Если изделие используется для выбора терапии, критерий должен учитывать риск неверного включения или исключения пациента из соответствующего лечебного маршрута.

Что правильно и что неправильно при подтверждении правильности и прецизионности

Схема 18.3 — Типовые ошибки при подтверждении правильности и прецизионности

— **Оценка правильности** — *Правильно: Использовать материал с установленным опорным значением или обоснованный метод сравнения; Неправильно: Сравнить результат с любым доступным набором без обоснования его статуса.*

— **Выбор уровней аналита** — *Правильно: Включить несколько уровней, включая клинически значимые зоны; Неправильно: Проверить только один удобный уровень.*

— **Пороговые тесты** — *Правильно: Оценить образцы около порога принятия решения; Неправильно: Использо-*

вать только явно отрицательные и явно положительные образцы.

— **Оценка прецизионности** — *Правильно:* Варьировать условия в соответствии с реальным применением изделия; *Неправильно:* Ограничиться повторными измерениями в одной серии, если изделие будет применяться шире.

— **Мониторинг** — *Правильно:* Оценить вариабельность относительно клинически значимого изменения; *Неправильно:* Рассчитать CV без объяснения, влияет ли он на интерпретацию динамики.

— **Несколько материалов** — *Правильно:* Проверить каждый заявленный материал или обосновать перенос данных; *Неправильно:* Подтвердить один материал и распространить вывод на все остальные.

— **Критерии приемлемости** — *Правильно:* Установить заранее и обосновать; *Неправильно:* Подобрать критерии после получения результатов.

— **Алгоритмический результат** — *Правильно:* Оценить устойчивость итоговой классификации или индекса; *Неправильно:* Проверить только входные данные, не оценивая итоговый вывод.

Эта схема полезна для практической оценки программы испытаний и отчета. Она показывает, что аналитическая характеристика подтверждается не самим фактом проведения измерений, а корректным дизайном, правильным выбором материала, уровней, условий и критериев приемлемости.

Документирование в отчете

В отчете по подтверждению правильности и прецизионности необходимо описать материалы, уровни аналита, количество измерений, условия исследования, источники вариабельности, метод сравнения, опорные значения, расчетные показатели и критерии приемлемости. Если использовались стандартные образцы или референтные материалы, нужно указать их характеристики и статус. Если использовался метод сравнения, необходимо обосновать его применимость.

Для прецизионности важно указать, какие условия варьировались: дни, операторы, приборы, серии реагентов, калибровки, площадки или версии программного обеспечения. Для каждого уровня аналита следует представить результаты расчета: среднее значение, SD, CV или иные применимые показатели. Для качественных и полуколичественных результатов нужно представить устойчивость классификации и частоту смены категории.

Если часть результатов не соответствует критериям приемлемости, это не должно скрываться. Необходимо объяснить, как это влияет на назначение изделия. Иногда требуется ограничить диапазон применения, изменить правило интерпретации, уточнить инструкцию, провести дополнительное исследование или пересмотреть критерии приемлемости, если они были изначально необоснованными. Но пересмотр критерия после получения неудобных данных без ме-

тодического обоснования является недопустимым.

В сильной доказательной базе правильность и прецизионность показывают, что изделие надежно получает результат в тех условиях, в которых оно будет применяться. Они не доказывают клинический смысл аналита и не заменяют клиническую эффективность, но без них невозможно доверять результату как основе медицинского решения. Именно поэтому эти характеристики занимают одно из центральных мест в подтверждении аналитической эффективности.

Итог главы

Правильность и прецизионность являются базовыми характеристиками аналитической эффективности медицинского изделия для диагностики *in vitro*. Правильность отражает близость результата к опорному значению и связана с систематической ошибкой. Прецизионность отражает близость повторных результатов между собой и связана со случайной вариабельностью. Эти характеристики подтверждаются через специально спланированные исследования с использованием релевантных материалов, уровней аналита, повторных измерений, варьируемых условий и заранее установленных критериев приемлемости. Их значение определяется и техническими расчетами, и влиянием на клиническую интерпретацию результата, особенно около порога, при мониторинге и при классификации пациента.

Глава 19. Линейность и диапазон измерений

Линейность и диапазон измерений относятся к тем характеристикам аналитической эффективности, которые особенно важны для количественных и некоторых полуколичественных медицинских изделий для диагностики *in vitro*. Они показывают, в каких пределах результат изделия может быть интерпретирован как надежное количественное отражение содержания аналита в образце. Если правильность и прецизионность отвечают на вопросы о близости результата к опорному значению и воспроизводимости измерений, то линейность и диапазон измерений помогают понять, в каком интервале эти результаты сохраняют аналитический и клинический смысл.

Линейность обычно понимается как способность аналитической системы давать результаты, пропорциональные количеству или концентрации аналита в образце в пределах определенного интервала. Если концентрация аналита увеличивается в два раза, результат метода в линейном участке должен изменяться предсказуемо и соразмерно. Это особенно важно для тех изделий, где числовой результат используется и для однократной классификации, и для оценки степени выраженности состояния, мониторинга динамики, кон-

троля терапии или сравнения результатов во времени.

Диапазон измерений связан с линейностью, но не сводится к ней. Диапазон показывает интервал значений, в пределах которого изделие может выдавать результат с приемлемыми характеристиками правильности, прецизионности и, при необходимости, линейности. Иными словами, диапазон измерений — это не просто интервал, в котором прибор «что-то показывает». Это область, где результат можно использовать в соответствии с назначением изделия. Если результат выходит за пределы подтвержденного диапазона, его нельзя интерпретировать так же, как результат внутри диапазона.

***Линейность** — это способность аналитической системы давать результаты, пропорциональные содержанию аналита в образце в пределах установленного интервала.*

***Диапазон измерений** — это интервал значений аналита, в пределах которого изделие обеспечивает результат с характеристиками, достаточными для заявленного применения. Диапазон измерений должен рассматриваться вместе с назначением изделия, клинически значимыми уровнями аналита и правилами интерпретации результата.*

Для медицинского изделия для диагностики *in vitro* принципиально важно различать технически наблюдаемый диапазон сигнала и диапазон, в котором результат пригоден для

медицинского применения. Аналитическая система может регистрировать сигнал за пределами подтвержденной области, но это не означает, что такой результат можно выдавать как количественный. В нижней части диапазона возрастает неопределенность, ухудшается прецизионность и появляется риск нестабильного различения низких концентраций. В верхней части диапазона возможны насыщение сигнала, отклонение от линейности, необходимость разведения, hook-эффект или иные явления, искажающие результат.

Назначение изделия определяет, какой диапазон является достаточным. Для теста, используемого для мониторинга вирусной нагрузки, диапазон должен покрывать уровни, имеющие значение для оценки динамики и ответа на терапию. Для гормонального теста важно, чтобы диапазон включал зоны, связанные с клиническими порогами и ожидаемыми физиологическими или патологическими концентрациями. Для неонатального скрининга диапазон должен быть пригоден для надежного различения значений около порога первичного отбора. Для терапевтического лекарственного мониторинга особенно важны значения в терапевтическом, субтерапевтическом и токсическом диапазонах.

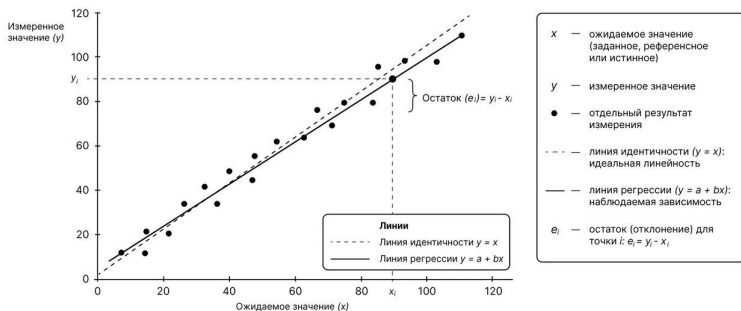
Таким образом, диапазон измерений не должен устанавливаться только из удобства метода. Он должен соответствовать клинической задаче. Если изделие заявлено для мониторинга, недостаточно показать, что оно измеряет аналит в узком интервале около одного значения. Если результат ис-

пользуется для выявления очень низких концентраций, нижняя часть диапазона становится критической. Если в клинической практике возможны очень высокие значения, необходимо понимать, как изделие ведет себя в верхней части диапазона и что следует делать при выходе результата за пределы измерения.

Диапазон измерений должен быть связан с аналитической системой, а также с клинической интерпретацией. Если диапазон не покрывает клинически важные уровни, изделие может быть аналитически работоспособным, но ограниченным для заявленного применения. Если правила интерпретации результатов вне диапазона не описаны, пользователь может ошибочно воспринимать непригодное значение как полноценный количественный результат.

Графическое представление линейности

Оценку линейности удобно представлять графически. По оси X откладывают ожидаемое или заданное значение аналита, а по оси Y — измеренное значение. Идеальная линейность соответствует прямой линии, при которой измеренные значения пропорциональны ожидаемым. Фактические результаты измерений сравнивают с этой линией и с заранее установленными допустимыми отклонениями.



Что показывает график

Как измеренные значения (y) соотносятся с ожидаемыми (x) по всему диапазону.

Линейность подтверждается, если точки близки к линии идентичности, а регрессионная прямая близка к $y = x$ (наклон $b \approx 1$, пересечение $a \approx 0$), без систематических отклонений по диапазону

Критерии линейности (примеры)

- Наклон b в пределах 0,95–1,05 (или согласно установленным критериям)
- Пересечение a близко к 0 (отсутствие систематического смещения)
- Остатки (e_i) случайно распределены вокруг 0, без тренда
- $|e_i|$ не превышают заранее заданные границы во всем диапазоне

Примечание — Ожидаемые значения могут быть заданы, например, референсным методом, аттестованными материалами или приготовленными смесями с известной концентрацией

Рисунок 19.1. Линейность: ожидаемое значение против измеренного.

Если точки расположены близко к ожидаемой линейной зависимости в пределах допустимого отклонения, линейность может считаться подтвержденной в данном интервале. Если на нижних или верхних уровнях наблюдается систематическое отклонение, диапазон должен быть ограничен или должны быть установлены специальные правила интерпретации. Принципиально важно оценивать и общий вид графика, и уровни, близкие к клинически значимым порогам.

Математически линейную зависимость часто описывают уравнением:

$$y = a + bx,$$

где: — ожидаемое, заданное или опорное значение аналита; — измеренное значение; — свободный член; — наклон линии.

Если система идеально соответствует ожидаемым значениям, свободный член должен быть близок к нулю, а наклон — к единице. Но в реальной аналитической системе возможны отклонения. Поэтому при оценке линейности важно анализировать не только коэффициенты уравнения, но и практическое значение отклонений на отдельных уровнях.

Иногда для оценки линейности используют коэффициент детерминации . Высокое значение может показывать хорошее соответствие линейной модели, но само по себе не является достаточным доказательством пригодности диапазона. Даже при высоком отдельные уровни могут иметь неприемлемое отклонение, особенно около нижней границы диапазона, верхней границы или клинического порога. Поэтому вывод о линейности нельзя сводить только к одному числу. Нужно оценивать отклонения по уровням и их клинический смысл.

Как подтверждают линейность

Линейность обычно проверяют на серии материалов с разными уровнями аналита, охватывающими предполагаемый диапазон измерений. Такие материалы могут быть получены путем смешивания образцов с высоким и низким содержанием аналита, последовательных разведений, использования

стандартных материалов, охарактеризованных панелей или клинических образцов. Выбранный подход должен соответствовать аналиту, матрице и назначению изделия.

При подготовке серии разведений важно учитывать, что разведение может менять свойства матрицы. Если образец разводят буфером или искусственной средой, полученные материалы могут отличаться от реального клинического образца. В некоторых методах это влияет на восстановление аналита, фон, связывание, интерференцию или стабильность. Поэтому линейность, подтвержденная на модельном материале, не всегда автоматически переносится на заявленную клиническую матрицу.

Обычно для оценки линейности выбирают несколько уровней, равномерно или целенаправленно покрывающих заявленный диапазон. Особое внимание уделяют нижней границе, верхней границе и уровням около медицински значимых решений. Для каждого уровня выполняют повторные измерения, рассчитывают среднее значение, отклонение от ожидаемого значения и оценивают соответствие заранее установленному критерию приемлемости.

Критерий приемлемости может выражаться в абсолютном отклонении, процентном отклонении или иной форме, обоснованной для данного аналита и назначения изделия. Например, для одних показателей допустимое отклонение может быть выражено как $\pm 10\%$, для других — как конкретная абсолютная величина. В нижней части диапазона процент-

ное отклонение может быть менее информативным, поэтому допустимо использовать абсолютные критерии или специальные подходы к оценке нижней границы количественного определения.

Пример 19.1 — Проверка линейности по серии разведений

Для оценки линейности количественного молекулярного теста использован образец с высоким содержанием аналита. Исходное значение принято как 1000 копий/мл. Подготовлены последовательные разведения: 1:2, 1:4, 1:8 и 1:16. Для каждого уровня рассчитано ожидаемое значение и получено среднее измеренное значение. Критерий приемлемости установлен заранее: отклонение измеренного значения от ожидаемого не должно превышать $\pm 10\%$.

Схема 19.1 — Пример оценки линейности по серии разведений

— **Без разведения** — Ожидаемое значение, копий/мл: 1000,0; Измеренное значение, копий/мл: 980,0; Отклонение, копий/мл: -20,0; Отклонение, %: -2,0 %; Оценка: Приемлемо.

— **1:2** — Ожидаемое значение, копий/мл: 500,0; Измеренное значение, копий/мл: 510,0; Отклонение, копий/мл: +10,0; Отклонение, %: +2,0 %; Оценка: Приемлемо.

— **1:4** — Ожидаемое значение, копий/мл: 250,0; Измеренное значение, копий/мл: 260,0; Отклонение, копий/мл: +10,0; Отклонение, %: +4,0 %; Оценка: Приемлемо.

— **1:8** — Ожидаемое значение, копий/мл: 125,0; Измеренное значение, копий/мл: 118,0; Отклонение, копий/мл: -7,0; Отклонение, %: -5,6 %; Оценка: Приемлемо.

— **1:16** — Ожидаемое значение, копий/мл: 62,5; Измеренное значение, копий/мл: 45,0; Отклонение, копий/мл: -17,5; Отклонение, %: -28,0 %; Оценка: Неприемлемо.

Отклонение от ожидаемого значения рассчитывают как:

$\Delta = x_{\text{meas}} - x_{\text{exp}}$, где: — измеренное значение; — ожидаемое значение.

Отклонение в процентах рассчитывают по формуле:

$$\Delta(\%) = ((x_{\text{meas}} - x_{\text{exp}}) / x_{\text{exp}}) \times 100$$

Расчет отклонения для уровня 1:8:

$$\Delta = 118,0 - 125,0 = -7,0 \text{ копий/мл.}$$

$$\Delta(\%) = (-7,0)/(125,0) \times 100\% = -5,6\%.$$

Так как отклонение составляет -5,6 %, а критерий приемлемости установлен как ± 10 %, уровень 1:8 считается приемлемым.

Для уровня 1:16:

$$\Delta = 45,0 - 62,5 = -17,5 \text{ копий/мл.}$$

$$\Delta(\%) = (-17,5)/(62,5) \times 100 = -28,0\%.$$

Отклонение составляет -28,0 %, что превышает критерий приемлемости ± 10 %. Следовательно, линейность на уровне 62,5 копий/мл не подтверждена.

Корректный вывод по этому примеру должен быть ограниченным:

Линейность теста подтверждена в интервале от 125 до 1000 копий/мл. На уровне 62,5 копий/мл наблюдается неприемлемое отрицательное отклонение, поэтому данный уровень не должен включаться в подтвержденный линейный диапазон без дополнительных данных. Нижняя граница линейного диапазона требует ограничения или дополнительного подтверждения.

Этот пример показывает, почему нельзя делать вывод о линейности только по общему тренду. Большинство уровней могут соответствовать критерию, но один уровень в нижней части диапазона может оказаться неприемлемым. Если именно этот уровень клинически важен, требуется пересмотр нижней границы диапазона, дополнительное исследование или уточнение инструкции.

Нижняя граница диапазона

В нижней части диапазона особенно важно различать обнаружение и количественное измерение. Изделие может обнаруживать наличие аналита на низком уровне, но не обеспечивать количественный результат с приемлемой правильностью и прецизионностью. Поэтому нижняя граница диапазона измерений должна быть согласована с пределом обнаружения и пределом количественного определения.

Предел обнаружения показывает минимальный уровень аналита, который может быть надежно отличен от отсутствия или фоновых сигнала. Предел количественного определения показывает уровень, начиная с которого аналит может

быть как обнаружен, так и измерен с приемлемыми характеристиками. Если изделие выдает числовые результаты ниже уровня количественной надежности, пользователь может ошибочно интерпретировать их как полноценные количественные значения. Это особенно опасно для мониторинга, где малые изменения результата могут восприниматься как динамика состояния.

Поэтому в нижней части диапазона необходимо оценивать как линейность, так и прецизионность, правильность и правила выдачи результата. Иногда корректнее указывать результат как «обнаружено ниже диапазона количественного определения» или «менее чем установленная граница», чем выдавать точное число, которое аналитически недостаточно надежно.

Верхняя граница диапазона и насыщение сигнала

В верхней части диапазона возможны другие проблемы. При высоких концентрациях аналита сигнал может перестать увеличиваться пропорционально концентрации. Это может происходить из-за насыщения детектора, ограничения реакционной системы, исчерпания реагента или иных свойств метода. В такой ситуации результат начинает занижаться относительно истинного уровня, а линейность теряется.

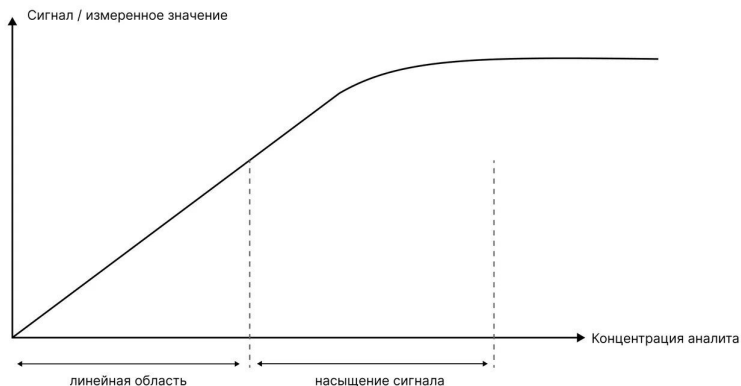


Рисунок 19.2. Потеря линейности при насыщении сигнала

На рисунке видно, что в начальной области сигнал растет пропорционально концентрации аналита. Затем рост замедляется и выходит на плато. Если область насыщения ошибочно включена в заявленный диапазон измерений, высокие концентрации могут быть интерпретированы как более низкие, чем они есть на самом деле. Поэтому верхняя граница диапазона должна быть подтверждена экспериментально.

Для некоторых иммунологических методов особенно важен hook-эффект. При очень высокой концентрации аналита сигнал может не просто выходить на плато, а снижаться. Основная причина состоит в избытке аналита относительно

связывающих реагентов: свободный аналит насыщает анти-тела и препятствует образованию полноценного иммунного комплекса, от которого зависит сигнал. В результате чрезвычайно высокий уровень аналита может дать результат, похожий на умеренный или низкий. Выраженность hook-эффекта зависит от соотношения аналита и реагентов, аффинности антител, времени инкубации, порядка добавления компонентов и разведения образца.

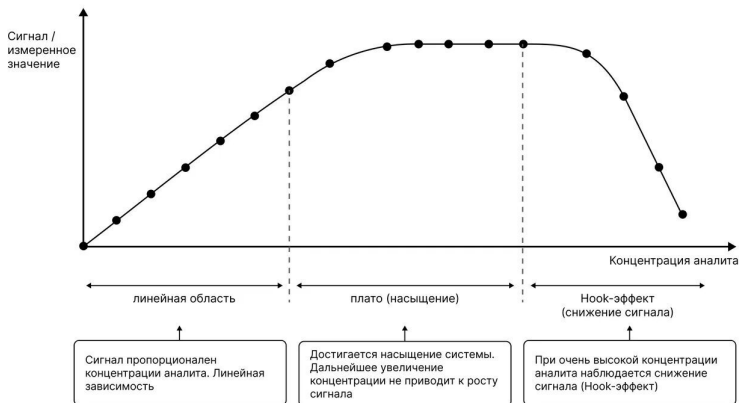


Рисунок 19.3. Хук-эффект при очень высокой концентрации аналита

Hook-эффект требует специальной проверки, если мето-

дологический принцип изделия делает его возможным. В таких случаях необходимо оценивать высокие концентрации аналита, описывать признаки подозрения на результат вне диапазона и устанавливать действия пользователя: разведение образца, повторное исследование, выдача результата «больше чем» или иные процедуры, предусмотренные изготовителем

Разведение образцов

Если для высоких значений предусмотрено разведение образца, эта процедура должна быть подтверждена. Нельзя автоматически считать, что разведение всегда сохраняет правильность результата. После разведения может измениться матрица, восстановление аналита, фон, интерференция или стабильность. Поэтому необходимо показать, что результат после разведения и пересчета соответствует ожидаемому значению в пределах установленного критерия.

В инструкции должно быть указано, когда требуется разведение, каким растворителем или материалом оно выполняется, какой коэффициент применяется, как пересчитывается результат и в каких пределах такой пересчет допустим. Если разведение подтверждено только для определенного интервала, нельзя распространять его на любые высокие значения.

Аналитический диапазон и клинически значимая область

Аналитический диапазон и клинически значимая область

не всегда совпадают. Аналитический диапазон показывает, где изделие способно выдавать надежный результат. Клинически значимая область показывает, где результат влияет на медицинское решение. Иногда клинически значимая область уже аналитического диапазона, и тогда не все измеряемые значения имеют одинаковое значение для решения. Иногда клинически значимая область шире подтвержденного диапазона, и тогда назначение изделия или правила интерпретации должны быть ограничены.

Результат вне диапазона не должен интерпретироваться как обычное количественное значение. Для нижней зоны может требоваться формулировка «ниже предела количественного определения». Для верхней зоны — «выше верхней границы диапазона», разведение или повторное исследование. Такие правила должны быть заранее установлены и отражены в документации.

Типовые проблемы линейности и диапазона

Схема 19.2 — Типовые проблемы линейности и диапазона измерений

— Неподтвержденная нижняя граница диапазона

— Как проявляется: Изделие выдает числовые результаты там, где прецизионность недостаточна; **Почему это важно:** Возникает риск ложной количественной интерпретации низких значений.

— Насыщение сигнала — Как проявляется: При

высоких концентрациях результат перестает возрастать

пропорционально; **Почему это важно:** Высокий уровень аналита может быть занижен.

— **Hook-эффект** — **Как проявляется:** Очень высокая концентрация аналита дает ложно низкий результат в некоторых иммунологических методах; **Почему это важно:** Возможна опасная неправильная интерпретация результата.

— **Некорректное разведение** — **Как проявляется:** Результат после разведения не соответствует ожидаемому восстановлению; **Почему это важно:** Пересчитанное значение может быть ошибочным.

— **Диапазон не покрывает клинически важные уровни** — **Как проявляется:** Подтвержденный интервал уже, чем область принятия решений; **Почему это важно:** Назначение изделия становится шире доказанных возможностей.

— **Отклонение от линейности около порога** — **Как проявляется:** Результат нестабилен в области клинического решения; **Почему это важно:** Возможна смена категории результата.

— **Использование неподходящей матрицы при проверке** — **Как проявляется:** Линейность подтверждена на модельном материале, не отражающем клинический образец; **Почему это важно:** Данные могут быть неприменимы к заявленному материалу.

Особенности для полуколичественных и алгорит-

мических изделий

Для полуколичественных тестов понятие линейности может иметь ограниченное значение, но диапазон категорий и устойчивость переходов между ними остаются важными. Если результат выражается как «низкий», «средний», «высокий» или в виде балльной шкалы, необходимо показать, что категории соответствуют ожидаемому изменению уровня аналита и воспроизводимо отделяются друг от друга. В таких случаях основным вопросом становится не идеальная линейная зависимость, а надежность классификации по уровням и клинический смысл каждой категории.

Для алгоритмических и программных изделий диапазон может относиться не только к концентрации аналита, но и к входным данным. Например, алгоритм может быть проверен на изображениях определенного качества, файлах секвенирования с определенной глубиной покрытия, сигналах в определенных пределах, значениях лабораторных показателей в определенном интервале. Если входные данные выходят за пределы области, на которой алгоритм был разработан и проверен, результат может стать менее надежным. Поэтому для таких изделий необходимо описывать область допустимых входных данных и правила обработки неподходящих или пограничных случаев.

Что правильно и что неправильно при подтверждении линейности и диапазона

Схема 19.3 — Типовые ошибки при подтверждении

линейности и диапазона измерений

— **Выбор уровней** — **Правильно:** Покрыть весь заявленный диапазон, включая нижнюю и верхнюю границу; **Неправильно:** Проверить только среднюю часть диапазона.

— **Клинические пороги** — **Правильно:** Включить уровни около медицински значимых решений; **Неправильно:** Не проверять область cut-off или зону принятия решения.

— **Материал** — **Правильно:** Использовать заявленную матрицу или обосновать модельный материал; **Неправильно:** Проверить линейность в буфере и автоматически перенести на клинический материал.

— **Разведение** — **Правильно:** Подтвердить восстановление после разведения; **Неправильно:** Считать разведение автоматически корректным.

— **Верхняя граница** — **Правильно:** Проверить насыщение, hook-эффект и правила выдачи результата выше диапазона; **Неправильно:** Выдавать точные значения выше подтвержденного диапазона.

— **Нижняя граница** — **Правильно:** Связать с пределом обнаружения, пределом количественного определения и прецизионностью; **Неправильно:** Выдавать числа ниже уровня количественной надежности.

— **Критерий** — **Правильно:** Оценивать отклонения по уровням и их клинический смысл; **Неправильно:** Считать высокий достаточным доказательством.

— **Полуколичественный результат** — *Правильно:*

Проверить устойчивость категорий и переходов между ними; Неправильно: Описать категории без подтверждения их воспроизводимости.

— **Алгоритмический результат** — *Правильно:* Опи-

сать допустимый диапазон входных данных; Неправильно: Применять алгоритм к данным за пределами проверенной области без дополнительной оценки.

Документирование в отчете

В отчете по подтверждению линейности и диапазона измерений необходимо описать заявленный диапазон, клинически значимые уровни, использованные материалы, способ подготовки образцов, число уровней, число повторов, ожидаемые и измеренные значения, расчет отклонений, графическое представление результатов и критерии приемлемости. Если использовались разведения, необходимо описать схему разведения, матрицу, коэффициенты пересчета и подтверждение восстановления.

Также нужно указать, какие зоны диапазона признаны подтвержденными, а какие требуют ограничения. Если нижняя часть диапазона не обеспечивает количественную надежность, это должно быть отражено в выводе и инструкции. Если верхняя часть требует разведения или выдачи результата «больше чем», это должно быть описано. Если выявлены признаки насыщения сигнала или hook-эффекта, необходимо указать, как они контролируются.

Заявленный диапазон в инструкции и регистрационной документации не должен быть шире подтвержденного. Если фактически подтвержден диапазон от 125 до 1000 копий/мл, нельзя заявлять количественный диапазон от 50 до 1000 копий/мл без дополнительных данных. Если результат ниже нижней границы может быть обнаружен, но не количественно определен, это должно быть ясно указано. Если результат выше верхней границы требует разведения, пользователь должен понимать, когда и как это делать.

Линейность и диапазон измерений, таким образом, являются как техническими характеристиками метода, так и частью доказательной связи между аналитическим результатом и клиническим применением. Они показывают, в каких пределах результат можно считать количественно надежным, где требуется осторожность, где нужно повторное исследование или разведение, а где результат не должен использоваться для заявленного вывода.

Итог главы

Линейность показывает способность аналитической системы давать результаты, пропорциональные содержанию аналита в пределах установленного интервала. Диапазон измерений определяет область, в которой результат изделия может использоваться с подтвержденной правильностью, прецизионностью и пригодностью для заявленного применения. Эти характеристики должны подтверждаться на релевантных материалах, в нескольких уровнях аналита, с за-

ранее установленными критериями приемлемости и с учетом клинически значимых зон. Заявленный диапазон не должен быть шире подтвержденного, а результаты вне диапазона должны иметь четко описанные правила обработки и интерпретации.

Глава 20. Предел обнаружения и предел количественного определения

Предел обнаружения и предел количественного определения относятся к ключевым характеристикам аналитической эффективности медицинских изделий для диагностики *in vitro*, особенно если изделие должно выявлять низкие уровни аналита или работать в области, близкой к клиническому порогу. Эти характеристики определяют нижнюю границу возможностей изделия: где оно еще не отличает сигнал от фона, где уже способно обнаружить аналит, и где может и обнаружить, и количественно измерить его с приемлемой надежностью.

В практической работе часто смешивают понятия «обнаружить» и «измерить». Для качественного теста может быть достаточно показать, что аналит выявляется (обнаруживается) при низкой концентрации с установленной вероятностью. Для количественного теста этого недостаточно. Если изделие выдает числовое значение (измеряется аналит), нужно показать, что на данном уровне результат имеет приемлемую правильность и прецизионность. Поэтому предел обнаружения и предел количественного определения отвечают на разные вопросы и должны быть подтверждены раз-

ными подходами.

Нижняя область аналитической системы обычно описывается через три связанных понятия: предел холостой пробы, предел обнаружения и предел количественного определения. Предел холостой пробы характеризует уровень сигнала, который может быть получен при отсутствии аналита. Предел обнаружения показывает минимальный уровень аналита, который можно надежно отличить от холостой пробы или фоновых сигнала. Предел количественного определения показывает минимальный уровень аналита, при котором результат можно количественно оценить с приемлемой точностью и прецизионностью.

Предел холостой пробы — это наибольший ожидаемый результат, который может быть получен для образца без аналита при заданной вероятности ошибки.

Предел обнаружения — это минимальный уровень аналита, который может быть надежно отличен от холостой пробы или фоновых сигнала.

Предел количественного определения — это минимальный уровень аналита, при котором результат может быть количественно измерен с заранее установленной приемлемой правильностью и прецизионностью.

Нижняя область аналитической системы

Эти понятия особенно важны для изделий, работающих с низкими концентрациями аналита. Если нижняя граница установлена неправильно, изделие может выдавать ложное

ощущение надежности там, где аналитическая система фактически работает нестабильно. Очень низкий уровень сигнала может быть принят за наличие аналита, хотя он не отличается от фона. Или наоборот, аналит может присутствовать, но изделие не обнаруживает его с достаточной вероятностью. Для количественного изделия дополнительный риск состоит в том, что низкое числовое значение может быть воспринято как точный результат, хотя на самом деле вариабельность в этой зоне слишком велика.

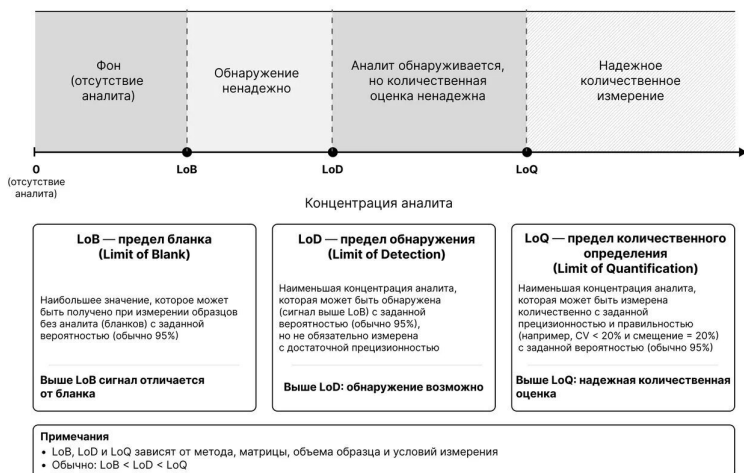


Рисунок 20.1. Шкала LoB/LoD/LoQ

Рисунок 20.1 показывает, что предел обнаружения и пре-

дел количественного определения не являются одним и тем же. Между LoD и LoQ может существовать зона, где аналит уже обнаруживается, но количественное значение еще недостаточно надежно. Для качественного изделия такая зона может быть приемлемой, если назначение связано только с обнаружением. Для количественного изделия необходимо ясно указать, где начинается область надежной количественной интерпретации.

Предел холостой пробы важен потому, что любой аналитический метод имеет фон. Даже при отсутствии аналита прибор или реакционная система могут давать небольшой сигнал. Этот сигнал может быть связан с шумом прибора, неспецифическим связыванием, фоновым флуоресцентным сигналом, загрязнением, матрицей или другими факторами. Поэтому нельзя считать любой ненулевой сигнал доказательством наличия аналита. Сначала нужно понять, какие значения характерны для холостых образцов.

Предел обнаружения строится уже относительно этого фона. Он должен быть достаточно выше холостой зоны, чтобы вероятность ошибочного обнаружения и вероятность пропуска низкого уровня аналита были приемлемыми. Для разных изделий это может подтверждаться разными способами. В одних случаях используют статистическую оценку распределений холостых и низкоположительных образцов. В других — серии разведений и расчет вероятности обнаружения. В молекулярной диагностике часто оценивают концен-

трацию, при которой аналит обнаруживается не менее чем в заданной доле повторов, например в 95 % случаев.

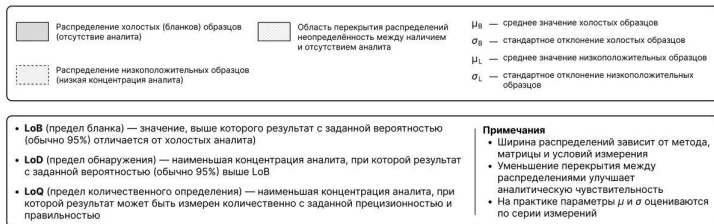
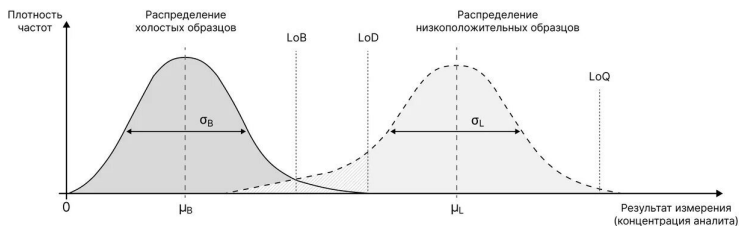


Рисунок 20.2. Распределение холостых и низкоположительных образцов

Рисунок 20.2 показывает общую идею: результаты холостых образцов имеют собственный разброс, а низкоположительные образцы дают распределение результатов выше фона. Если эти распределения сильно перекрываются, надежное обнаружение невозможно. Чем лучше отделены низкоположительные результаты от холостых, тем увереннее можно установить предел обнаружения.

Расчетные подходы к пределу холостой пробы и пределу обнаружения

Существуют разные подходы к расчету LoB и LoD. Конкретная методика зависит от типа изделия, распределения данных, количества измерений, вида результата и применимых стандартов или руководств. В учебных целях удобно показать общий статистический подход.

Если результаты холостых образцов приблизительно нормально распределены, предел холостой пробы может быть оценен как:

$$\text{LoB} = \bar{x}_{\text{blank}} + c \times \text{SD}_{\text{blank}},$$

где: — среднее значение результатов холостых образцов;
— стандартное отклонение результатов холостых образцов;
— коэффициент, выбранный в зависимости от заданной вероятности ошибки.

Для приближенной оценки часто используют коэффициент около 1,645 для одностороннего 95-го перцентиля при нормальном распределении:

$$\text{LoB} = \bar{x}_{\text{blank}} + 1,645 \times \text{SD}_{\text{blank}}.$$

Предел обнаружения может быть оценен с учетом вариабельности низкоположительных образцов:

$$\text{LoD} = \text{LoB} + c \times \text{SD}_{\text{low}}$$

где — стандартное отклонение результатов низкоположительных образцов.

При использовании того же коэффициента 1,645 формула может быть записана как:

$$\text{LoD} = \text{LoB} + 1,645 \times \text{SD}_{\text{low}}.$$

В практической работе важно понимать, что эти формулы являются не «магической процедурой», а способом оценить границу между фоном и низким уровнем аналита. Если данные не соответствуют предпосылкам выбранного метода, распределение резко асимметрично, число наблюдений мало или матрица не соответствует клиническому материалу, расчет может быть методологически слабым.

Пример 20.1 — Условный расчет предела холостой пробы и предела обнаружения

Для оценки нижней области метода исследовали холостые образцы, не содержащие аналит, и низкوپоложительные образцы с малым содержанием аналита. Получены следующие данные:

- *среднее значение холостых образцов: 0,20 Ед/мл;*
- *стандартное отклонение холостых образцов: 0,05 Ед/мл;*
- *стандартное отклонение низкوپоложительных образцов: 0,08 Ед/мл.*

Предел холостой пробы:

$$\text{LoB} = 0,20 + 1,645 \times 0,05$$

$$\text{LoB} = 0,20 + 0,082 = 0,282 \text{ Ед/мл.}$$

Предел обнаружения:

$$\text{LoD} = 0,282 + 1,645 \times 0,08$$

$$\text{LoD} = 0,282 + 0,132 = 0,414 \text{ Ед/мл.}$$

В этом примере результат ниже 0,282 Ед/мл находится

в зоне, где он может соответствовать фоновому сигналу. Результат около 0,414 Ед/мл и выше может рассматриваться как уровень, на котором аналит надежно отличается от холостой пробы при заданных условиях оценки. Однако это еще не означает, что начиная с 0,414 Ед/мл результат можно надежно выдавать как количественный. Для этого требуется подтверждение предела количественного определения.

Предел количественного определения

Предел количественного определения имеет особое значение для изделий, выдающих числовой результат. Он показывает, начиная с какого уровня аналит можно измерять с заранее установленной приемлемой правильностью и прецизионностью. Если изделие только обнаруживает наличие аналита, LoQ может не быть центральной характеристикой. Но если результат заявлен как количественный, LoQ становится критически важным.

На практике LoQ устанавливают через оценку низких уровней аналита. Для каждого уровня проводят повторные измерения и оценивают, соответствуют ли результаты критериям правильности и прецизионности. Например, производитель может заранее установить, что на уровне LoQ смещение не должно превышать $\pm 20\%$, а коэффициент вариации не должен превышать 20% . Эти значения приведены только как пример: реальные критерии должны быть обоснованы назначением изделия, клиническим смыслом результата

и применимыми требованиями.

Схема 20.1 — Различие LoB, LoD и LoQ

— **LoB** — *На какой вопрос отвечает: Какой сигнал возможен при отсутствии аналита?; Что требуется для подтверждения: Измерения холостых образцов и оценка фоновое распределения.*

— **LoD** — *На какой вопрос отвечает: Какой минимальный уровень аналита можно надежно обнаружить?; Что требуется для подтверждения: Измерения низкоположительных образцов и отделение от фона.*

— **LoQ** — *На какой вопрос отвечает: С какого уровня аналит можно надежно количественно измерять?; Что требуется для подтверждения: Подтверждение правильности и прецизионности на низких уровнях.*

Важно не заявлять количественный результат ниже LoQ. Если аналит обнаруживается ниже предела количественного определения, это может быть описано как «обнаружено, ниже диапазона количественного определения» или аналогичной формулировкой, если такая интерпретация предусмотрена назначением изделия. Но выдача точного числового значения ниже LoQ может создать ложное впечатление надежной количественной оценки.

Для LoQ удобно формулировать не одну универсальную расчетную формулу, а условие приемлемости: минимальный уровень может быть принят как предел количественного определения только тогда, когда одновременно выполня-

ются критерии по смещению и прецизионности:

$$\text{Bias}(\%) \mid \leq \text{Bias_max}$$

$$\text{CV}(\%) \leq \text{CV_max}.$$

Пример 20.2 — Установление LoQ по критериям правильности и прецизионности

Для количественного теста исследованы четыре низких уровня аналита. Критерии приемлемости для LoQ установлены заранее: смещение не более $\pm 20\%$, коэффициент вариации не более 20% .

Схема 20.2 — Пример оценки предела количественного определения

— **0,30** — Смещение, %: -28% ; CV, %: 32% ; Оценка: Неприемлемо.

— **0,50** — Смещение, %: -18% ; CV, %: 24% ; Оценка: Неприемлемо по CV.

— **0,75** — Смещение, %: -12% ; CV, %: 18% ; Оценка: Приемлемо.

— **1,00** — Смещение, %: -8% ; CV, %: 12% ; Оценка: Приемлемо.

В этом примере минимальный уровень, на котором одновременно выполняются критерии по смещению и прецизионности, составляет $0,75$ Ед/мл. Следовательно, LoQ может быть установлен как $0,75$ Ед/мл при условии, что дизайн исследования, материал и число повторов достаточны для такого вывода.

Корректный вывод должен быть сформулирован ограни-

ценно:

Предел количественного определения подтвержден на уровне 0,75 Ед/мл. На уровнях 0,30 и 0,50 Ед/мл количественная надежность не подтверждена. Результаты ниже 0,75 Ед/мл не должны выдаваться как точные количественные значения без специальных правил интерпретации.

Особенности качественных тестов

Для качественных тестов центральным обычно является предел обнаружения, а не предел количественного определения. Если изделие выдает результат «обнаружено/не обнаружено», задача состоит в том, чтобы показать, при каком уровне аналита положительный результат получается с установленной вероятностью. При этом важно использовать не только один низкоположительный образец, а серию концентраций около предполагаемого предела обнаружения.

Для качественного теста недостаточно показать, что аналит был обнаружен один раз на очень низком уровне. Нужно оценить воспроизводимость обнаружения. Если на концентрации 10 копий/мл результат положительный в 1 из 20 повторов, такой уровень нельзя считать пределом обнаружения. Если на концентрации 50 копий/мл результат положительный в 19 из 20 повторов, такой уровень может быть рассмотрен как кандидатный LoD при условии, что критерий приемлемости установлен заранее.

В качественных тестах особенно важна зона около LoD. Слабоположительные образцы могут обнаруживаться постоянно. Это не всегда означает ошибку, но должно быть отражено в характеристиках изделия. Пользователь должен понимать, что результат вблизи предела обнаружения имеет вероятностную природу: чем ниже концентрация аналита, тем выше риск непостоянного обнаружения.

Особенности ПЦР

Для молекулярных тестов, включая ПЦР и другие методы, предел обнаружения часто выражается в копиях на реакцию, копиях на миллилитр, международных единицах на миллилитр или иной единице, соответствующей материалу и назначению изделия. Здесь особенно важно ясно указать, к какому этапу относится значение: к исходному образцу, экстракту, реакции амплификации или пересчитанной концентрации в клиническом материале. При использовании ГСО содержания копий ДНК, разработанных ВНИИМ им. Д.И. Менделеева, необходимо показывать связь аттестованного значения стандартного образца с расчетной концентрацией копий в образце, экстракте или реакционной смеси.

Для молекулярных тестов предел обнаружения часто подтверждают как минимальную концентрацию аналита, обнаруживаемую с заданной вероятностью, например в 95 % повторов. Это связано с тем, что при низких концентрациях молекулы распределяются вероятностно, а отдельная реакция может не получить целевую копию даже при наличии

аналита в исходном материале. Поэтому LoD для молекулярного теста имеет вероятностный характер.

Схема 20.3 — Пример подтверждения LoD для качественного NAAT-теста

— **10** — *Число повторов: 20; Положительных результатов: 8; Доля обнаружения: 40 %.*

— **25** — *Число повторов: 20; Положительных результатов: 14; Доля обнаружения: 70 %.*

— **50** — *Число повторов: 20; Положительных результатов: 19; Доля обнаружения: 95 %.*

— **100** — *Число повторов: 20; Положительных результатов: 20; Доля обнаружения: 100 %.*

В данном примере концентрация 50 копий/мл обнаружена в 19 из 20 повторов, то есть в 95 % случаев. Если заранее установленный критерий состоит в обнаружении не менее 95 % повторов, уровень 50 копий/мл может быть принят как подтвержденный предел обнаружения. Уровни 10 и 25 копий/мл не соответствуют критерию.

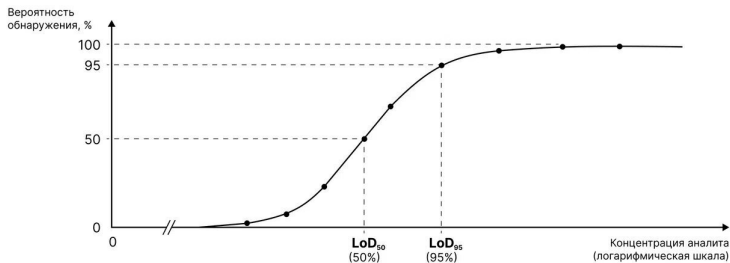
Однако такой вывод возможен только при корректном дизайне исследования. Нужно учитывать матрицу, способ подготовки образцов, экстракцию, объем, внесенный в реакцию, число повторов, партии реагентов, приборы, операторов и возможные ингибиторы. Если LoD подтвержден только в буфере, его нельзя автоматически переносить на плазму, мазок, мокроту, кал или другой клинический материал.

Вероятность обнаружения и LoD 95

Для оценки вероятности обнаружения при разных концентрациях может использоваться пробит-анализ или иной подход моделирования зависимости «концентрация — вероятность положительного результата». Такой подход особенно полезен, когда нужно оценить концентрацию, при которой ожидаемая вероятность обнаружения составляет, например, 95 %.

Рисунок 20.3 показывает типичную логику: при очень низкой концентрации аналит обнаруживается непостоянно, а при увеличении концентрации вероятность положительного результата возрастает. LoD95 соответствует уровню, при котором ожидаемая вероятность обнаружения составляет 95 %. В отчете важно указать не только оценку LoD95, но и доверительный интервал, если он рассчитывался, а также диапазон концентраций, на которых построена модель.

Пробит-анализ не должен использоваться как формальная замена хорошего экспериментального дизайна. Если концентрации выбраны неудачно, повторов мало, матрица нерелевантна или результаты нестабильны из-за технических факторов, модель может дать недостоверную оценку. Поэтому статистическая модель должна опираться на корректно спланированное исследование.



LoD₉₅ (предел обнаружения при 95%) — наименьшая концентрация аналита, при которой результат с заданной вероятностью (обычно 95%) превышает LoB (предел бланка).
 Определяется по экспериментальным данным как концентрация, соответствующая 95 % вероятности обнаружения

Обозначения

- кривая вероятности обнаружения
- — экспериментальные данные
- — уровни вероятности (50%, 95%, 100%)
- — определение LoD₅₀ и LoD₉₅

Примечания

- LoD₉₅ соответствует 95%-ной вероятности того, что результат измерения выше LoB
- LoD₅₀ (медиана вероятности обнаружения) соответствует 50%-ной вероятности
- Шкала концентрации - логарифмическая (обычно порядок величины)
- Значение LoD зависит от метода, матрицы, объема образца и условий измерения

Рисунок 20.3. Вероятность обнаружения и LoD₉₅

Материал и матрица при оценке нижней границы

Оценка LoB, LoD и LoQ должна выполняться в материале, соответствующем заявленному назначению изделия. Матрица может существенно влиять на фоновый сигнал, ингибирование реакции, извлечение аналита, неспецифическое связывание и стабильность. Поэтому предел обнаружения, установленный в буфере или искусственной среде, не всегда отражает предел обнаружения в реальном клиническом материале.

Если изделие заявлено для нескольких материалов, на-

пример плазмы, сыворотки и цельной крови, необходимо подтвердить нижнюю границу для каждого материала или обосновать перенос данных. Для молекулярных тестов особенно важно учитывать этап экстракции и возможные ингибиторы. Для иммунологических тестов — матричные эффекты, перекрестные реакции и фоновый сигнал. Для сухих пятен крови — гематокрит, объем пятна, условия высушивания и хранения.

Типовые ошибки при подтверждении LoD и LoQ

Схема 20.4 — Типовые ошибки при подтверждении предела обнаружения и предела количественного определения

— **Различение LoD и LoQ** — *Правильно: Отдельно указать уровень обнаружения и уровень надежного количественного измерения; Неправильно: Считать, что обнаружение автоматически означает возможность количественного измерения.*

— **Выбор материала** — *Правильно: Использовать заявленную клиническую матрицу или обосновать модельный материал; Неправильно: Подтвердить LoD в буфере и перенести на клинические образцы без обоснования.*

— **Число повторов** — *Правильно: Использовать достаточное число повторов около предполагаемого предела; Неправильно: Делать вывод по одному или нескольким единичным обнаружениям.*

— **Качественный тест** — *Правильно: Оценивать долю*

обнаружения на нескольких концентрациях; **Неправильно:** Указывать самую низкую концентрацию, где был хотя бы один положительный результат.

— **Количественный тест** — **Правильно:** Подтвердить правильность и прецизионность на уровне LoQ ; **Неправильно:** Выдавать точные числовые значения ниже уровня количественной надежности.

— **НААТ-тест** — **Правильно:** Указать копии/реакцию, копии/мл или другие единицы и способ пересчета; **Неправильно:** Неясно указывать LoD без связи с объемом образца и этапом реакции.

— **Пробит-анализ** — **Правильно:** Использовать корректно выбранные концентрации и достаточные повторы; **Неправильно:** Строить модель на недостаточных или нерелевантных данных.

— **Несколько материалов** — **Правильно:** Подтвердить каждый материал или обосновать сопоставимость; **Неправильно:** Проверить один материал и заявить все остальные.

— **Интерпретация результата** — **Правильно:** Описать правила «ниже LoD », «обнаружено ниже LoQ », «ниже диапазона»; **Неправильно:** Выдавать пользователю неоднозначные низкие значения.

Эта схема показывает, что нижняя граница метода требует не только расчета, но и правильной интерпретации. Ошибки в этой области особенно опасны, потому что пользователь

может придать медицинское значение результату, который аналитически недостаточно надежен.

Документирование в отчете

В отчете по подтверждению предела обнаружения и предела количественного определения необходимо описать назначение изделия, тип результата, материал, используемые образцы, способ приготовления низких концентраций, число уровней, число повторов, партии реагентов, приборы, операторов, условия выполнения и статистический подход. Если использовались холостые образцы, нужно описать их состав и соответствие заявленной матрице. Если использовались низкоположительные образцы, необходимо указать, как установлены их концентрации.

Для LoV нужно представить данные холостых образцов и расчет фоновой границы. Для LoD — данные низкоположительных уровней, расчет или подтверждение вероятности обнаружения. Для LoQ — данные по правильности и прецизионности на низких уровнях и критерии приемлемости. Для NAAT-тестов нужно отдельно указать единицы выражения LoD, объем исходного материала, объем экстракта, объем реакции и способ пересчета, если он используется.

Если предел обнаружения или предел количественного определения подтвержден только для определенного материала, это должно быть отражено. Если для разных материалов получены разные значения, нельзя указывать одно общее значение без обоснования. Если результат ниже LoQ мо-

жет быть обнаружен, но не должен выдаваться как количественный, это должно быть ясно описано в инструкции и отчете.

Заявления в инструкции должны соответствовать подтвержденным данным. Если LoD подтвержден как 50 копий/мл при 95 % обнаружения, нельзя заявлять надежное обнаружение 10 копий/мл только потому, что на этом уровне были отдельные положительные результаты. Если LoQ подтвержден как 0,75 Ед/мл, нельзя выдавать точные количественные результаты ниже этого уровня без дополнительных данных. Нижняя область метода должна быть описана так, чтобы пользователь понимал границы результата.

Итог главы

Предел холостой пробы, предел обнаружения и предел количественного определения описывают нижнюю область аналитических возможностей медицинского изделия для диагностики *in vitro*. Предел холостой пробы характеризует фоновый сигнал при отсутствии аналита, предел обнаружения — минимальный уровень, который можно надежно отличить от фона, а предел количественного определения — минимальный уровень, на котором аналит можно количественно измерить с приемлемой правильностью и прецизионностью. Для качественных тестов центральным обычно является надежное обнаружение, для количественных — надежное количественное измерение. Эти характеристики должны подтверждаться в релевантной матрице, с достаточным чис-

лом повторов, заранее установленными критериями и ясными правилами интерпретации результатов ниже LoD, около LoD и ниже LoQ.

Глава 21. Аналитическая специфичность и интерференция

Аналитическая специфичность и интерференция относятся к тем характеристикам аналитической эффективности, которые определяют, насколько результат изделия устойчив к влиянию посторонних, близких, сопутствующих или мешающих факторов. Если правильность, прецизионность, линейность и предел обнаружения показывают, насколько надежно изделие получает результат при заданных условиях, то аналитическая специфичность и интерференция отвечают на другой вопрос: не получает ли изделие ошибочный результат из-за того, что реагирует не только на заявленный аналит или потому, что на результат влияет постороннее вещество, состояние матрицы или иной фактор.

Для медицинских изделий для диагностики *in vitro* это имеет принципиальное значение. Ложноположительный или ложноотрицательный результат может возникнуть не только из-за недостаточной чувствительности или плохой воспроизводимости. Он может быть связан с перекрестной реактивностью, ингибированием реакции, матричным эффектом, лекарственным препаратом, эндогенным веществом, близким микроорганизмом, сходной последовательностью нуклеиновой кислоты, аутоантителами, гемолизом, липемией,

высоким уровнем белка, антикоагулянтом, консервантом, качеством изображения или особенностью входных данных алгоритма. Поэтому оценка аналитической специфичности и интерференции должна быть связана с реальными условиями применения изделия.

Аналитическая специфичность — это способность медицинского изделия для диагностики *in vitro* определять заявленный аналит или объект исследования без значимого влияния со стороны других веществ, организмов, последовательностей, признаков или факторов.

Перекрестная реактивность — это реакция изделия с объектами, сходными с целевым аналитом или мишенью, но не являющимися заявленным объектом определения.

Интерференция — это влияние вещества, компонента матрицы, лекарственного препарата, физиологического состояния, технического фактора или иного условия на результат исследования.

Эти понятия связаны между собой, но не совпадают полностью. Перекрестная реактивность обычно относится к ситуации, когда изделие ошибочно реагирует на другой объект как на целевой аналит. Например, антитело в иммунологическом тесте связывается со структурно сходным белком, а молекулярный тест амплифицирует близкую последовательность другого микроорганизма. Интерференция шире: мешающий фактор может не быть похожим на аналит, но изменять результат. Например, гемолиз может влиять на

биохимическое измерение, билирубин — на фотометрический сигнал, гепарин — на амплификацию, а высокая вязкость или клеточность материала — на обработку образца.

Аналитическая специфичность особенно важна для качественных и молекулярных тестов, где результат часто интерпретируется как наличие или отсутствие объекта. Если изделие выявляет заявленную мишень, а также близкий организм, близкую последовательность или сходный антиген, возникает риск ложноположительного результата. Для количественных тестов специфичность также имеет значение: если метод измеряет и заявленный аналит, и метаболиты, фрагменты, родственные молекулы или лекарственные вещества, числовой результат может быть завышен или занижен.

Интерференция может приводить как к ложноположительным, так и к ложноотрицательным результатам. В одном случае мешающий фактор повышает сигнал и создает видимость наличия или более высокого уровня аналита. В другом — подавляет реакцию и приводит к занижению результата или пропуску аналита. В третьем — увеличивает неопределенность и делает результат нестабильным. Поэтому при оценке интерференции важно анализировать как наличие статистического различия, так и клиническое значение изменения результата.

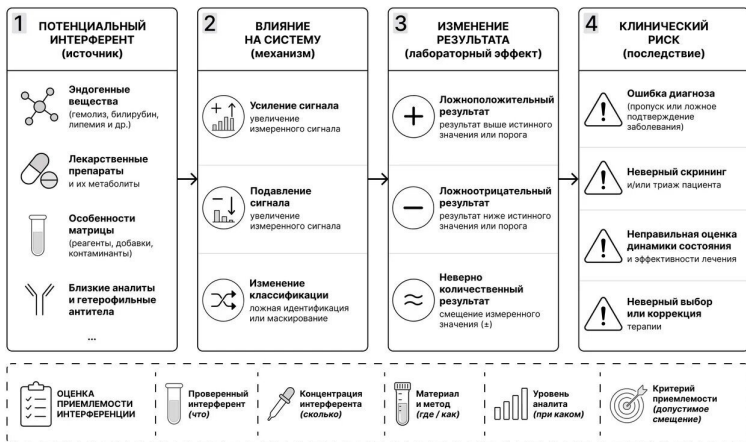


Рисунок 21.1. Интерферент – влияние на систему – изменение результата – клинический риск

Типы влияющих факторов

Потенциальные интерференты и перекрестно реагирующие объекты выбирают исходя из назначения изделия, материала, метода и целевой популяции. Универсального списка, одинакового для всех изделий, не существует. Для биохимического анализа крови важны одни факторы, для молекулярного теста — другие, для иммунологического анализа — третьи, для цифрового алгоритма — четвертые.

Эндогенные интерференты — это вещества или состояния, присутствующие в организме или образце пациента. К

ним могут относиться гемоглобин при гемолизе, билирубин при иктеричности, липиды при липемии, высокий уровень белка, ревматоидный фактор, аутоантитела, гетерофильные антитела, эндогенные метаболиты, гормоны, антитела после вакцинации или инфекционного контакта. Их значение зависит от метода: один и тот же фактор может быть критичным для одного изделия и не иметь значения для другого.

Экзогенные интерференты — это вещества, поступающие извне. Это могут быть лекарственные препараты, антикоагулянты, консерванты, компоненты транспортных сред, дезинфицирующие вещества, контрастные препараты, биотин, добавки, вещества из пробирок, реагенты пробоподготовки или загрязнители. Для изделий, применяемых у пациентов, получающих терапию, лекарственные интерференции могут иметь особое значение. Для тестов, используемых в определенной матрице, важны добавки и условия транспортировки.

К биологически близким объектам относятся микроорганизмы, последовательности, антигены, антитела или молекулы, которые могут быть сходны с целевой мишенью. В инфекционной диагностике это близкородственные виды, штаммы, непатогенные представители нормальной микрофлоры или организмы, часто присутствующие в том же локусе. В молекулярной диагностике — сходные участки генома, псевдогены, родственные варианты, последовательности других организмов. В иммунологических тестах — структурно близкие белки, фрагменты, изоформы или метаболиты.

ТЫ.

Схема 21.1 — Основные типы факторов, влияющих на аналитическую специфичность и интерференцию

— **Эндогенные вещества** — *Примеры:* Гемоглобин, билирубин, липиды, белок, метаболиты; *Возможное влияние:* Изменение сигнала, фон, матричный эффект.

— **Иммунологические факторы** — *Примеры:* Гетерофильные антитела, аутоантитела, ревматоидный фактор; *Возможное влияние:* Ложноположительный или ложноотрицательный результат в иммунологических тестах.

— **Экзогенные вещества** — *Примеры:* Лекарства, биотин, антикоагулянты, консерванты, транспортные среды; *Возможное влияние:* Искажение реакции, подавление сигнала, изменение матрицы.

— **Близкие биологические объекты** — *Примеры:* Родственные микроорганизмы, нормальная микробиота, близкие антигены; *Возможное влияние:* Перекрестная реактивность или ложноположительный результат.

— **Сходные последовательности** — *Примеры:* Псевдогены, близкие участки генома, варианты последовательности; *Возможное влияние:* Неспецифическая амплификация или пропуск мишени.

— **Матричные факторы** — *Примеры:* Вязкость, точность, pH, соли, ингибиторы, деградация материала; *Возможное влияние:* Снижение эффективности реакции

или нестабильность результата.

— **Технические факторы** — *Примеры:* Качество изображения, шум сигнала, артефакты, глубина покрытия; *Возможное влияние:* Ошибка классификации или вычисленного результата.

Эта схема не должна использоваться как универсальный обязательный перечень. Она показывает направления анализа. Для каждого изделия нужно определить, какие факторы действительно вероятны и значимы. Например, биотин важен для некоторых иммунологических систем, но не является универсальным интерферентом для всех методов. Гемолитичность критична для ряда биохимических показателей, но может быть незначимой для других. Близкородственные микроорганизмы особенно важны для инфекционных тестов, но их перечень зависит от мишени и материала.

Планирование испытания интерференции и специфичности

Испытание аналитической специфичности и интерференции должно начинаться с анализа риска. Сначала определяют, какие вещества, организмы, последовательности, матричные состояния или технические факторы могут присутствовать в заявленном материале и повлиять на результат. Затем выбирают условия испытания: уровни аналита, концентрации интерферентов, тип матрицы, число повторов, контрольные образцы и критерии приемлемости.

Поскольку экспериментальная проверка всех потенциаль-

ных интерферентов физически невозможна, отбор факторов для программы испытаний целесообразно проводить в два шага: сначала оценить вероятность встречаемости фактора именно в заявленной популяции и материале, затем — потенциальную величину его влияния на результат в клинически значимой зоне. Приоритет для экспериментальной проверки отдаётся факторам, сочетающим высокую вероятность встречаемости с высоким потенциальным влиянием.

Для оценки интерференции обычно сравнивают результат образца с потенциальным интерферентом и результат сопоставимого контрольного образца без интерферента. Если изделие количественное, оценивают величину отклонения. Если качественное — оценивают изменение категории результата. Если полуколичественное или алгоритмическое — оценивают переход между категориями, индексами или классами.

Важно проверять интерференцию на уровнях аналита, где влияние может быть клинически значимым. Для качественных тестов это область около предела обнаружения или порога принятия решения. Для количественных тестов — низкий, средний, высокий уровни и области клинических порогов. Для мониторинговых тестов — уровни, где интерференция может имитировать динамику состояния. Если интерференцию проверить только на высоком уровне аналита, можно не заметить влияние, которое приводит к ложному отрицательному результату около порога.

Концентрации интерферентов должны быть обоснованы. Они должны отражать клинически возможные или максимально ожидаемые уровни в заявленной популяции и материале. Если концентрация слишком низкая, испытание не покажет реального риска. Если она чрезмерно высокая и маловероятная, выявленное влияние может быть трудно интерпретировать. В любом случае выбор концентраций должен быть описан в программе испытаний.

Для оценки перекрестной реактивности выбирают объекты, которые могут быть сходны с целевым анализом или встречаться в том же материале. В инфекционной диагностике это могут быть близкородственные микроорганизмы, возбудители со сходной клиникой или представители нормальной микрофлоры соответствующего локуса. В молекулярных тестах — близкие последовательности, псевдогены, варианты в области праймеров или зондов. В иммунологических тестах — родственные белки, фрагменты, изоформы или антитела к близким антигенам.

Расчет влияния интерферента

Для количественных изделий влияние интерферента часто оценивают через процент отклонения результата образца с интерферентом от результата контрольного образца без интерферента:

$$\text{Deviation (\%)} = [(X_{\text{test}} - X_{\text{control}}) / X_{\text{control}}] \times 100.$$

где: X_{test} — результат образца с потенциальным интерферентом, X_{control} — результат контрольного образца без

интерферента.

Если результат выражается в логарифмических единицах, отношениях, индексах или иных шкалах, подход к расчету должен соответствовать типу результата. Для качественных тестов вместо процента отклонения может оцениваться доля смены категории: например, сколько слабоположительных образцов стали отрицательными при наличии интерферента.

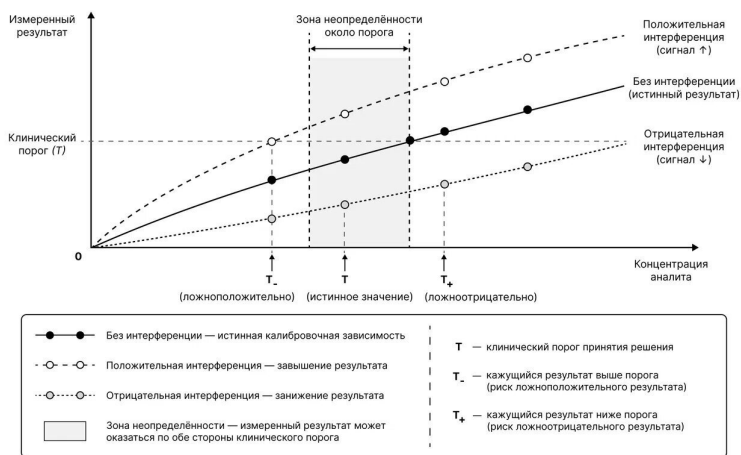


Рисунок 21.2. График влияния интерферента около клинического порога

Пример 21.1 — Оценка влияния интерферента на количественный результат

При оценке интерференции исследовали образец с установленной концентрацией аналита около клинического порога. Контрольный образец без интерферента дал средний результат 100 Ед/мл. Образец с добавлением потенциального интерферента дал средний результат 87 Ед/мл. Заранее установленный критерий приемлемости: отклонение не должно превышать $\pm 10\%$.

Расчет:

$$\text{Deviation (\%)} = [(87 - 100) / 100] \times 100 = -13\%.$$

Отклонение составляет -13% , что превышает критерий $\pm 10\%$. Следовательно, интерферент оказывает значимое отрицательное влияние на результат при данном уровне аналита. Корректный вывод не должен ограничиваться фразой «интерференция выявлена». Нужно указать, при какой концентрации интерферента и на каком уровне аналита выявлено влияние, может ли оно привести к клинически значимому изменению результата и какие ограничения или предупреждения должны быть внесены в документацию.

Если тот же интерферент на высоком уровне аналита давал бы отклонение только -4% , это не отменяло бы проблему около порога. Поэтому уровни аналита при испытании интерференции должны выбираться с учетом клинической интерпретации результата.

Особенности иммунологических изделий

Для иммунологических изделий характерны риски перекрестной реактивности, матричных эффектов и влияния иммунологических факторов. Антитела, используемые в тесте, могут взаимодействовать с родственными молекулами, фрагментами, изоформами или структурно сходными белками. Гетерофильные антитела, ревматоидный фактор, аутоантитела и другие иммунологические компоненты образца могут нарушать связывание реагентов и приводить к ложноположительным или ложноотрицательным результатам.

Отдельное значение имеет биотин, если метод использует биотин-стрептавидиновую систему. В зависимости от формата анализа избыток биотина может приводить как к завышению, так и к занижению результата. Поэтому для таких систем необходимо оценивать влияние биотина на клинически значимых уровнях аналита и указывать ограничения, если они выявлены.

Для иммунологических методов также важен hook-эффект, уже рассмотренный в главе о диапазоне измерений. Он связан не столько с интерференцией внешнего вещества, сколько с особенностью реакции при очень высокой концентрации аналита. Однако с практической точки зрения он также является источником ложного результата и должен учитываться при оценке специфичности и надежности изделия.

Особенности молекулярных изделий

Для молекулярных изделий аналитическая специфичность связана прежде всего с выбором мишени, праймеров, зондов, условий амплификации, алгоритмов анализа последовательности и устойчивости к вариабельности генома. Тест должен выявлять заявленную мишень и не выявлять нерелевантные последовательности. Для инфекционных НААТ-тестов это означает проверку близкородственных организмов, организмов с похожей клинической картиной и микроорганизмов, которые могут присутствовать в заявленном материале.

Не менее важен риск ложноотрицательного результата из-за вариаций последовательности в области праймера или зонда. Если мутация, делеция или генетическая вариабельность нарушает связывание, мишень может не обнаруживаться. Поэтому для молекулярных тестов аналитическая специфичность включает как «не реагировать на чужое», так и «не пропускать заявленное» при ожидаемой вариабельности целевой последовательности.

Интерференция в молекулярных тестах часто связана с ингибиторами амплификации, свойствами матрицы и пробоподготовкой. Кровь, кал, мокрота, мазки, ткань и другие материалы могут содержать вещества, подавляющие реакцию. Поэтому внутренний контроль, контроль экстракции и контроль ингибирования играют важную роль. Если интерферент приводит к ложноотрицательному результату, изде-

лие должно либо выявлять такую ситуацию через контроль, либо содержать ограничения и правила повторного исследования.

Особенности биохимических и химико-аналитических изделий

Для биохимических и химико-аналитических изделий интерференция часто связана с эндогенными веществами, лекарственными препаратами и матричными эффектами. Гемолиз, липемия и иктеричность являются классическими источниками влияния, но не исчерпывают перечень возможных факторов. В зависимости от аналита могут иметь значение метаболиты, лекарственные вещества, антикоагулянты, консерванты, pH, осмолярность, белковый состав и иные свойства образца.

Существенно важно оценивать интерференцию в области медицинских решений. Например, если биохимический показатель используется для стратификации пациента по порогу, даже умеренное смещение может привести к неправильной категории. Если показатель используется для мониторинга, интерференция может имитировать улучшение или ухудшение состояния. Поэтому критерий приемлемости должен быть связан с клинически значимым изменением результата.

Особенности микробиологических изделий

В микробиологических изделиях аналитическая специфичность зависит от того, насколько тест способен отличать

целевой микроорганизм или группу микроорганизмов от близких или сопутствующих видов. Для культуральных, биохимических, иммунологических и молекулярных методов перечень проверяемых объектов будет различаться. Важно учитывать и таксономически близкие виды, и микроорганизмы, которые часто встречаются в том же клиническом материале.

Например, тест для выявления возбудителя в материале из дыхательных путей должен учитывать микроорганизмы, которые могут присутствовать в этом локусе как нормальная микробиота, колонизаторы или возбудители заболеваний со сходной клинической картиной. Для нестерильных локусов особенно важно различать клинически значимое обнаружение и присутствие сопутствующей флоры, если это связано с назначением изделия.

Особенности алгоритмических и цифровых изделий

Для алгоритмических и цифровых изделий понятия специфичности и интерференции требуют адаптации. Здесь «интерферентом» может быть не химическое вещество, а фактор, влияющий на входные данные или классификацию: качество изображения, артефакты окрашивания, шум сигнала, неравномерное освещение, низкая глубина секвенирования, ошибки разметки, нестандартный формат файла, необычная морфология, редкая комбинация признаков или сдвиг распределения данных относительно обучающей вы-

борки.

Если алгоритм классифицирует изображение, необходимо оценивать, как он работает при различном качестве изображения и наличии артефактов. Если алгоритм интерпретирует данные секвенирования, важны глубина покрытия, качество чтений, ошибки выравнивания, псевдогены, сложные участки генома и правила фильтрации вариантов. Если алгоритм рассчитывает риск на основе нескольких показателей, нужно оценить устойчивость результата к допустимым вариациям входных данных.

Для таких изделий аналитическая специфичность означает и способность распознать нужный признак, и способность не принимать посторонние артефакты или нерелевантные паттерны за целевой результат. Поэтому оценка должна включать как «идеальные» данные, так и реалистичные пограничные и трудные случаи.

Схема 21.2 — Примеры факторов для оценки интерференции и специфичности в разных типах МИ ИВД

— **Иммунологический тест** — *Что следует учитывать:* Гетерофильные антитела, ревматоидный фактор, биотин, родственные антигены, hook-эффект, матричные эффекты.

— **Биохимический тест** — *Что следует учитывать:* Гемолиз, липемия, билирубин, лекарственные препараты, антикоагулянты, pH, белковый состав.

— **Молекулярный НААТ-тест** — *Что следует учитывать:* Близкие последовательности, родственные микроорганизмы, ингибиторы амплификации, вариабельность мишени, контаминация.

— **Микробиологический тест** — *Что следует учитывать:* Близкородственные виды, нормальная микрофлора, сопутствующие возбудители, колонизация, нестерильные локусы.

— **Генетический тест** — *Что следует учитывать:* Псевдогены, гомологичные участки, сложные варианты, качество ДНК, глубина покрытия, правила классификации.

— **Цифровое или алгоритмическое изделие** — *Что следует учитывать:* Артефакты изображения, шум данных, качество файла, нестандартные случаи, сдвиг распределения входных данных.

Что правильно и что неправильно при оценке специфичности и интерференции

Схема 21.3 — Типовые ошибки при подтверждении аналитической специфичности и интерференции

— **Выбор интерферентов** — *Правильно:* Выбирать вещества и факторы исходя из материала, метода, популяции и назначения; *Неправильно:* Использовать формальный универсальный список без связи с изделием.

— **Концентрации интерферентов** — *Правильно:* Обосновать клинически возможные или максимально ожи-

даемые концентрации; **Неправильно:** Проверять слишком низкие или необоснованные концентрации.

— **Уровни аналита** — **Правильно:** Оценивать влияние около порога и в клинически значимых зонах; **Неправильно:** Проверять только высокий уровень аналита.

— **Матрица** — **Правильно:** Использовать заявленный клинический материал или обосновать модельную матрицу; **Неправильно:** Проверять в буфере и переносить вывод на клинические образцы.

— **Качественный тест** — **Правильно:** Оценивать изменение категории результата и слабopоложительные образцы; **Неправильно:** Ограничиться явно положительными образцами.

— **Количественный тест** — **Правильно:** Рассчитывать отклонение и сравнивать с заранее установленным критерием; **Неправильно:** Делать вывод без числового критерия приемлемости.

— **Перекрестная реактивность** — **Правильно:** Проверять близкие и клинически вероятные объекты; **Неправильно:** Проверять только нерелевантные или случайные объекты.

— **Молекулярный тест** — **Правильно:** Учитывать вариабельность мишени и ингибиторы; **Неправильно:** Проверить только *in silico* без экспериментального подтверждения, если риск требует испытания.

— **Алгоритмическое изделие** — **Правильно:** Проверить

рять реалистичные артефакты и трудные случаи; **Неправильно:** Оценивать только идеальные входные данные.

— **Вывод** — **Правильно:** Указывать ограничения и условия, при которых влияние выявлено; **Неправильно:** Писать «интерференция отсутствует» без указания проверенных факторов и концентраций.

Оценка специфичности и интерференции должна быть предметной. Недостаточно заявить, что «интерференты проверены». Нужно показать, почему выбраны именно эти факторы, на каких концентрациях, в какой матрице, на каких уровнях аналита и с каким результатом.

Документирование в отчете

В отчете по оценке аналитической специфичности и интерференции необходимо указать перечень проверенных интерферентов, перекрестно реагирующих объектов или влияющих факторов, обоснование их выбора, использованные концентрации, тип материала, уровни аналита, число повторов, контрольные образцы, критерии приемлемости и полученные результаты. Для количественных изделий следует представить величину отклонения. Для качественных — долю смены категории или частоту ложноположительных и ложноотрицательных результатов. Для алгоритмических изделий — изменение классификации, индекса или иного итогового вывода.

Если влияние не выявлено, вывод должен быть сформулирован в пределах проведенного испытания. Корректно пи-

сать не «интерференция отсутствует вообще», а «при проверенных веществах, концентрациях, матрицах и уровнях аналита значимого влияния на результат не выявлено». Если влияние выявлено, нужно указать его характер, величину, условия возникновения и возможное клиническое значение. При необходимости должны быть внесены ограничения в инструкцию, предупреждения, правила повторного исследования или требования к непригодности образца.

Важно согласовать отчет с инструкцией по применению. Если гемолиз, липемия, биотин, антикоагулянт, ингибиторы амплификации, качество изображения или иной фактор влияет на результат, пользователь должен знать, как это распознать и что делать. В некоторых случаях достаточно указать ограничение. В других — требуется встроенный контроль, повторное исследование, альтернативный материал или подтверждающий метод.

Если изделие заявлено для нескольких материалов или популяций, оценка интерференции и специфичности должна покрывать их или содержать обоснование переноса данных. Например, интерференция, оцененная в сыворотке, не всегда применима к цельной крови или сухому пятну крови. Специфичность инфекционного теста, проверенная на одном наборе организмов, может быть недостаточной для другого клинического материала. Устойчивость алгоритма, проверенная на изображениях одного сканера, может не распространяться на другую техническую среду.

В сильной доказательной базе раздел о специфичности и интерференции показывает, что изделие способно выявлять заявленный аналит, а также устойчиво к тем факторам, которые реально могут исказить результат в заявленном применении. Это делает аналитический результат более надежным и снижает риск ошибочной клинической интерпретации.

Итог главы

Аналитическая специфичность и интерференция показывают, насколько результат медицинского изделия для диагностики *in vitro* защищен от влияния посторонних, сходных, сопутствующих или мешающих факторов. Аналитическая специфичность связана со способностью изделия определять именно заявленный аналит или объект исследования. Перекрестная реактивность отражает риск реакции с близкими объектами. Интерференция показывает влияние веществ, матрицы, лекарственных препаратов, физиологических состояний, технических факторов или особенностей входных данных на результат. Эти характеристики должны подтверждаться в релевантной матрице, на клинически значимых уровнях аналита, с обоснованным выбором факторов и заранее установленными критериями приемлемости. Вывод об отсутствии или наличии интерференции должен всегда ограничиваться проверенными условиями и быть согласован с инструкцией по применению изделия.

Глава 22. Стабильность и сохранение аналитической пригодности

Стабильность относится к тем характеристикам аналитической эффективности, которые показывают, сохраняет ли медицинское изделие для диагностики *in vitro* заявленную пригодность во времени и при установленных условиях обращения. Если правильность и прецизионность отвечают на вопрос о надежности результата в заданных условиях выполнения исследования, то стабильность отвечает на другой вопрос: сохранится ли эта надежность после хранения, транспортирования, вскрытия упаковки, подготовки реагента, хранения клинического образца или изменения версии программного обеспечения.

Для медицинского изделия для диагностики *in vitro* это имеет принципиальное значение. Пользователь работает не с идеальным изделием в день разработки метода, а с конкретной серией реагентов, калибраторов, контрольных материалов, расходных компонентов и программной версии. Реагенты хранятся, транспортируются, вскрываются и используются в течение заявленного срока. Клинические образцы могут ожидать исследования, охлаждаться, замораживаться, размораживаться или перевозиться. Программные изделия

могут обновляться. Все эти факторы способны изменить результат, даже если исходные аналитические характеристики были подтверждены корректно.

Поэтому стабильность в доказательной оценке не должна рассматриваться как формальное указание срока годности. Она показывает, при каких условиях изделие, реагент, калибратор, контрольный материал, образец или программный результат сохраняют аналитическую пригодность для заявленного применения. Если такие условия не подтверждены, пользователь может получить результат, который выглядит полноценным, но уже не поддерживается доказательной базой.

***Стабильность** — это способность изделия, реагента, калибратора, контрольного материала, клинического образца, экстракта или программного результата сохранять заявленные характеристики во времени и при установленных условиях хранения, транспортирования и применения.*

Связь стабильности с ранее подтвержденными аналитическими характеристиками

Стабильность не является заменой правильности, прецизионности, предела обнаружения, линейности или интерференционных исследований. Она проверяет, сохраняются ли эти характеристики во времени и при заявленных условиях. Например, реагент может демонстрировать приемлемую правильность и прецизионность в начале срока годности, но

терять чувствительность после длительного хранения. Контрольный материал может иметь установленное значение, но постепенно изменяться при хранении. Клинический образец может быть пригоден сразу после взятия, но становится непригодным после несоблюдения температурного режима.

Именно поэтому стабильность следует связывать с теми характеристиками, которые критичны для назначения изделия. Для качественного теста особенно важно, сохраняется ли выявление слабоположительных образцов около предела обнаружения. Для количественного теста важно, не возникает ли значимое смещение результата и не ухудшается ли прецизионность на клинически значимых уровнях. Для изделий с программным компонентом необходимо понимать, сохраняется ли результат при неизменной версии алгоритма и как влияет обновление версии.

В настоящей главе термин «сопоставимость результата» используется в прикладном смысле: как сохранение аналитического вывода при изменении серии, срока хранения, транспортных условий, открытого флакона, калибровочной системы или версии программного обеспечения. Такая сопоставимость важна, но методически она должна быть связана либо со стабильностью, либо с оценкой прецизионности, рассмотренной ранее.

Виды стабильности

Стабильность не является одной характеристикой. В зависимости от изделия и назначения требуется оценивать раз-

ные виды стабильности. Для реагентного набора важно подтвердить срок годности в закрытой упаковке и стабильность после вскрытия. Для калибратора или контрольного материала — стабильность установленного значения в течение заявленного периода. Для клинического образца — условия, при которых аналит сохраняется до исследования. Для транспортируемого образца — устойчивость к температуре, времени и условиям перевозки. Для замороженного материала — влияние циклов замораживания и оттаивания. Для программного изделия — стабильность результата при неизменной версии алгоритма и сопоставимость после изменений.

Для каждого изделия нужно определить, какие виды стабильности действительно относятся к заявленному применению. Нельзя механически переносить один набор испытаний на все изделия. Для одноразового экспресс-теста важны условия хранения, транспортирования и стабильность в пользовательской среде. Для лабораторной системы с калибраторами и контролями важны стабильность реагентов, калибровки, открытого флакона и межсерийная сопоставимость. Для молекулярного теста важны стабильность нуклеиновой кислоты в материале, условия транспортировки, устойчивость экстракта и влияние циклов замораживания и оттаивания.

Схема 22.1 — Основные виды стабильности в изделиях для диагностики *in vitro*

— **Стабильность закрытого изделия** — **Что оценивается:** Сохранение характеристик в течение срока годности; **Пример вопроса:** Работает ли набор реагентов после хранения до конца заявленного срока?.

— **Стабильность открытого флакона** — **Что оценивается:** Сохранение характеристик после вскрытия упаковки; **Пример вопроса:** Можно ли использовать реагент 30 дней после открытия?.

— **Транспортная стабильность** — **Что оценивается:** Влияние перевозки и температурных колебаний ; **Пример вопроса:** Сохраняются ли характеристики после транспортирования при заявленных условиях?.

— **Стабильность клинического образца** — **Что оценивается:** Сохранение аналита до исследования; **Пример вопроса:** Можно ли хранить образец 24 часа при 2–8 °C?.

— **Стабильность после замораживания/оттаивания** — **Что оценивается:** Влияние циклов замораживания и оттаивания; **Пример вопроса:** Меняется ли результат после двух циклов замораживания?.

— **Стабильность калибратора или контроля** — **Что оценивается:** Сохранение установленного значения; **Пример вопроса:** Остается ли контрольный материал в допустимых пределах в течение срока применения?.

— **Стабильность программной версии** — **Что оценивается:** Сохранение результата при неизменной версии ПО; **Пример вопроса:** Получает ли алгоритм тот же резуль-

тат при повторной обработке тех же данных?.

— **Сопоставимость после изменения** — **Что оценивается:** Влияние обновления реагента, прибора, серии или алгоритма; **Пример вопроса:** Сохраняется ли аналитический вывод после смены версии или серии?.

Этапы, на которых стабильность может влиять на результат

Стабильность может нарушаться не только при хранении готового изделия. Она может влиять на результат на разных этапах: производство, хранение у изготовителя, транспортирование, хранение у пользователя, вскрытие упаковки или подготовка реагента, получение и хранение клинического образца, пробоподготовка, измерение, обработка данных и интерпретация результата.

При проведении испытаний следует учитывать условия, заявленные в инструкции по применению: сроки хранения образца и реагентов, допустимое время от взятия образца до исследования, условия подготовки материала, а также версию программного обеспечения или алгоритма, если результат зависит от программной обработки. Эти параметры не всегда могут быть полноценно подтверждены в рамках конкретного испытания, но они должны быть зафиксированы в протоколе и учтены при оценке применимости полученных результатов.

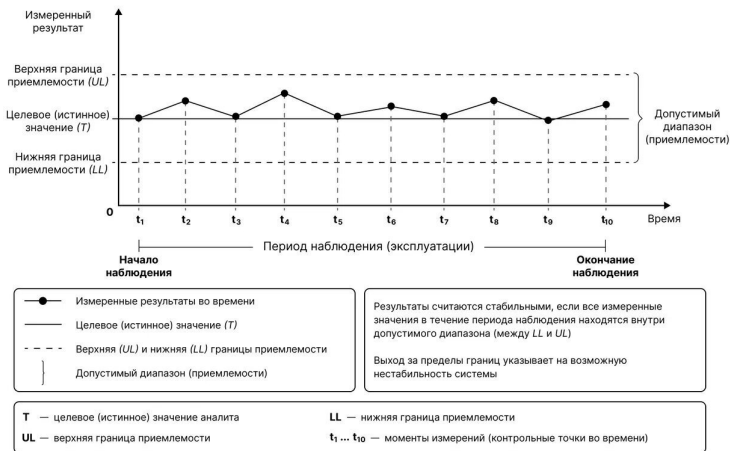


Рисунок 22.1. Стабильность результата во времени с границами приемлемости

Как подтверждают стабильность

Исследование стабильности должно соотноситься с заявленными условиями хранения, транспортирования и применения изделия. В рамках испытаний необходимо фиксировать фактические условия хранения реагентов и образцов, срок от подготовки или взятия образца до исследования, а также иные параметры, способные повлиять на результат. Полное подтверждение срока годности или всех условий транспортирования обычно относится к данным изготовителя и регистрационной документации; в конкретной програм-

ме испытаний могут проверяться только отдельные критические условия, если это предусмотрено ее целью.

Стабильность обычно оценивают во времени. В выбранные точки измеряют контрольные материалы, образцы или реагенты и сравнивают результаты с исходным значением или установленными пределами. Важно включать уровни аналита, которые имеют клиническое значение: низкий, близкий к порогу, средний и высокий. Для качественных тестов особенно важны слабоположительные образцы около предела обнаружения. Для количественных — уровни около клинических решений и границ диапазона.

Критерии приемлемости должны быть установлены заранее. Они могут выражаться как допустимое процентное изменение, абсолютное отклонение, сохранение результата в пределах контрольного интервала, отсутствие смены категории или отсутствие клинически значимого влияния. Выбор критерия зависит от типа результата и назначения изделия.

Для количественных результатов изменение относительно исходного значения может рассчитываться по формуле:

$$\Delta_t(\%) = (X_t - X_0) / (X_0) \times 100.$$

где: $\Delta_t(\%)$ — изменение результата в момент времени t относительно исходного значения; X_t — результат в момент времени t ; X_0 — исходный результат.

Пример 22.1 — Оценка стабильности контрольного материала при хранении

Контрольный материал с исходным значением 100 Ед/мл

хранили при 2–8 ° С . Результаты измеряли в начале исследования и через 1, 3, 6 и 12 месяцев. Критерий приемлемости установлен заранее: изменение относительно исходного значения не должно превышать $\pm 10\%$.

Схема 22.2 — Пример оценки стабильности контрольного материала

— 0 месяцев — *Результат, Ед/мл: 100; Изменение относительно исходного значения: 0 %; Оценка: Исходное значение.*

— 1 месяц — *Результат, Ед/мл: 98; Изменение относительно исходного значения: -2 %; Оценка: Приемлемо.*

— 3 месяца — *Результат, Ед/мл: 96; Изменение относительно исходного значения: -4 %; Оценка: Приемлемо.*

— 6 месяцев — *Результат, Ед/мл: 93; Изменение относительно исходного значения: -7 %; Оценка: Приемлемо.*

— 12 месяцев — *Результат, Ед/мл: 88; Изменение относительно исходного значения: -12 %; Оценка: Неприемлемо.*

Расчет для 6 месяцев:

$$\Delta_{-6}(\%) = (93 - 100)/(100) \times 100 = -7\%.$$

Изменение составляет -7 %, что находится в пределах критерия $\pm 10\%$, поэтому стабильность на сроке 6 месяцев подтверждена.

Расчет для 12 месяцев:

$$\Delta_{-12}(\%) = (88 - 100)/(100) \times 100 = -12\%.$$

Изменение составляет -12 %, что превышает критерий

$\pm 10\%$. Следовательно, стабильность на сроке 12 месяцев не подтверждена.

Корректный вывод:

Стабильность контрольного материала при хранении 2–8 °C подтверждена до 6 месяцев включительно. На сроке 12 месяцев выявлено изменение -12% , превышающее установленный критерий приемлемости $\pm 10\%$. Заявление срока стабильности 12 месяцев по данным настоящего исследования не обосновано; требуется дополнительное исследование, изменение срока годности или пересмотр условий хранения.

Этот пример показывает, что срок стабильности нельзя заявлять по последней желаемой точке, если данные ее не подтверждают. Если стабильность подтверждена до 6 месяцев, а на 12 месяцах результат неприемлем, вывод должен быть ограничен

Ускоренная и реальная стабильность

Для подтверждения стабильности могут использоваться исследования в реальном времени и ускоренные исследования. Исследование в реальном времени показывает, как изделие или материал ведет себя при заявленных условиях хранения в течение фактического срока. Это наиболее прямой способ подтверждения срока годности. Ускоренные исследования используют более жесткие условия, например повышенную температуру, чтобы получить предварительную оценку стабильности быстрее.

Ускоренные исследования полезны на этапе разработки

или предварительного обоснования, но они не всегда могут полностью заменить данные реального времени. Биологические материалы, белки, нуклеиновые кислоты, антитела, ферменты и сложные реагентные системы могут деградировать не строго предсказуемо. Поэтому, если срок годности изначально обоснован ускоренными данными, в дальнейшем обычно требуется подтверждение в реальном времени. Важно различать два разных предмета ускоренных испытаний. Ускоренное старение по модели, аналогичной ASTM F1980 (Standard Guide for Accelerated Aging of Sterile Barrier Systems for Medical Devices), относится к целостности упаковки и стерильного барьера как физической системы и не является моделью потери аналитической активности реагента. Кинетика потери активности белковых, антительных и ферментных компонентов определяется иными механизмами (термическая денатурация, агрегация) и оценивается методами, изложенными в CLSI EP25; единого стандарта, аналогичного ASTM F1980, для моделирования потери активности сложных биологических реагентов не существует. В отчете нужно четко указать, какие данные являются реальными, какие — ускоренными, и какие ограничения имеет такой подход.

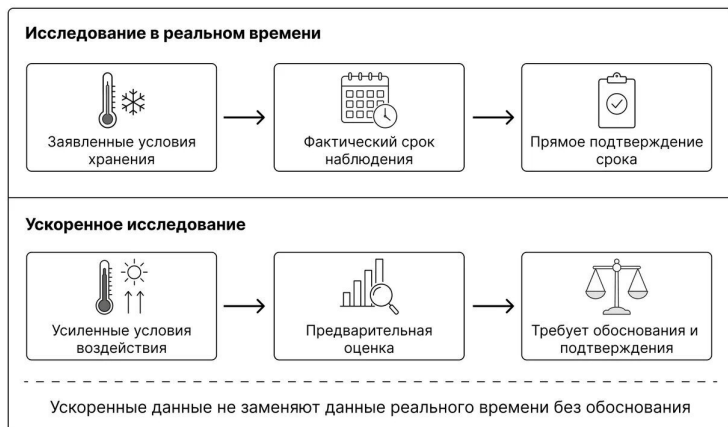


Рисунок 22.2. Реальная и ускоренная стабильность

Стабильность клинического образца и пробоподготовки

Стабильность клинического образца имеет самостоятельное значение. Даже если изделие и реагенты стабильны, анализ может разрушаться, связываться, изменять форму, подвергаться деградации или становиться недоступным для реакции в самом образце. Это особенно важно для РНК-мишеней, белков, клеточных маркеров, метаболитов, нестабильных гормонов, ферментов и материалов, чувствительных к температуре или времени.

Если инструкция допускает хранение образца до иссле-

дования, транспортирование, замораживание, размораживание или хранение экстракта, эти условия должны быть подтверждены. Нельзя указывать удобный для пользователя срок хранения без данных. Для качественных тестов необходимо проверять слабоположительные образцы, потому что именно они первыми могут перейти в отрицательную категорию при деградации аналита. Для количественных тестов важно оценивать смещение и изменение результата на клинически значимых уровнях.

Стабильность для качественных, количественных и молекулярных изделий

Для качественных и пороговых тестов стабильность проявляется через сохранение категории результата. Если слабоположительный образец после хранения становится отрицательным, стабильность недостаточна. Если пограничный образец после транспортирования переходит через порог, это может изменить медицинское решение. Поэтому для таких изделий критерий приемлемости может включать и изменение сигнала, и сохранение классификации результата.

Для количественных изделий стабильность обычно оценивают через изменение результата относительно исходного значения или установленного контрольного интервала. Однако процентное изменение само по себе не всегда достаточно. Нужно учитывать, влияет ли это изменение на клиническую интерпретацию. Если результат используется для мониторинга, даже умеренное смещение при хранении может

имитировать динамику состояния пациента.

В молекулярных изделиях стабильность связана с сохранностью нуклеиновой кислоты, эффективностью экстракции, ингибированием, условиями транспортировки, замораживанием и оттаиванием, а также хранением экстракта. Для РНК-мишеней стабильность особенно чувствительна к деградации. Для ДНК-мишеней могут быть важны фрагментация, качество материала и условия фиксации, особенно при работе с FFPE-образцами.

Сопоставимость результата при изменении серии, условий или версии

После подтверждения стабильности важно понимать, сохраняется ли аналитический вывод при изменениях, которые неизбежно возникают в жизненном цикле изделия. Это может быть новая серия реагента, новый калибратор, изменение условий транспортирования, использование открытого флакона, обновление программного обеспечения или изменение правил обработки данных. Такие изменения не всегда относятся к стабильности в узком смысле, но они влияют на сохранение аналитической пригодности результата.

Если новая серия реагента дает систематически иной результат, пользователь может ошибочно принять межсерийное различие за изменение состояния пациента. Если обновление алгоритма меняет классификацию, прежние данные об аналитической эффективности могут быть применимы только после оценки сопоставимости версий. Если открытый

флаконе после нескольких дней использования дает иной уровень сигнала, это должно быть учтено в сроке применения после вскрытия.

Поэтому в документации важно указать, какие изменения контролируются изготовителем и какие данные подтверждают сохранение результата. Для реагентных изделий это может быть межсерийная сопоставимость и стабильность открытого флакона. Для приборных систем — условия калибровки и технического обслуживания. Для программных изделий — управление версиями, оценка влияния обновлений и правила применения прежних данных после изменения алгоритма.

Стабильность программной версии

Для программных и алгоритмических изделий стабильность имеет особую форму. Если на вход подаются одни и те же данные при неизменной версии алгоритма, результат должен быть воспроизводимым в бытовом смысле этого слова, то есть повторная обработка тех же данных не должна произвольно менять итоговый вывод. Если результат меняется из-за обновления программной версии, изменения базы знаний, порогов, правил фильтрации или модели машинного обучения, необходимо оценить сопоставимость версий.

В таких изделиях стабильность результата зависит не от химической деградации, а от управления версиями, входными данными и правилами обработки. Поэтому в отчете нужно фиксировать, какая версия программного обеспечения

была использована при аналитической оценке. Если в обращении находится другая версия, необходимо показать, что изменение не влияет на клинически значимый результат или провести дополнительную оценку.

Что правильно и что неправильно при подтверждении стабильности

Схема 22.3 — Типовые ошибки при подтверждении стабильности и сохранения аналитической пригодности

— **Срок годности** — *Правильно: Подтвердить заявленный срок данными в реальном времени или обоснованной программой стабильности; Неправильно: Заявить срок без данных или только по краткому наблюдению.*

— **Открытый флакон** — *Правильно: Проверить стабильность после вскрытия при заявленных условиях применения; Неправильно: Считать срок закрытого изделия применимым после открытия.*

— **Транспортирование** — *Правильно: Смоделировать заявленные температурные и временные условия; Неправильно: Не оценивать влияние перевозки.*

— **Клинический образец** — *Правильно: Подтвердить условия хранения образца до исследования; Неправильно: Указывать срок хранения образца без данных.*

— **Замораживание/оттаивание** — *Правильно: Проверить заявленное число циклов; Неправильно: Допускать повторные циклы без подтверждения.*

— **Пороговые тесты** — *Правильно: Проверять слабоположительные и пограничные образцы; Неправильно: Использовать только явно положительные образцы.*

— **Критерии приемлемости** — *Правильно: Связать критерии с назначением и клинически значимой ошибкой; Неправильно: Оценивать только статистическое изменение без клинического смысла.*

— **Ускоренная стабильность** — *Правильно: Отделять предварительное ускоренное обоснование от данных реального времени; Неправильно: Представлять ускоренные данные как полную замену реальной стабильности без обоснования.*

— **Программная версия** — *Правильно: Фиксировать версию и оценивать влияние обновлений; Неправильно: Считать изменения ПО технически незначимыми без анализа.*

— **Вывод** — *Правильно: Ограничивать вывод проверенными условиями, сроками, материалами и версиями; Неправильно: Распространять вывод на непроверенные условия.*

Документирование в отчете

В отчете по стабильности необходимо указать, что именно оценивалось: реагент, набор, калибратор, контрольный материал, клинический образец, экстракт, открытый флакон, транспортные условия или программный результат. Нужно описать условия хранения, температуру, влажность при необходимости, временные точки, число серий, уровни анализа, тип материала, число повторов, критерии приемлемости.

сти и результаты по каждой точке.

Если стабильность подтверждается ускоренным исследованием, это должно быть указано отдельно. Необходимо описать модель ускорения, условия воздействия и ограничения вывода. Если данные реального времени еще продолжают накапливаться, это должно быть отражено в плане наблюдения. Нельзя представлять ускоренные данные как полностью эквивалентные реальному сроку без обоснования.

Если стабильность или сопоставимость результата подтверждены только для определенных условий, вывод должен быть ограниченным. Если стабильность подтверждена только при 2–8 °С, нельзя заявлять хранение при комнатной температуре. Если стабильность клинического образца подтверждена только для плазмы, нельзя автоматически переносить вывод на цельную кровь, сухое пятно крови или ликвор. Если версия алгоритма изменилась после испытаний, нужно оценить применимость прежних данных.

В инструкции по применению должны быть отражены подтвержденные условия хранения, транспортирования, срок годности, срок использования после вскрытия, допустимые условия обращения с образцом, правила замораживания и оттаивания, требования к экстракту, ограничения программной версии и действия пользователя при нарушении условий. Выводы отчета и инструкция не должны расходиться.

Итог главы

Стабильность показывает, сохраняет ли медицинское изделие для диагностики *in vitro*, его реагенты, калибраторы, контрольные материалы, клинические образцы, экстракты или программные результаты аналитическую пригодность во времени и при установленных условиях обращения. Она отличается от прецизионности: повторяемость, внутрилабораторная прецизионность и воспроизводимость относятся к условиям оценки прецизионности и рассмотрены ранее. В настоящей главе стабильность рассматривается как доказательство того, что ранее подтвержденные аналитические характеристики сохраняются в течение срока годности, после транспортирования, после вскрытия, при хранении образца, после замораживания и оттаивания, а также при управляемом изменении серии или версии изделия. Выводы о стабильности должны быть строго ограничены проверенными условиями, сроками, материалами, уровнями аналита и версиями изделия и должны быть согласованы с инструкцией по применению.

Глава 23. Прослеживаемость и калибровка

Прослеживаемость и калибровка являются ключевыми элементами аналитической эффективности медицинских изделий для диагностики *in vitro*, особенно для количественных измерений. Они позволяют понять, откуда берется числовое значение результата, с чем оно связано и насколько оно сопоставимо с результатами, полученными в другой серии, на другом приборе, в другой лаборатории или в другой системе измерений. Без прослеживаемости и корректной калибровки количественный результат может выглядеть точным, но быть слабо сопоставимым и плохо интерпретируемым.

Для диагностики *in vitro* это имеет особое значение. Многие результаты используются и как разовая лабораторная информация, и как основание для мониторинга, классификации пациента, оценки динамики, выбора терапии или сравнения с клиническими порогами. Если результат глюкозы, гормона, вирусной нагрузки, белкового маркера, метаболита или лекарственного вещества зависит от произвольной калибровки, меняется между сериями или не связан с признанной системой значений, его клиническая надежность снижается. Пользователь может получить число, но не пони-

мать, насколько оно сопоставимо с клиническим порогом или предыдущим результатом.

Метрологическая прослеживаемость — это свойство результата измерения, при котором результат может быть связан с признанной референтной основой через документированную непрерывную цепь калибровок или сравнений, каждое из которых вносит свой вклад в неопределенность.

Калибровка — это установление зависимости между значениями, приписанными калибраторам, и сигналом аналитической системы, позволяющее получать результат измерения для неизвестных образцов.

Прослеживаемость отвечает на вопрос, к какой системе значений относится результат, калибровка — как аналитическая система переводит сигнал в числовое значение, а неопределенность — насколько надежно это значение выражает измеряемую величину с учетом всех существенных источников вариабельности

Прослеживаемость не следует понимать как формальную ссылку на калибратор. Важно наличие калибратора, а также происхождение значения, которое ему приписано. Значение может быть связано с международным стандартом, референтным материалом, референтной методикой, стандартным образцом, внутренней системой изготовителя или консенсусным значением. Чем яснее и надежнее эта связь, тем

сильнее доказательная основа количественного результата. Если происхождение значения не раскрыто, результат может быть трудно сопоставить с результатами других систем.

Калибровка, в свою очередь, не является простой технической процедурой настройки прибора. Она определяет, как сигнал превращается в результат. Если калибровочная зависимость построена некорректно, если калибраторы не охватывают рабочий диапазон, если материал калибратора не ведет себя как клинический образец или если стабильность калибровки не подтверждена, результат пациента может быть смещен. Поэтому калибровка является частью доказательной базы аналитической эффективности.

Для большинства современных белковых и нуклеиновых биомаркеров признанная международная референтная система в настоящее время отсутствует — это типичная, а не исключительная ситуация. Изложенная далее полная цепь прослеживаемости — ориентир, реально достижимый для ограниченного круга давно стандартизированных аналитов; для остальных случаев задача состоит не в достижении полной цепи, а в честном описании системы значений изготовителя и её ограничений.

Метрологическая цепь прослеживаемости

В идеальном случае результат измерения связан с признанной референтной системой. Такая система может включать единицы СИ, первичный референтный материал, референтную методику, референтную лабораторию, вторичные

калибраторы, рабочие калибраторы изготовителя и калибраторы, используемые пользователем. В реальной практике для многих аналитов такая полная система отсутствует. Особенно это характерно для сложных биологических маркеров, антител, молекулярных мишеней, клеточных признаков и некоторых алгоритмических результатов. Но даже в этих случаях необходимо описать, к какой системе значений привязан результат и какие ограничения это создает.



Рисунок 23.1. Метрологическая цепь прослеживаемости результата

Рисунок 23.1 показывает, что результат пациента находится в конце цепи, а не существует сам по себе. Если верх-

ние звенья цепи надежны, документированы и применимы к данному аналиту, результат легче интерпретировать и сопоставлять. Если цепь разорвана или не описана, пользователь фактически получает результат, привязанный только к внутренней системе изготовителя. Иногда это допустимо, но должно быть прямо указано, особенно если отсутствует международная референтная система.

Для изделий, результаты которых используются в мониторинге, прослеживаемость особенно важна. Если пациент наблюдается во времени, изменение результата должно отражать изменение состояния, а не изменение калибровочной системы, партии реагента или пересмотра значения калибратора. Для изделий, применяемых в разных лабораториях, прослеживаемость помогает обеспечить сопоставимость результатов между площадками. Для калибраторов и контрольных материалов прослеживаемость является одной из центральных характеристик, потому что они сами используются для переноса значений.

Калибраторы, контрольные материалы, стандартные образцы и референтные материалы

В практике МИ ИВД часто смешивают понятия калибратора, контрольного материала, стандартного образца и референтного материала. Это создает методические ошибки. Калибратор используется для установления зависимости между сигналом и значением. Контрольный материал используется для контроля стабильности работы системы, но обыч-

но не предназначен для калибровки, если это специально не установлено. Стандартный образец или референтный материал имеет установленное значение и неопределенность, а его статус определяется процедурой аттестации и назначением. Термин «референтный материал» относится к международной терминологии; в российском тексте его целесообразно пояснять через понятие стандартного образца.

Схема 23.1 — Различие калибратора, контрольного материала, стандартного образца и референтного материала

— **Калибратор** — *Основное назначение:* Установление калибровочной зависимости между сигналом и значением аналита; *Что важно для доказательной оценки:* Источник приписанного значения, стабильность, диапазон, матрица, прослеживаемость.

— **Контрольный материал** — *Основное назначение:* Контроль стабильности и пригодности аналитической системы; *Что важно для доказательной оценки:* Установленные контрольные пределы, стабильность, матрица, применимость к методу.

— **Стандартный образец** — *Основное назначение:* Передача или подтверждение значения с установленными характеристиками; *Что важно для доказательной оценки:* Аттестованное значение, неопределенность, однородность, стабильность, назначение.

— **Референтный материал** — *Основное назначение:*

*Материал с достаточно однородными и стабильными свойствами для применения в измерительном процессе; **Что важно для доказательной оценки:** Коммутативность, пригодность к методу, значение и неопределенность при наличии.*

— Международный стандарт — Основное назначение: Гармонизация результатов и единиц для определенного аналита или биологической активности; **Что важно для доказательной оценки:** Связь с международной единицей, область применения, ограничения матрицы.

Контрольный материал нельзя автоматически использовать как калибратор, а калибратор нельзя рассматривать как полноценный референтный материал без подтверждения его статуса. Для каждого объекта нужно понимать, для какой цели он предназначен и какое значение ему приписано.

Калибровочная зависимость

Калибровка строится на серии калибраторов с известными значениями. Аналитическая система измеряет их сигналы, после чего устанавливается зависимость между значением аналита и сигналом. Затем по этой зависимости рассчитываются значения неизвестных образцов. В простейшем случае зависимость может быть линейной:

$$y = a + bx,$$

где: x — значение аналита; y — сигнал аналитической системы; a — свободный член; b — наклон калибровочной зависимости.

Однако в МИ ИВД многие методы имеют нелинейную калибровочную зависимость. Это особенно характерно для иммунологических анализов, некоторых ферментативных реакций, флуоресцентных сигналов, хемилюминесцентных методов и широкодиапазонных количественных систем. В таких случаях могут использоваться нелинейные модели, например логистические кривые. В учебном контексте важно не название модели, а принцип: выбранная модель должна адекватно описывать связь между сигналом и значением во всем заявленном диапазоне.

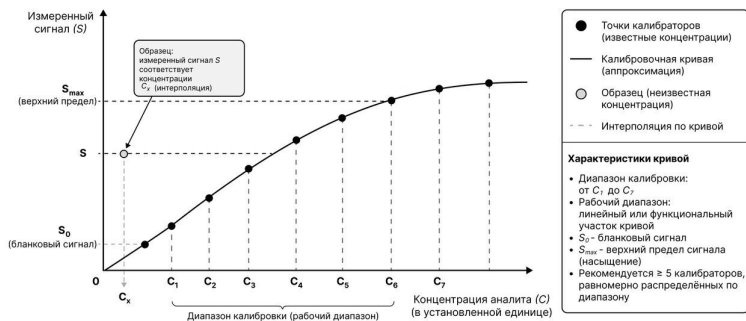


Рисунок 23.2. Калибровочная кривая

Калибраторы должны покрывать рабочий диапазон изде-

лия. Если клинически важные решения принимаются в нижней части диапазона, калибровочная система должна быть надежна именно там. Если изделие используется для высоких концентраций, верхняя часть диапазона также должна быть поддержана калибровочными точками или обоснованной моделью. Недопустимо строить калибровку только в средней части диапазона, а затем без подтверждения интерпретировать результаты на нижней или верхней границе.

Пример 23.1 — Расчет концентрации неизвестного образца по калибровочной зависимости

Предположим, что по результатам калибровки получена линейная зависимость:

$$y = 5 + 2x,$$

где y — сигнал прибора, x — концентрация аналита, Ед/мл. Для неизвестного образца получен сигнал:

$$y = 205.$$

Тогда концентрацию можно рассчитать:

$$205 = 5 + 2x,$$

$$2x = 200,$$

$$x = 100 \text{ Ед/мл.}$$

Следовательно, расчетное значение неизвестного образца составляет 100 Ед/мл.

Этот пример иллюстрирует только принцип пересчета сигнала в результат. В реальной аналитической системе необходимо учитывать модель калибровки, рабочий диапазон, неопределенность калибраторов, правильность, пре-

цизионность, матрицу, стабильность калибровки и критерии приемлемости. Если сигнал неизвестного образца выходит за пределы области, покрытой калибраторами, результат нельзя автоматически считать надежным.

Коммутативность материалов

Коммутативность означает, что калибратор, контрольный материал или стандартный образец ведет себя в аналитической системе подобно реальному клиническому образцу. Это особенно важно для МИ ИВД, потому что искусственно приготовленный материал может иметь то же номинальное значение аналита, но взаимодействовать с методом иначе. Если материал некоммутативен, он может давать правильную калибровку для себя, но смещать результаты клинических образцов.

Например, белковый аналит в сыворотке пациента может находиться в иной молекулярной форме, окружении или матрице, чем очищенный белок в буферном растворе. Нуклеиновая кислота в клиническом образце может отличаться от синтетического материала по способу экстракции, стабильности и взаимодействию с ингибиторами. Контрольный материал на искусственной матрице может быть удобным для контроля, но не полностью отражать поведение реального образца.

Поэтому при оценке прослеживаемости и калибровки важно понимать, какие материалы использованы и насколько они применимы к заявленной матрице. Некоммутатив-

ность не всегда делает материал непригодным, но ограничивает его доказательную роль. В отчете следует указать, для чего материал применялся: для калибровки, контроля, проверки линейности, оценки правильности или внутренней стабильности.

Неопределенность и перенос значения

Каждое звено цепи прослеживаемости вносит неопределенность (u). Значение референтного материала имеет неопределенность. Передача значения к рабочему калибратору также имеет неопределенность. Измерения калибраторов, стабильность калибровочной кривой, прецизионность прибора и матричные эффекты добавляют вклад в итоговую неопределенность результата. для настоящего издания важно не перегружать этот раздел расчетами неопределенности, но нужно показать сам принцип: числовой результат не является абсолютно точным, он имеет границы надежности.

В практической оценке МИ ИВД это означает, что изготовитель должен понимать и контролировать основные источники неопределенности, особенно если результат используется для клинических порогов или мониторинга. Если неопределенность результата велика относительно клинически значимого изменения, интерпретация динамики становится ненадежной. Если неопределенность около порога велика, возрастает риск неправильной классификации.

Калибровка и контроль качества

Калибровка устанавливает зависимость между сигналом

и значением, а контроль качества проверяет, сохраняется ли система пригодной после калибровки. Эти действия связаны, но не совпадают. Контрольный материал не заменяет калибровку, если система требует калибровки. Калибровка не заменяет контроль качества, потому что после калибровки система может изменяться во времени, между сериями, при замене реагентов или при изменении условий.

В доказательной документации важно показать, как часто требуется калибровка, когда она должна повторяться, какие материалы используются, как контролируется приемлемость калибровочной кривой и что происходит при несоответствии. Если изделие поставляется с заводской калибровкой, необходимо описать, как она устанавливается, как переносится к пользователю и какие условия должны выполняться для сохранения ее действительности.

Особенности разных типов МИ ИВД

Для биохимических и иммунохимических количественных тестов прослеживаемость и калибровка обычно имеют центральное значение. Результат выражается в концентрации или активности, а пользователь сопоставляет его с референсными интервалами, клиническими порогами или динамикой пациента. Здесь важны калибраторы, контрольные материалы, стабильность калибровки, коммутативность и прослеживаемость к признанной системе значений.

Для молекулярных количественных тестов, например вирусной нагрузки, прослеживаемость может быть связана с

международными стандартами или материалами, выраженными в международных единицах, копиях/мл или иных единицах. Важно ясно указывать, как пересчитывается результат, к какому материалу относится значение и как учитываются экстракция, объем образца и объем реакции. Если разные системы используют разные единицы или калибровочные основы, результаты могут быть несопоставимы.

Для качественных тестов калибровка может быть менее заметной, но это не значит, что она отсутствует. Многие качественные системы используют порог сигнала, внутренние контроли, калибровочные настройки или программные правила классификации. В таких случаях важно подтвердить, что порог установлен обоснованно, стабилен и связан с заявленным пределом обнаружения или клиническим правилом интерпретации.

Для алгоритмических изделий аналогом калибровки может быть настройка порога классификации, шкалы риска, модели вероятности или правила интерпретации. Если алгоритм выдает риск или категорию, нужно понимать, как этот результат калиброван относительно клинической вероятности или референтной разметки. Если модель систематически завышает или занижает риск, это аналогично смещению количественного метода.

Что правильно и что неправильно при прослеживаемости и калибровке

Схема 23.2 — Типовые ошибки при подтверждении прослеживаемости и калибровки

— **Описание прослеживаемости** — *Правильно:* Указать цепь передачи значения и источник приписанного значения; *Неправильно:* Написать «прослеживаемость обеспечена» без раскрытия цепи.

— **Использование калибратора** — *Правильно:* Обосновать значение, матрицу, стабильность и область применения; *Неправильно:* Использовать калибратор без описания происхождения значения.

— **Контрольный материал** — *Правильно:* Использовать для контроля стабильности системы; *Неправильно:* Подменять контрольным материалом калибровку без обоснования.

— **Диапазон калибровки** — *Правильно:* Покрыть клинически значимые уровни и рабочий диапазон; *Неправильно:* Калибровать только среднюю часть диапазона.

— **Коммутативность** — *Правильно:* Оценить применимость материала к клинической матрице; *Неправильно:* Считать любой искусственный материал эквивалентным клиническому образцу.

— **Заводская калибровка** — *Правильно:* Описать перенос калибровки и условия ее действительности; *Неправильно:* Не указывать, как пользователь контролирует пригодность системы.

— **Версия ПО** — *Правильно:* Учитывать влияние ал-

горитма и версии на расчет результата; **Неправильно:** Менять алгоритм пересчета без оценки сопоставимости.

— **Единицы результата** — **Правильно:** Указать системе единиц и способ пересчета; **Неправильно:** Использовать единицы без объяснения их связи с калибровочной основой.

— **Ограничения** — **Правильно:** Указать отсутствие полной референтной системы, если она отсутствует; **Неправильно:** Создавать видимость полной прослеживаемости без основания.

Документирование в отчете

В отчете по прослеживаемости и калибровке необходимо описать источник приписанных значений калибраторов, цепь прослеживаемости, использованные материалы, их матрицу, стабильность, неопределенность при наличии, область применения и ограничения. Если применяются международные стандарты, стандартные образцы или референтные материалы, нужно указать их идентификацию, назначение и способ передачи значения.

Для калибровки необходимо описать количество калибровочных уровней, их значения, модель калибровочной зависимости, рабочий диапазон, критерии приемлемости калибровки, частоту калибровки и действия при ее неприемлемости. Если зависимость нелинейная, следует указать используемую модель и показать, что она пригодна для заявленного диапазона. Если используются заводские калибро-

вочные коэффициенты или программная обработка, необходимо указать, как обеспечивается их контроль и защита от неконтролируемых изменений.

Конец ознакомительного фрагмента.

Текст предоставлен ООО «Литрес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на Литрес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.