

ГЛАВА 1.1.2.

ПРИНЦИПЫ И МЕТОДЫ ВАЛИДАЦИИ ДИАГНОСТИЧЕСКИХ ИССЛЕДОВАНИЙ НА НАЛИЧИЕ ИНФЕКЦИОННЫХ ЗАБОЛЕВАНИЙ

ВВЕДЕНИЕ¹

Валидация – это процесс определения пригодности анализа², разработанного, оптимизированного и стандартизированного надлежащим образом для предусмотренного назначения. Все диагностические анализы (лабораторные и в полевых условиях) должны быть валидированы для видов, в отношении которых они будут использоваться. Валидация включает оценку аналитических и диагностических характеристик теста. В контексте данной главы анализ, который прошел все три стадии пути валидации (см. Рисунок 1 ниже), включая характеристику рабочих свойств, может считаться «валидированным для первоначального предназначения»³. Для сохранения статуса валидированного анализа, однако, необходимо проводить тщательный мониторинг эффективности анализа в условиях рутинного использования, часто путем отслеживания поведения контролей анализа со временем. Это гарантирует, что анализ, валидированный первоначально, будет постоянно сохранять свои рабочие характеристики. В случае если данный анализ показывает результаты, не соответствующие первоначальным данным валидации, анализ может быть признан непригодным для своего предусмотренного назначения. Таким образом, валидированный анализ необходимо постоянно оценивать, чтобы обеспечить сохранение его пригодности для цели путем оценки результатов контроля анализа в каждом цикле и посредством оценки во время рутинного использования в целевой популяции.

Анализ, применяемый в отношении отдельных особей или популяций, имеет множество целей, например, помощь при: документальном подтверждении статуса благополучия по определенной болезни в стране или регионе, предотвращение распространения болезни во время торговли, искоренение инфекции в регионе или стране, подтверждение диагноза клинических случаев, оценка превалентности инфекции для содействия проведению анализа риска, выявление инфицированных животных для осуществления мер борьбы с болезнями и классификация животных для оценки состояния здоровья стада или иммунного статуса после вакцинации. Один анализ можно валидировать для одного или нескольких предусмотренных целей путем оптимизации его рабочих характеристик для каждой цели, например, установления высокой диагностической чувствительности (D_{Se}) наряду с более низкой диагностической специфичностью (D_{Sp}) для скринингового анализа или наоборот – установление высокой диагностической специфичности с более низкой диагностической чувствительностью для подтверждающего анализа.

В результате постоянного появления новых уникальных диагностических реагентов, совмещенных с множеством новых аналитических платформ и протоколов, возникают постоянные дискуссии о том, как правильно валидировать эти анализы. Уже недостаточно предложить простые примеры из серологических анализов, например, иммуноферментный анализ, чтобы разработчики анализов могли понять, как валидировать более сложные

¹ Версия, принята Всемирной Ассамблеей Делегатов МЭБ в мае 2013 г.

² «Анализ», «метод исследования» и «тест» это синонимичные термины в контексте данной главы и, соответственно, заменяют друг друга.

³ Доступно на:

http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/GUIDELINE_3.6.0_INTRODUCTION.pdf

тесты, такие как тесты на выявление нуклеиновых кислот. Чтобы внести согласованность в процесс валидации всех видов анализов, в этой главе будут описаны критерии, которые необходимо выполнять во время разработки анализа и валидации всех видов анализа. Включение разработки анализа в процесс его валидации, возможно, противоречит логике, но в реальности, три критерия валидации (определение предусмотренного назначения, оптимизация, стандартизация), которые нужно оценивать для получения валидированного анализа, включают этапы процесса его разработки. Таким же образом, процесс разработки анализа органично переходит к процессу его валидации, при этом оба этих процесса включают критерии валидации, обязательные для соблюдения. Кроме того, более подробные рекомендации представлены в Руководстве МЭБ по валидации⁴, специально разработанном для нескольких фундаментально разных типов анализа (например, обнаружение нуклеиновых кислот, антител или антигенов) и обеспечивает большей информацией по особым вопросам, относящимся к валидации диагностических анализов. Для получения специальной информации по диким видам, смотрите Руководство МЭБ по валидации 3.6.7⁵ Принципы и методы валидации диагностических тестов на инфекционные болезни, применимых к диким животным. Информация, представленная в Руководстве МЭБ по валидации 3.6.7, по диким животным может также быть полезной в отношении валидации тестов, проводимых на домашних животных, например, в тех случаях, когда количество и доступность проб являются ограниченными.

ПРЕДВАРИТЕЛЬНЫЕ СООБРАЖЕНИЯ ПРИ РАЗРАБОТКЕ И ВАЛИДАЦИИ АНАЛИЗОВ

Все лаборатории должны соответствовать требованиям Главы 1.1.1 (*Руководство по водным животным*) или 1.1.5 (*Руководство по наземным животным*) по Управлению качеством в ветеринарных исследовательских лабораториях. Это минимизирует влияние факторов, которые не зависят от самого теста, таких как оборудование, ошибка оператора, выбор реактива (как химического, так и биологического) и калибровка, сосуды и платформы для проведения реакции, качество воды, рН и степень ионизации буферных растворов и разбавителей, температура инкубации и ее длительность, ошибки в технических характеристиках анализа.

Первый этап в разработке анализа – определить цель анализа, поскольку она будет направлять все последующие этапы процесса валидации. Критерии валидации анализа являются характерными чертами анализа, представляющими собой решающие факторы, параметры или стандарты, на основании которых делаются выводы или принимается решение. После рассмотрения переменных, которые влияют на характеристики анализа, критерии, которые необходимо оценить в процессе валидации, становятся более понятными. Переменные можно разделить на категории: (а) образец – индивидуальный или объединенный, состав матрицы и взаимодействия хозяин/организм, количественным или качественным образом влияющие на целевой аналит; (б) система анализа – физические, химические, биологические или человеческие факторы, влияющие на способность анализа обнаруживать специфичный аналит в образце; и (с) интерпретация результатов теста – способность результата теста, полученного с помощью системы анализа, точно предсказывать статус отдельной особи или популяции в отношении цели, для которой он применяется.

Отбор, сбор, подготовка, консервирование и контроль образцов являются критическими

⁵ : http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/3.6.07_WILDLIFE.pdf

переменными при планировании и разработке анализа с целью обеспечения обоснованных результатов исследования. Другие переменные, такие как транспортировка, передача и хранение, прослеживание образцов и система управления информацией в лаборатории являются основными источниками отклонений/ошибок, что является особенно важным, когда анализ применяется для проведения рутинного исследования. Целостность результатов эксперимента во время разработки и валидации анализа зависит от качества используемых образцов. Предвидение факторов, которые могут отрицательно сказываться на качестве образца, должно предшествовать деятельности по валидации анализа. Референтные образцы, используемые в разработке и валидации анализа, должны быть той же матрицы, что и матрица, используемая в анализе (например, сыворотка, ткани, цельная кровь), и представлять виды животных, исследуемые с использованием анализа. Референтные материалы должны соответственно представлять диапазон концентрации аналита, который должен быть обнаружен посредством анализа. Информация о сборе, подготовке, консервации, управлении и транспортировке доступна в главах 1.1.2 и 1.1.3 *Руководства по наземным животным*.

Матрица, в которой может находиться целевой аналит (сыворотка, фекалии, ткань и т.д.) может содержать эндогенные или экзогенные ингибиторы, которые могут мешать работе некоторых анализов. Это особенно касается таких фермент-зависимых тестов, как полимеразная цепная реакция (ПЦР) или иммуноферментный анализ (ИФА). Другие факторы, которые влияют на концентрацию и состав целевого аналита (особенно антитела) в образце, могут быть преимущественно связаны с хозяином и являются либо присутствующими изначально (например, возраст, пол, порода, упитанность, беременность, иммунологическая реактивность) или приобретенными (например, пассивно приобретенные антитела, активный иммунитет в результате вакцинации или инфекции). Факторы, не связанные с хозяином, такие как контаминация или порча образца, также потенциально влияют на способность анализа обнаруживать специфичный целевой аналит в образце. Также важно, чтобы биологические реактивы были свободны от чужеродных агентов, иначе это может привести к ошибочным результатам.

КРИТЕРИИ РАЗРАБОТКИ И ВАЛИДАЦИИ АНАЛИЗА

На рабочие характеристики анализа влияют множество факторов, начиная с оптимизации анализа. После первоначальной оптимизации для предусмотренного назначения проводят тестирование рабочих характеристик. На основании результатов валидации для анализа может потребоваться дополнительная оптимизация анализа или же он может быть признан, как соответствующий целевому назначению.

Критерии валидации и разработки анализа

- i) Определение предусмотренного назначения
- ii) Оптимизация
- iii) Стандартизация
- iv) Повторяемость
- v) Аналитическая чувствительность
- vi) Аналитическая специфичность
- vii) Пороговые значения (точки раздела)
- viii) Диагностическая чувствительность
- ix) Диагностическая специфичность
- x) Воспроизводимость
- xi) Соответствие предусмотренному назначению

1. ПУТЬ РАЗРАБОТКИ АНАЛИЗА

1.1. Определение предусмотренного назначения анализа

В *Стандарте МЭБ по управлению и техническим требованиям для лабораторий, осуществляющих исследования на инфекционные болезни* (Всемирная организация по охране здоровья животных, 2008)⁶ указано, что методы тестирования и связанные с этим процедуры должны соответствовать определенным сферам применения для диагностики, чтобы результаты тестов имели значение. Другими словами, анализ должен «быть пригоден для цели». Качественная или количественная оценка способности положительного или отрицательного результата теста точно прогнозировать статус инфицирования или заражения животного или популяции животных является окончательным фактором, который учитывают при валидации анализа. Эта способность зависит от разработки тщательно оптимизированного и стандартизированного (Раздел А.2.3) анализа, который путем накопления данных валидации обеспечивает уверенность в том, что анализ можно провести в соответствии с предусмотренной целью. Чтобы гарантировать, что результаты тестов представляют полезные диагностические выводы о животных и популяциях в отношении предусмотренной цели, процесс валидации должен распространяться на документацию по первоначальной разработке и рабочим характеристикам анализа, а также на непрерывную оценку программ контроля и обеспечения качества. На Рисунке 1 показан процесс валидации анализа с момента проектирования анализа посредством разработки и валидации до применения, внедрения и материально-технического обеспечения анализа.

Первым этапом в разработке анализа является выбор типа анализа, который соответствует и который, по-видимому, можно валидировать для определенной цели (соответствие цели).

Самыми распространенными задачами анализа являются:

- 1) Дополнительные данные для демонстрации отсутствия инфекции в определенной популяции (стране/зоне/компарimente/стаде) (превалентность очевидно нулевая):
 - a) «Благополучие» с применением и/или без применения вакцинации,
 - b) Восстановление благополучия по инфекции после вспышек,
- 2) Подтверждение отсутствия инфекции или наличия возбудителя у отдельных животных или продуктов для целей торговли/перемещения.
- 3) Содействие искоренению болезни или инфекции из определенных популяций.
- 4) Подтверждение диагноза случаев с подозрением или с клиническими признаками (включает подтверждение положительного скринингового теста).
- 5) Оценка превалентности инфекции или заражения для анализа риска (наблюдения, статус здоровья стада, меры контроля болезни).
- 6) Определение иммунного статуса отдельных животных или популяций (пост-вакцинальный).

⁶ Это специфическая интерпретация наиболее общих требований ISO/IEC 17025:2005 международного стандарта качества для исследовательских лабораторий (2005). В данной публикации указано, что для того, чтобы метод считался подходящим, он должен быть должным образом валидирован и данная валидация должна соответствовать принципам, указанным в данном документе, Стандарт МЭБ по валидации.

Эти цели включают более узкие и более специфичные области применения анализов (см. Рекомендации Руководства МЭБ по валидации диагностических тестов: сноска 3) для каждого типа анализа). Такие особые области применения и их уникальные цели нужно четко определять в контексте полностью валидированного анализа.

Для предусмотренного назначения анализ должен быть определен относительно видов целевых животных, целевых патогенов или условий и матрицы пробоотбора.

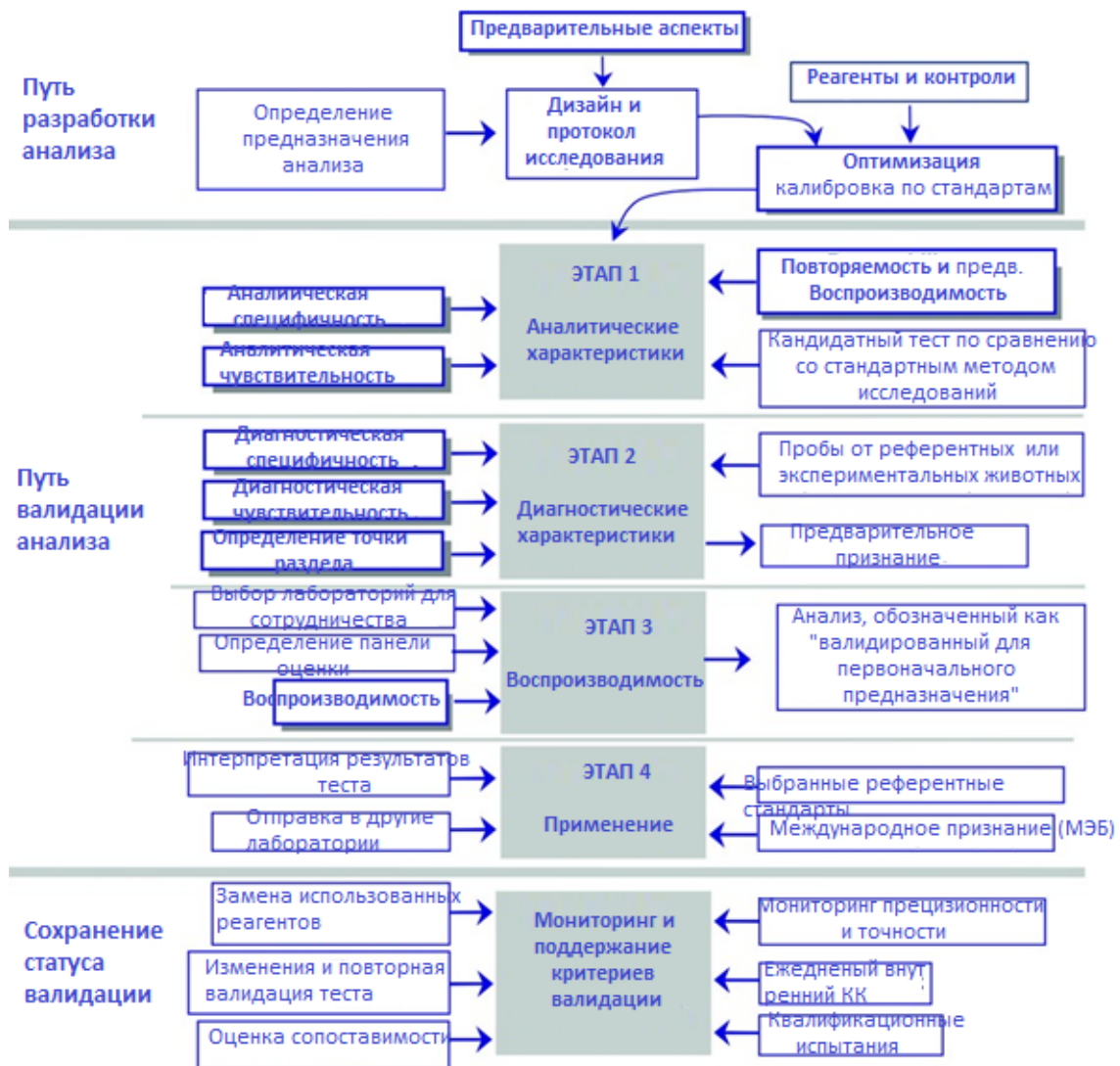


Рисунок 1. Пути разработки и валидации анализа. Критерии валидации анализа выделены жирным шрифтом в затененных прямоугольниках.

1. 2. Разработка анализа – экспериментальные исследования

1.2.1. Дизайн метода тестирования и доказательство правильности концепции

Первоначальные знания, соображения и планирование должны быть направлены на проектирование всех этапов нового анализа, предназначенного для валидации, или

модификации существующего анализа. Помощь предлагается в Руководстве по валидации диагностических тестов (смотри сноску 3), в котором описаны наилучшие практики разработки и валидации анализов для обнаружения различных аналитов (например, обнаружение антител, антигенов и нуклеиновых кислот).

Разработка любого анализа зависит от референтных образцов, которые отражают целевой аналит и матрицу, в которой данный аналит обнаруживают и популяцию, для которой данный анализ предназначен. Референтными образцами могут быть сыворотки, жидкости (включая мясные соки) или ткани которые содержат указанный аналит или геномную конструкцию, совпадающую с целевым аналитом. Эти референтные материалы используются в эксперименте на протяжении всего процесса разработки и переносятся на этап валидации анализа.

1.2.2. Рабочий диапазон анализа

Рабочий диапазон анализа - интервал концентраций аналита или титров, в котором метод обеспечивает нужную точность и прецизионность. Точность это близость исследуемой величины к предполагаемой (достоверной) величине (средней или срединной) для референтного стандарта реактива известной концентрации или титра. Прецизионность⁷ - это степень дисперсии (вариация, стандартное отклонение или коэффициент вариаций) в рамках серии измерений того же самого тестируемого образца при указанных условиях. Во время разработки анализа устанавливают нижние и верхние пределы рабочего диапазона. Для формального определения данного диапазона выбирают сильно положительный референтный образец. (В идеале этот образец должен быть в то же время одним из трех образцов, описанных в разделе А.2.3 ниже). Данный сильно положительный образец серийно разводят до исчезновения ответа анализа в аналит - отрицательной матрице того же состава, что и матрица проб от животных в популяции, для которой предназначен анализ. Результаты представляют в виде кривой реакции, при этом ответ (например, оптическая плотность, пороговое значение цикла и т.д.) является функцией концентрации (количества) аналита. Кривая определяет рабочий диапазон анализа. Если обнаружено, что диапазон является неприемлимым для предусмотренного назначения, то может понадобиться дополнительная оптимизация. Типичная калибровочная кривая для большинства анализов имеет сигмоидальную форму. Данные трансформируются в аппроксимируемую линейную зависимость между ответом и концентрацией с использованием подходящего алгоритма (Findlay & Dillard, 2007).

1.2.3. Стандартизация и оптимизация

Оптимизация – это процесс, с использованием которого оценивают наиболее важные физические, химические и биологические параметры анализа и корректируют их для обеспечения того, чтобы рабочие характеристики анализа наилучшим образом подходили для предусмотренного применения. Полезно выбрать минимум три хорошо определенных референтных образца, представляющих аналит в диапазоне от сильно положительного до отрицательного (например, отрицательный, слабо и сильно положительный). Эти образцы в идеале должны представлять животных с известным заражением и незараженных животных в популяции, которые в результате будут целью анализа. Однако не всегда возможно получить такие референтные образцы, особенно для анализов на обнаружение

⁷ Лабораторные источники вариации, которые влияют на прецизионность анализа, включают: в течение одного цикла теста, 2) между циклами, 3а) между циклами анализа в разное время в тот же день или в разные дни при тех же условиях, 3б) между циклами анализа в разные дни с учетом разных операторов, 4 между лабораториями. В данной главе категории 1-3 являются оценками повторяемости и категория 4 синонимична воспроизводимости.

нуклеиновых кислот и антигенов. Альтернативой является подготовка референтных образцов с добавлением возбудителей, выращенных методом культивирования, или положительных сывороток, но такой способ хуже, так как эти образцы недостоверно представляют естественное взаимодействие матрицы и возбудителя (см. Руководство МЭБ по валидации 3.6.6⁸ *Отбор и использование референтных образцов и панелей*). Но если других вариантов нет, добавление к пробе известного количества аналита или агента, полученного в результате культивирования, или разведение сильно положительной сыворотки в отрицательной сыворотке того же вида может быть единственной альтернативой. В любом случае, обязательно, чтобы матрица, в которую вносят аналит или в которой его разводят, была идентична или как можно больше походила на пробы, для которых, в конечном счете, будет предназначен анализ. В идеале, референтные образцы должны быть хорошо охарактеризованы одним или лучше как минимум двумя альтернативными методами. Эти образцы можно использовать в экспериментах для определения того, может ли анализ дифференцировать разное количество аналита, различать целевой и близкородственные аналиты и для оптимизации концентрации реактивов и усовершенствования протокола. В принципе для всех типов анализа желательно подготовить и хранить достаточное количество каждого референтного образца в аликвотах для использования в каждом цикле анализа-кандидата, так как его оценивают на протяжении всего процесса разработки и валидации. При смене референтных образцов во время процесса валидации возникает неустранимая переменная, которая может сильно осложнить интерпретацию экспериментальных данных и, следовательно, подорвать надежность процесса разработки и валидации.

Трудоемкий процесс оптимизации анализа является определяющим и важным этапом для достижения надежного и предсказуемого проведения анализа. Научные суждения и использование наилучших научных практик, представленных в Разделе 3.6 *Наземного кодекса* (сноска 3), рекомендуются для руководства процессом оптимизации всех элементов разработки и валидации анализа. Такой подход является надежным основанием для разработки надежного анализа. Часто прототипные анализы разрабатывают с использованием реактивов и оборудования, имеющихся в лаборатории. Однако если анализ предназначен для проведения рутинных лабораторных тестов в нескольких лабораториях, стандартизация становится особенно важной. Состав каждого химического реактива и буферного раствора должен быть полностью описан. Все реактивы должны быть определены в отношении чистоты и классификации (включая воду). Для таких параметров, как pH, молярность и т.д., должны быть установлены и задокументированы приемлемые рабочие диапазоны. Для биологических препаратов необходимо определить стандарты для качества, чистоты, концентрации и реактивности. Также необходимо учесть сроки годности и условия хранения химических и биологических препаратов. Обязательно нужно установить приемлемые диапазоны для времени реакции и температуры. Оборудование, необходимое для эффективного проведения анализа, должно быть подробно описано, включая рабочие спецификации и калибровку. Технологический контроль (контроль качества) должен стать неотъемлемой частью процесса оптимизации с самого начала, а не при завершении разработки анализа, как обычно происходит. Кроме вышеуказанных аспектов, стандартизация и оптимизация также могут потребоваться для последующих действий, таких как сбор, обработка и интерпретация данных. И наконец, после оптимизации все эти параметры должны быть полностью описаны в протоколе метода тестирования.

Во время осуществления оптимизации анализа важно делать записи об этапах процедуры и

⁸ Доступно на: http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/3.6.06_REFERENCE_SAMPLES.pdf

параметрах с узким диапазоном, в котором они оптимально работают, так как это и есть критические точки, которые могут повлиять на надежность анализа (см. Раздел 1.2.7). В некоторых типах анализа определенные этапы процедуры могут в большей степени влиять на окончательное проведение анализа, чем другие этапы (См. Раздел 2.5 ниже и Главу 3.6.8 *Сопоставимость анализов после минимальных изменений в валидированном методе тестирования Руководства по наземным животным*⁹ для получения дополнительной информации по определению сопоставимости при смене реактивов или процессов).

В следующих разделах описаны несколько референтных образцов аналита и другие методы контроля процесса, которые обычно используются в тест-системах. Они обеспечивают важные функции мониторинга анализа, которые требуют особого внимания во время оптимизации анализа. Кроме этого, необходимо обеспечить правильное приготовление и хранение всех биологических реактивов и референтных материалов для поддержания их стабильности (см. Главу 1.1.2 *Руководства по наземным животным*).

1.2.4. Ингибирующие факторы в матрице пробы

Каждая матрица, которая должна использоваться в анализе должна применяться в процессе валидации. Некоторые матрицы пробы включают ингибирующие факторы, которые мешают выполнению некоторых типов анализов. Сыворотка, особенно гомогенизированная, может содержать факторы, токсичные к клеткам, используемым в реакции вирус нейтрализации, в то время как эндогенные вещества, обнаруженные в некоторых тканях и жидкостях, могут мешать или ингибировать анализы связывания с лигандом или анализы, основанные на ферментативной реакции, такие как ИФА. Пробы фекалий, автолизированных тканей и спермы содержат больше интерферирующих веществ и, таким образом, являются более проблематичными для проведения анализа, чем сыворотка, кровь или свежие ткани.

1.2.5. Устойчивость

Устойчивость относится к способности анализа оставаться неизменным под влиянием небольших вариаций в ситуации тестирования, которые могут происходить во время тестирования в одной лаборатории. Оценка устойчивости должна начинаться во время разработки анализа и этапов оптимизации. Намеренное изменение параметров метода может производиться в экспериментах после определения оптимальных условий анализа. Однако если для оптимизации анализа используется титрация реактивов по нескольким факторам, могут появиться признаки нарушения устойчивости. Если небольшие различия в условиях или концентрации реактивов приводят к неприемлемым вариациям, вероятно, анализ не будет устойчивым. Своевременное выявление таких ситуаций становится важным моментом для принятия решения о том, целесообразно ли продолжать валидацию анализа, поскольку если анализ не будет устойчивым в одной лаборатории при сравнительно идеальных условиях, он вряд ли будет обладать должной воспроизводимостью при передаче в другие лаборатории.

К факторам, которые с наибольшей вероятностью будут влиять на устойчивость анализа, относятся pH, температура, партия реактивов, торговая марка титрационных микропланшетов и факторы водных и органических матриц (Dejaegher & Vander Heyden, 2006). Когда оптимизация завершена, устойчивость анализа становится частью оценки

9

http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/3.6.08_COMPARABILITY_ASSAYS_AFTER_CHAN_GES.pdf

повторяемости.

1.2.6. Калибровка анализа по стандартным реактивам

1.2.6.1 Международные и национальные эталонные стандарты

В идеале, эталонные стандарты МЭБ и другие международные стандарты, содержащие известную концентрацию или титр аналита, это реактивы, по которым стандартизируют все анализы (смотри Руководство МЭБ 3¹⁰а также главу 3.6.6.). Такие стандарты готовят и распространяют референтные лаборатории МЭБ или другие международные референтные лаборатории. Национальные эталонные образцы калибруют путем сравнения с международным эталонным стандартом, если возможно; их готовит и распространяет национальная референтная лаборатория. В случае если международный эталонный образец отсутствует, национальный эталонный образец становится образцом сравнения для анализа-кандидата. Такие стандарты тщательно характеризуют с применением комплексного анализа, и, желательно, чтобы методы для их характеристики, подготовки и хранения были опубликованы в рецензируемых журналах.

1.2.6.2. Внутрилабораторный стандарт

Внутрилабораторный референтный стандарт обычно калибруют относительно международного или национального стандарта. В отсутствие одного из этих калибровочных стандартов и по возможности внутренний стандарт характеризуют тем же способом, что и международные и национальные стандарты (смотри главу 3.6.6.). Местный внутрилабораторный стандарт, следовательно, становится наилучшим из имеющихся стандартов и должен храниться в достаточном количестве в виде аликвот для периодического использования в качестве стандарта, по которому калибруются рабочие стандарты.

1.2.6.3. Рабочий стандарт

Один или более рабочих стандартов, обычно известных как контроли аналитов или контроли процесса, калибруют по международному, национальному или внутреннему стандарту и готовят в больших количествах, делят на аликвоты и хранят для рутинного использования при каждом проведении диагностического анализа.

1.2.7. «Нормализация» результатов тестов по рабочему стандарту

Из-за внутренних вариаций в необработанных результатах тестов, которые часто наблюдаются между разными проведениями одного и того же анализа или между лабораториями при использовании одного и того же анализа и похожих анализов, практически невозможно напрямую сравнить (полу-)количественные данные. Чтобы заметно улучшить сопоставимость результатов тестов как внутри одной лаборатории, так и между лабораториями, в каждый цикл анализа включают один или более рабочих стандартов. Необработанные тестовые значения по каждой исследуемой пробе затем можно будет конвертировать в единицы активности относительно рабочего стандарта(ов) с помощью процесса, который называется «нормализацией». «Нормализованные»

¹⁰ Доступно на:

http://www.oie.int/fileadmin/Home/eng/Our_scientific_expertise/docs/pdf/GUIDELINE_3_REF_STANDARDS_ANG.pdf

значения можно выражать разными способами, например, в виде процента положительного контроля (например, в ИФА) или в виде оцененной концентрации или титра аналита, полученных по стандартной кривой. Это является хорошей практикой включать рабочие стандарты во все циклы анализа во время разработки и валидации анализа, так как это позволяет «нормализовать» данные для получения валидного средства непосредственного сравнения результатов между циклами анализа. Обязательно осуществлять контроль (абсолютной) вариации нормализованных стандартов, так как иначе нормализация может быть причиной отклонений.

Дополнительную информацию см. в главах 3.6.1¹¹ *Разработка и оптимизация анализа по обнаружению антител*, 3.6.2¹² *Разработка и оптимизация анализа по обнаружению антигенов* и 3.6.3¹³ *Разработка и оптимизация анализов по обнаружению нуклеиновой кислоты*.

1.2.8. Предварительное изучение повторяемости

Оценка повторяемости должна начинаться во время этапов разработки и оптимизации. Своевременное выявление таких ситуаций становится важным моментом для принятия решения о том, целесообразно ли продолжать валидацию анализа.

Повторяемость далее верифицируют на этапе 1 валидации анализа (Раздел 2.1.1.). Когда оптимизированный тест проводится в рутинных лабораторных или полевых условиях (Этап 4 валидации анализа), то повторяемость постоянно контролируют как часть процедуры контроля в течение всего срока использования анализа (смотри Раздел 2.5.1).

2. ПУТЬ ВАЛИДАЦИИ АНАЛИЗА

Валидация – это процесс определения пригодности анализа, разработанного, оптимизированного и стандартизированного надлежащим образом для предусмотренного назначения. Валидация включает оценку аналитических и диагностических характеристик метода. В контексте данного документа, анализ который завершил все три стадии пути валидации (Рисунок 1), включая характеристику рабочих свойств, может считаться «валидированным для первоначального предназначения».

2.1. Этап 1 – Аналитические характеристики

В идеале, дизайн исследований, указанных в следующих разделах, должен состояться с помощью статистика и эксперта по болезням, чтобы размер образца и экспериментальный подход были валидными. Есть возможность спланировать эксперименты, которые эффективно предоставляют информацию о вероятных внутри- и межлабораторных источниках вариации в прецизионности анализа (см. сноску 4 в Разделе А.2.2, выше), в которой дано определение рабочих характеристик анализа. Выбор организмов, штаммов или серотипов для оценки аналитической специфичности должен отражать имеющиеся знания и, следовательно, сообщать наилучший из возможных экспериментальных дизайнов для нацеливания на определенные аналиты.

2.1.1. Повторяемость

¹¹ Доступно на: http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/3.6.01_ANTIBODY_DETECT.pdf

¹² Доступно на: http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/3.6.02_ANTIGEN_DETECT.pdf

¹³ Доступно на: http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/3.6.03_NAD_ASSAYS.pdf

Повторяемость – это уровень согласованности между результатами дублирующих проб как внутри, так и между циклами одного метода тестирования в данной лаборатории.

Повторяемость определяют путем оценки вариаций по результатам дублирующих проб. Количество дублирующих проб предпочтительно определять вместе со статистиком. Рекомендуется как минимум три дублирующие пробы, представляющие активность аналита в рамках рабочего диапазона анализа. Аликвоты каждого из этих образцов помещают в соответствующее количество отдельных сосудов в виде трех одинаковых копий первоначальной пробы, содержащей оригинальный аналит и концентрацию матрицы (См. Руководство МЭБ по валидации 3.6.6 (сноска 8). Затем каждую копию прогоняют по всем этапам анализа, включая приготовление рабочего разведения, как если бы это была исследуемая проба, полученная из популяции, для которой предназначен анализ. Нельзя готовить конечное рабочее разведение пробы в одной пробирке, из которой разведенные аликвоты помещают с помощью пипетки в реакционные сосуды или делать дублирующие пробы из одной экстракции нуклеиновой кислоты, вместо экстракции каждой дублирующей пробы до разведения в реакционных сосудах. Такие «пробы» не считаются валидными дубликатами для изучения повторяемости. Вариацию между циклами определяют с использованием одних и тех же проб в нескольких циклах с участием двух или более операторов в разные дни. Вариации по результатам дублирующих проб могут выражаться в виде стандартных отклонений, коэффициентов вариантов (стандартное отклонение+среднее значение дублирующих проб) или других возможных вариантов (См. Главу 3.6.4¹⁴ *Руководство по наземным животным Неопределенность измерений для оценки повторяемости*).

2.1.2. Аналитическая специфичность

Аналитическая специфичность (ASp) – это способность анализа дифференцировать целевой аналит (например, антитело, организм или геномную последовательность) от нецелевых аналитов, включая компоненты матрицы. Оценка является качественной и выбор и источники типов, организмов и последовательностей для оценки ASp должен отражать цель теста и тип анализа. В главах 3.6.1., 3.6.2. и 3.6.3 *Руководства по наземным животным* (сноски 11, 12 и 13) можно найти информацию по анализам на обнаружение антитела, антигена и нуклеиновой кислоты соответственно. ASp документируют во время Этапа 1 валидации и идентифицируют перекрестные реакции. Допустима кросс-реактивность (ASp менее 100%) в зависимости от целевого назначения анализа. Влияние кросс-реактивности далее документируют на этапе 2 (Обнаружение диагностической специфичности) и оценивается во время выполнения Этапа 4.

2.1.2.1. Селективность

Селективность относится к степени, в которой метод может точно определить количество целевого аналита в присутствии: 1) интерферирующих веществ, таких как компоненты матрицы (например, ингибиторов ферментов в реакционной смеси); 2) дегрантов (например, токсических факторов); 3) неспецифического связывания реагирующих веществ с твердой фазой (например, конъюгата ИФА, абсорбированного к лунке титрационного микропланшета); 4) антител к вакцинации, которые можно перепутать с антителами к активному заражению. Такие интерферирующие вещества могут вызывать ложно пониженный или ложно повышенный ответ в анализе, который отрицательно влияет на его аналитическую специфичность. Vessman *et al.* (2001) представили полезный обзор по селективности для аналитической химии, из которого взято изменение, описанное в данном разделе, для применения в отношении диагностических тестов.

¹⁴ Доступно на:

http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/3.6.04_MEASUREMENT_UNCERT.pdf

2.1.2.2. Эксклюзивность

Эксклюзивность – это способность анализа обнаруживать аналит или геномную последовательность, уникальную для данного целевого организма, и исключать все другие известные организмы, которые могут перекрестно реагировать. Это определение также относится к подтверждающему анализу.

2.1.2.3. Инклюзивность

Инклюзивность – это способность анализа обнаруживать несколько штаммов или сероваров вида, нескольких видов рода или похожей группы близкородственных организмов или антител к ним. Инклюзивность характеризует область действия для скринингового анализа.

2.1.3. Аналитическая чувствительность

Предел обнаружения (ПО) – это измерение аналитической чувствительности (ASe) анализа. ПО является оцененным количеством аналита в указанной матрице, которое будет давать положительный результат, по крайней мере, в течение указанного процента времени. Обычно оцененный ПО будет основан на обогащении целевой матрицы аналитом. Выбор аналита (например, видов, штаммов) является частью определения ASe. Он должен быть зарегистрирован должным образом. Данные эксперименты могут быть разработаны для прецизионной и точной оценке вероятности (например, 50% или 100%), но при некоторых обстоятельствах приемлема консервативная оценка ПО (например, 100%). Например, при титрации с использованием 10-кратных разведений все копии во всех разведениях могут показывать как 100%, так и 0% ответ. На данном этапе выбор состоит из двух опций. Последующее разведение, показывающее 100% ответ, может быть принято в качестве консервативной оценки нижнего предела обнаружения. Более точная оценка может быть получена на втором этапе эксперимента с использованием более узких интервалов в схеме разведения с фокусом на области от 100% до 0%. Методы статистической оценки данных ПО находятся в Главе Руководства по наземным животным 3.6.5¹⁵ *Статистические подходы к валидации*.

2.1.4. Аналитическая точность вспомогательных тестов или процедур

Некоторые методы или процедуры тестирования можно квалифицировать для использования в качестве аналитических инструментов в диагностической лаборатории. Это обычно вторичные вспомогательные тесты или процедуры, которые применяются к аналиту, обнаруженному в первичном анализе. Цель таких аналитических инструментов – дополнительная характеристика аналита, обнаруженного в первичном анализе. Примеры таких вспомогательных тестов включают нейтрализацию вируса для типирования выделенного вируса и молекулярное секвенирование.

Такие вспомогательные тесты нужно валидировать для определения аналитических характеристик (Разделы А.2 – В.1.3 выше). Однако они отличаются от диагностических тестов, потому что для них не требуется валидация для определения диагностических характеристик (Разделы В.2 – В.4 ниже), если их результаты не используются для проведения окончательной диагностики в отношении предусмотренного назначения. Аналитическую точность таких методик можно установить путем сравнения с эталонным стандартом реактива или по характеристикам, присущим самому методу (например, титрация в конечной точке). Во всех этих примерах целевой аналит должен подвергаться

¹⁵ Доступно на:
http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/3.6.04_MEASUREMENT_UNCERT.pdf

дальнейшей количественной или качественной характеристике с помощью аналитического метода.

2.2. Этап 2 – диагностические характеристики анализа

Оценка диагностической чувствительности (доля проб от известных инфицированных эталонных животных с положительным результатом в анализе) и диагностической специфичности (доля эталонных точно не зараженных животных с отрицательным результатом в анализе) – это первичные показатели рабочих характеристик, установленные во время валидации анализа. Эти оценки – основа расчета других параметров, из которых делают выводы о результатах теста (например, прогнозируемые показатели положительных или отрицательных результатов тестов). Следовательно, обязательно, чтобы оценки диагностической чувствительности и диагностической специфичности были точными насколько возможно. В идеале, их получают в результате тестирования панели проб от эталонных животных с известной историей и статусом инфекции относительно данной болезни/инфекции, и имеющих значение для страны или региона, в которых будет использоваться тест. Анализ области под рабочей характеристической кривой – это полезное дополнение к оценке диагностической чувствительности и специфичности, так как он оценивает общую точность всех тестов по возможным показателям анализа (Greiner et al., 2000; Zweig & Campbell, 1993). Этот подход подробно описан в главе 3.6.5 (сноска 15).

Обозначенное количество известных положительных и известных отрицательных проб будет зависеть от вероятных значений диагностической чувствительности и диагностической специфичности анализов-кандидатов и желаемого уровня достоверности показателей (Таблица 1 и Jacobson, 1998). В таблице 1 представлены две панели теоретического количества необходимых проб, когда при оценке диагностической чувствительности и специфичности допускается 5% или 2% ошибок. Достаточно большое количество проб необходимо для достижения высокой степени уверенности (обычно 95%) в оценке диагностической чувствительности и специфичности, если желательно минимальное количество ошибок при оценке. Например, сравнение 2% ошибок против 5% ошибок для 90% и 95% уверенности в диагностической чувствительности и диагностической специфичности показывает значительное увеличение (864 против 138) в количестве требуемых образцов.

Ввиду логистических и финансовых ограничений может оказаться, что для оценки потребуется количество проб ниже оптимального, и в этом случае доверительный интервал, рассчитанный для диагностической чувствительности и диагностической специфичности, будет означать меньший диагностический доверительный уровень результатов. Объем выборки также может быть ограниченным ввиду возможного отсутствия эталонных популяций и эталонных стандартов МЭБ (смотри главу 3.6.5 *Руководства по наземным животным* для более подробной информации) (сноска 15). Следовательно, возможно сначала придется использовать количество проб ниже оптимального. Однако желательно повысить доверительный уровень и снизить допустимый уровень ошибок при оценке диагностической чувствительности и диагностической специфичности путем добавления дополнительного количества проб (эквивалентного статуса в исходную панель) по мере их поступления.

Таблица 2.1. Теоретическое количество проб от животных с известным статусом заражения, необходимое для определения диагностической чувствительности и диагностической специфичности в зависимости от вероятного значения диагностической чувствительности, диагностической специфичности, допустимой погрешности и доверительном уровне

Оценка диагностической	2% ошибка, допустимая при оценке диагностической	5% ошибка, допустимая при оценке диагностической
---------------------------	---	---

чувствительности или диагностической специфичности	чувствительности или диагностической специфичности			чувствительности или диагностической специфичности		
	Доверительный уровень			Доверительный уровень		
	90%	95%	99%	90%	95%	99%
90%	610	864	1493	98	138	239
92%	466	707	1221	75	113	195
94%	382	542	935	61	87	150
95%	372	456	788	60	73	126
96%	260	369	637	42	59	102
97%	197	279	483	32	45	77
98%	133	188	325	21	30	52
99%	67	95	164	11	15	26

Ниже приведены примеры эталонных популяций и методов, которые могут помочь при определении рабочих характеристик валидируемого теста.

2.2.1. Популяции эталонных животных

В идеале, при выборе эталонных животных необходимо, чтобы важные переменные факторы хозяев в целевой популяции представлены животными, выбранными для заражения или подвергания воздействию целевым возбудителем, или животными, которые никогда не инфицировались или никогда не подвергались воздействию. Такие переменные факторы включают вид, возраст, пол, породу, стадию инфекции, историю вакцинации и историю болезни стада (для дальнейшей информации смотри главу 3.6.6 *Руководства по наземным животным МЭБ* [сноска ⁸])

2.2.1.1. Отрицательные эталонные образцы

Истинные отрицательные образцы от животных, которые не подвергались заражению или воздействию возбудителем, вряд ли легко найти. Часто есть возможность получить такие образцы из стран, в которых болезнь была искоренена или никогда не регистрировалась. Такие образцы будут полезны, пока целевая популяция для анализа будет достаточно похожей на популяцию-источник образцов.

2.2.1.2. Положительные эталонные образцы

Обычно трудно найти достаточное количество действительно положительных животных, статус которых определен по результатам выделения патогенного организма. Возможно, потребуются образцы от животных, которые были протестированы с использованием других тестов, например, с использованием системы обнаружения нуклеиновых кислот. Кандидатный тест применяется с использованием эталонных образцов и результаты (положительные и отрицательные) классифицированы перекрестным образом в таблице 2x2. Это называется «моделью золотого стандарта», так как она предполагает, что эталонный стандарт является совершенным. Расчет пробы продемонстрирован в таблице 2, Разделе 2.2.5).

2.2.2. Образцы, отобранные от животных с неизвестным статусом

Когда так называемый эталонный стандарт является несовершенным, как обычно бывает с любым диагностическим тестом, оценка диагностической чувствительности и диагностической специфичности кандидатного теста, основанного на данном стандарте,

может быть ошибочной. Эту проблему можно частично решить путем проведения анализа латентных классов совокупных результатов двух тестов, предполагая, что ни один из тестов не является совершенным.

В моделях латентных классов не предполагается, что эталонный тест является совершенным, а проводится оценка точности кандидатного теста и эталонного стандарта с использованием совокупных результатов тестов (Branscum et al., 2005; Enoe et al., 2000; Georgiadis et al., 2003; Hui & Walter, 1980). Если используется байесовский анализ латентного класса, то в анализ должны быть включены предыдущие знания об эффективности эталонного теста и кандидатного теста.

Поскольку такие статистические модели достаточно сложны и требуют критических допущений, необходимо обратиться к помощи специалистов по статистике для проведения анализа и описания формы отбора проб из целевой популяции(й), характеристик других тестов, включенных в анализ, надлежащего выбора модели и методов оценки на основании рецензируемой литературы (см. главу 3.6.5 *Руководства по наземным животным МЭБ* [сноска ¹⁵] для получения подробной информации).

2.2.3. Экспериментально зараженные или вакцинированные эталонные животные

Пробы, полученные последовательно от экспериментально инфицированных или вакцинированных животных, полезны для определения кинетики гуморального ответа или присутствия/отсутствия антигена или организмов в пробах от таких животных. Однако множество серийных результатов, полученных до и после заражения от отдельных животных, не подходят для оценки диагностической чувствительности и специфичности, потому что это нарушает статистические требования независимых наблюдений. Для этой цели подходит только отбор проб в определенные моменты времени от отдельных животных (например, одна проба, произвольно отобранная от каждого животного). Тем не менее, необходимо отметить, что, не прямые методы обнаружения аналита, подвергание воздействию организмов в экспериментальных условиях или вакцинация могут стимулировать выработку гуморального ответа, который с качественной или количественной точки зрения может отличаться от естественной инфекции в целевой популяции (Jacobson, 1998). Штамм организма, доза и способ введения экспериментальным животным это переменные, которые могут внести ошибку, когда показатели диагностической чувствительности и специфичности будут экстраполироваться на целевую популяцию. В тех случаях, когда из-за невозможности получить подходящие эталонные пробы от животных, подвергнувшихся инфицированию в естественных условиях, приходится использовать пробы от экспериментальных животных для исследований по валидации, полученные показатели диагностической чувствительности и диагностической специфичности нужно считать менее совершенной оценкой истинной диагностической чувствительности и специфичности.

2.4. Определение точки раздела (порогового значения)

Для получения показателей диагностической чувствительности и специфичности кандидатного теста, который измеряется с использованием непрерывной шкалы, результаты теста сначала нужно свести к двум (положительной и отрицательной) или трем (положительной, промежуточной [сомнительной] или отрицательной) категориям результатов тестирования. Это завершается размещением одной или двух точек раздела (пороговое значение или точка принятия решения) на непрерывной шкале результатов теста. Выбор точки (ек) раздела должен отражать предусмотренное назначение анализа и его применение, а также поддерживать необходимую диагностическую чувствительность и специфичность анализа. Существуют разные варианты и описательные методы для

определения наилучшего способа выражения диагностической чувствительности и специфичности (Branscum *et al.*, 2005; Georgiades *et al.*, 2003; Greiner *et al.*, 1995; Greiner *et al.*, 2000; Jacobson, 1998; Zweig & Campbell, 1993; и глава 3.6.5). Если в распределении показателей тестов животных с известным заражением или точно не инфицированных животных наблюдается существенное перекрытие, невозможно выбрать одну точку раздела, которая точно поможет классифицировать этих животных в соответствии с их инфекционным статусом. Вместо одной лучше выбрать две точки раздела, определяющих высокую диагностическую чувствительность (например, включение 99% значений от инфицированных животных) и высокую диагностическую специфичность (например, 99% значений от неинфицированных животных), чем одну (Greiner *et al.*, 1995).

Основную сложность при установке точек раздела на основании диагностических характеристик представляет отсутствие необходимого количества хорошо охарактеризованных проб. Альтернативы обсуждаются в Разделе 2.2.6 о предварительном принятии анализа во время получения данных для подкрепления оценки диагностической чувствительности и специфичности».

2.5. Расчет диагностической чувствительности и специфичности на основании результатов тестирования эталонных сывороток

Типичным методом для определения значений диагностической чувствительности и специфичности является тестирование эталонных образцов с помощью нового анализа и составление комбинированной 2×2 таблицы категорийных результатов тестов. В гипотетическом примере разработчик теста может сделать допущение, что показатели диагностической чувствительности и специфичности должны быть 97% и 99%, соответственно, с желаемым доверительным уровнем 95% для обоих показателей. Допустимая ошибка в оценках составляет 2%. В таблице 2.1. показано, что для оценки диагностической чувствительности нужно 279 проб от животных с известным заражением, а для оценки диагностической специфичности – 95 известно отрицательных проб. Затем пробы тестируют в новом анализе. Таблица 2 представляет собой гипотетический набор результатов, на основании которых были получены показатели диагностической чувствительности и специфичности.

В данном примере получены ожидаемые показатели диагностической чувствительности, но диагностическая специфичность оказалась намного ниже (92%) ожидаемого значения 99%. Как следствие, ширина доверительного интервала для диагностической специфичности получилась больше, чем ожидалась. Повторный просмотр Таблицы 2.1. показывает, что для получения границы ошибки $\pm 2\%$ при диагностической специфичности 92% нужно 707 проб, но такое увеличение объема выборки может быть невозможным (см. главу 3.6.5 *Руководства по наземным животным МЭБ* [сноска ¹⁵] для получения дальнейшей информации).

Таблица 2.2. Показатели диагностической чувствительности и специфичности, рассчитанные по гипотетическому набору результатов протестированных проб от животных с известным заражением и неинфицированных животных

		Необходимое количество эталонных проб*			
		Известные положительные (279)		Известные отрицательные (95)	
Результаты теста	Положительные	270	ИП	ЛП	7
	Отрицательные	9	ЛО	ИО	88

	Диагностическая чувствительность* ИП/(ИП + ЛО) 96,8% (94,0 – 98,5%)**	Диагностическая специфичность* ИО/(ИО + ЛП) 92,6% (85,4 – 97,0%)**
--	---	--

* На основании таблицы 2.1. для анализа со следующими параметрами:

- 1) перед тестированием, показатель диагностической чувствительности 97%, а специфичности – 99%
- 2) 95% = требуемый доверительный уровень для показателей диагностической чувствительности и специфичности
- 3) 2% = допустимый предел ошибки при оценке диагностической чувствительности и специфичности

ИП и ЛП = истинно положительные и ложноположительные результаты, соответственно

ИО и ЛО = истинно отрицательные и ложноотрицательные результаты, соответственно

** 95% точный биномиальный доверительный уровень для рассчитанных показателей диагностической чувствительности и специфичности (информацию о доверительных уровнях смотрите главу 3.6.5 *Руководства по наземным животным МЭБ* [сноска¹⁵])

2.6. Предварительное признание анализа¹⁶

Существуют ситуации, когда невозможно или нежелательно выполнить Этап 2 программы валидации, вследствие недостаточного количества соответствующих проб от целевых животных и трудностей доступа к животным (например, трансграничные инфекционные болезни или болезни диких животных).

Опыт показал, что наибольшим препятствием при прохождении Этапа 2 программы валидации является количество определенных образцов, требуемых для подсчета диагностической специфичности и диагностической чувствительности. Формула является хорошо известной и существуют таблицы для определения количества образцов, требуемых для оценки различных уровней диагностической специфичности и диагностической чувствительности, в зависимости от желаемого предела ошибки и доверительного уровня показателей (Таблица 1 Jacobson, 1998). Формула предполагает, что учитываются мириады факторов хозяина/организма, которые могут повлиять на результаты тестирования. Так как данное предположение может быть под вопросом, предполагаемые размеры образца являются, мягко говоря, минимальными. Что касается болезни, которая не является эндемичной или широко распространенной, возможно будет невозможным первоначально получить требуемое количество образцов, но с течением времени получение дополнительной информации позволит скорректировать точку раздела (пороговое значение) или если такая корректировка не требуется, повысить доверительный уровень показателей.

Предварительное признание определяет анализ, который был оценен на Этапе 1 на основные критические параметры анализа (чувствительность анализа, специфичность анализа, повторяемость), а также на предварительные показатели диагностической специфичности и диагностической чувствительности, на основании панели хорошо охарактеризованных образцов малого выбора, содержащей целевой аналит и предварительные показатели воспроизводимости. Это представляет собой частичное завершение Этапа 2. Предварительные показатели воспроизводимости кандидатного анализа могут быть получены с использованием

¹⁶ Предварительное признание не подразумевает утверждение МЭБ. Однако это является признанием решение органов на местном, государственном или международном уровнях об условном разрешении частично валидированного анализа

панели выбора хорошо охарактеризованных образцов для усиления предварительного статуса принятия анализа. Кандидатный метод тестирования обычно дублируют в лабораториях, по крайней мере, в двух институтах и панель образцов оценивают с использованием кандидатного анализа в каждой из лабораторий с использованием того же протокола, тех же реагентов, которые указаны в протоколе и сопоставимого оборудования.

Это уменьшенный вариант изучения воспроизводимости на Этапе 3 валидации анализа. При соблюдении данной процедуры предварительного признания протокол тестирования не должен быть изменен.

Предварительное признание анализа государственными или национальными органами означает, что анализ не проходил оценку на диагностические характеристики. По существу для удовлетворения этого требования лаборатория должна разработать протокол добавления и оценки образцов, как только они становятся доступными, и следовать ему. В идеале данный процесс должен быть ограничен определенными временными рамками. В течение этого времени такое накопление должно быть направлено на выполнение Этапов 2 и 3 программы валидации и определенные ситуации (чрезвычайные ситуации, мелкие виды, отсутствие других тестов и т.д.).

2.3. Этап 3 – Воспроизводимость и повышенные показатели повторяемости

2.3.1. Воспроизводимость

Воспроизводимость – это способность метода тестирования показывать постоянные результаты, определяемые по оценке прецизионности, в отношении аликвот одних и тех же проб, протестированных в разных лабораториях, предпочтительно расположенных в определенных или разных регионах или странах, с использованием одного и того же анализа (протокол, реагенты и контролей). Для оценки воспроизводимости анализа, каждая из, по крайней мере, трех лабораторий проводит тестирование одной и той же панели проб (вслепую), содержащих минимум 20 проб с одинаковыми аликвотами, которые поступили в каждую лабораторию (см. главу 3.6.6 *Руководства по наземным животным* МЭБ [сноска ⁸]). Такое испытание также позволяет получить предварительные данные по неслучайным эффектам, связанным с применением анализа в других лабораториях. Кроме этого, показатели внутрилабораторной повторяемости возрастают с дублирующими пробами, используемыми в исследованиях воспроизводимости. Показатели прецизионности можно оценить как для данных воспроизводимости, так и повторяемости (см. главу 3.6.4 *Руководства по наземным животным* МЭБ [сноска ¹⁴]) для дальнейших пояснений и по применению).

Для полевых испытаний воспроизводимость нужно оценивать в условиях назначенной цели.

2.3.2. Назначение валидированного анализа

По завершении Этапа 3 валидации при условии, что предварительные этапы были полностью завершены с удовлетворительными результатами, анализ может быть определен как «валидированный для первоначального целевого назначения». Сохранение данного определения зависит от постоянного мониторинга характеристик анализа, как описано в Разделе 2.5.1.

2.4. Этап 4 – Осуществление программы

Успешное применение анализа обеспечивает дополнительными и ценными доказательствами его рабочих характеристик в соответствии с прогнозами. Более того, (истинная) превалентность диагностической особенности целевой популяции является важным фактором, который необходимо учитывать в соответствии с нижеуказанным.

2.4.1. Соответствие назначению

Вместе с тем, что глава посвящена валидации и соответствию назначению с научной точки зрения, необходимо также отметить, что другие практические факторы влияют на использование анализа с точки зрения его предполагаемого назначения. Эти факторы включают не только диагностическое соответствие анализа, но и его приемлемость для научных и регламентирующих сообществ, клиентов и осуществимость при наличии доступных ресурсов лаборатории. В отношении некоторых болезней многочисленные анализы могут быть доступны для использования в комбинации с контролем болезней и программами надзора и таким образом, применение анализа, возможно, требуется оценить посредством оценки изменений в диагностической специфичности, диагностической чувствительности и прогностических значениях комбинированных тестов.

Из-за неспособности соответствовать операционным требованиям к анализу он так же может быть непригодным для предусмотренного назначения. Такие требования могут включать: затраты на обеспечение результативности, доступность оборудования, уровень технических знаний и умений интерпретирования, доступность набор/реагента, срок годности, требования к транспорту, безопасность, биобезопасность, пропускная способность по исследованиям образцов, время цикла получения результатов, аспекты контроля качества и обеспечения качества, а также возможность использования анализа другими лабораториями. Тест-наборы, используемые в полевых условиях очень востребованы с точки зрения простоты использования, но из-за того, что их применяют за пределами контролируемой лабораторной среды, необходимо соблюдать дополнительные меры предосторожности для сохранения соответствия предусмотренному назначению (Crowther et al., 2006).

2.4.2. Интерпретация результатов

Прогнозируемые показатели результатов теста: Положительное прогнозируемое значение (PPV)– это вероятность того, что животное, результаты тестирования которого были положительными, действительно является положительным с точки зрения фактического диагностического статуса. Отрицательное прогнозируемое значение (NPV) это вероятность того, что животное, результаты тестирования которого были отрицательными, действительно является отрицательным с точки зрения фактического диагностического статуса.

Прогнозируемые показатели результатов теста рассчитываются при помощи теоремы Байеса указанным ниже способом:

$$PPV = \frac{P \times DSe}{P \times DSe + (1 - P) \times (1 - DSp)}$$

and

$$NPV = \frac{(1 - P) \times DSp}{P \times (1 - DSe) + (1 - P) \times DSp}$$

где,

PPV – прогнозируемые показатели положительных результатов теста

NPV – прогнозируемые показатели отрицательных результатов тестирования

P – превалентность инфекции

DSe – диагностическая чувствительность

DSp – диагностическая специфичность

По сравнению с диагностической чувствительностью и диагностической специфичностью на прогнозируемые показатели влияет фактическая превалентность фактического диагностического статуса целевой популяции. Другими словами прогнозируемые показатели не являются присущими характеристиками определенного диагностического теста, но являются функцией его диагностической чувствительности и диагностической специфичности и местной превалентностью инфекции в определенной популяции в определенный отрезок времени.

Прогнозируемые показатели представляют важность для ветеринаров, работающих в полевых условиях для интерпретации результатов. Например, PPV равно 0,9 означает, что животное, у которого наблюдается положительная реакция на тест, имеет 90% шанс того, что оно действительно является инфицированным и 10% вероятность быть ложноположительным.

Прогнозируемые показатели положительных результатов теста имеют огромную важность для ветеринарных служб, ответственных за управление контролем или программы по искоренению. Если мы будем учитывать обратное значение прогнозируемых показателей положительных результатов теста (например, $1/PPV$), то это даст информацию о количестве денег, потраченных на отбраковку фактически и ложноположительных животных на каждое фактически положительное животное, обнаруженное в ходе надзора. Другими словами если PPV равно 0,67, это означает, что два положительных животных из трех являются фактически положительными, а остальные ложноположительными. Так как во время применения программы контроля превалентность инфекции постоянно меняется, мониторинг PPV является способом оценки затрат программы.

Более того, во время применения программы по контролю обычно предпочтительно изменять чувствительность теста, на основании вариации превалентности инфекции в целевой популяции и задачи программы, PPV может быть использован для внесения изменений в диагностическую чувствительность или диагностическую превалентность из экономических соображений. Другими словами, когда возникает необходимость внесения изменений в диагностическую чувствительность и в диагностическую специфичность, вдоль кривой ROC валидации теста могут быть установлены предполагаемые точки раздела и соответствующие величины диагностической чувствительности и диагностической специфичности для каждой точки раздела могут быть использованы для оценки предполагаемых затрат на отбраковку каждого инфицированного животного.

Если задачей является установление доказательств свободы от болезни, NPV является наиболее важной мерой. NPV в значительной мере зависит от диагностической чувствительности.

2.4.3. Признание на международном уровне

Традиционно, за признание анализов на международном уровне отвечает МЭБ, который назначает их в качестве предписанных или альтернативных тестов для международной торговли. Часто это происходит на основании данных их полезности на национальном, региональном или международном уровне. Для коммерческих тест-наборов, которые уже были подвергнуты валидации и сертификации диагностических анализов,

заключительным этапом является включение в список тестов в Реестре МЭБ. Тесты, указанные в Реестре, сертифицированы как пригодные для специальной цели, если они завершили этапы валидации 1, 2 и 3. Цель Реестра – предоставить потенциальным пользователям информативный и объективный источник информации о тест-наборе и его рабочих характеристиках для предусмотренного назначения. Реестр доступен на веб-сайте МЭБ: www.oie.int/en/our-scientific-expertise/certification-of-diagnostic-tests/the-register-of-diagnostic-tests.

2.4.4. Внедрение анализа

Последнее доказательство полезности анализа – это его успешное применение в других лабораториях и включение в национальные, региональные и/или международные программы. Референтные лаборатории играют важную роль в этом процессе. Естественный прогресс в области диагностических и технологических разработок способствует тому, что новые анализы становятся новыми стандартными методами, с которыми сравнивают другие анализы. Как следствие, они могут постепенно получать признание на национальном, региональном и международном уровне. В качестве признанного стандарта эти анализы также можно использовать для разработки эталонных реактивов для целей контроля качества, повышения квалификации и гармонизации. Такие эталонные реактивы могут также стать международными стандартами.

Оценку воспроизводимости следует проводить повторно, когда тест переносят из лаборатории-разработчика в полевые условия, относится ли это к использованию в местных лабораториях или к анализу проб в полевых условиях. Предсказуемые изменения, например, крайние температуры и уровень опыта оператора, нужно оценивать как дополнительные источники вариации в результатах анализа, которые могут повлиять на показатели воспроизводимости.

2.5. Мониторинг характеристик анализа после первоначальной валидации

2.5.1. Мониторинг анализа

Для сохранения статуса валидированного анализа необходимо обеспечить, чтобы первоначально валидированный анализ постоянно поддерживал рабочие характеристики, определенные во время валидации анализа. Это может быть определено в программах по обеспечению качества посредством мониторинга ежедневных результатов анализа первоначально посредством показателей прецизионности и точности для внутренних контролей, а также других тенденций статистического выброса. Мониторинг результативности можно осуществлять графическим способом посредством составления графика контрольных параметров на контрольных картах¹⁷. Отклонения от ожидаемых характеристик следует анализировать, чтобы принять корректирующие меры при необходимости. Такой мониторинг позволяет получить критические доказательства, что анализ сохраняет свое «валидированное» назначение во время фазы внедрения анализа. Воспроизводимость оценивают с использованием программ внешнего контроля качества, таких как квалификационные испытания. Если анализ прекращает демонстрировать результаты, которые соответствуют первоначальным данным по валидации, то анализ должен считаться непригодным для предусмотренного назначения. Так, валидированный анализ может быть подвержен постоянной оценке для обеспечения его пригодности предусмотренному назначению.

¹⁷ Контрольная карта: Графическое представление данных повторного измерения контрольных образцов, исследуемых в различных циклах анализа в течение периода времени.

2.5.2. Модификации и улучшения – соображения по поводу изменений в анализе

Со временем может возникнуть необходимость в модификации анализа в связи с изменениями в предусмотренном назначении (т.е. модификации анализа для корректировки диагностической результативности) или технических модификациях для улучшения эффективности или рентабельности анализа. Для изменения предусмотренного назначения анализа обязательно проведение пересмотренной валидации начиная со 2 этапа.

Если анализ будет применяться в другом географическом регионе и/или популяции, рекомендуется провести повторную валидацию при новых условиях. Линии или сублинии инфекционного агента, выделенного от животных в разных географических регионах, отличаются и требуют повторной валидации анализа для указанной целевой популяции.

Особенно это относится к системам обнаружения нуклеиновых кислот (NAD), поскольку у многих возбудителей инфекционных болезней (например, в РНК вирусах) происходят точечные мутации. Мутации, которые могут происходить в сайтах праймеров или зондов, могут влиять на эффективность анализа и даже делать недействительными рабочие характеристики. Также рекомендуется регулярно подтверждать целевую последовательность в выбранных областях генома для национальных или региональных изолятов возбудителей инфекционных болезней. Особенно это касается сайтов праймеров или зондов: необходимо обеспечить, чтобы они оставались стабильными, а показателям диагностической чувствительности и специфичности анализа ничего не угрожало.

Похожая ситуация может произойти с появлением новых подтипов существующих патогенных организмов. В данных обстоятельствах, возможно, потребуется модифицировать существующие анализы.

2.5.2.1. Технические модификации и оценки сопоставимости

Технические модификации валидированного анализа, например, изменения в приборах, протоколах экстрагирования и конверсии анализа в полуавтоматическую или автоматическую систему с использованием робототехники обычно не требуют проведения полной повторной валидации анализа. Скорее, сравнительное исследование методов проводится для определения того, могут ли относительно незначительные модификации анализа повлиять на ранее задокументированные рабочие характеристики анализа. Сопоставимость можно установить путем одновременного проведения модифицированной процедуры и оригинальной процедуры с такой же панелью образцов в обоих тестах в нескольких циклах. Панель, выбранная для этого сравнения должна представлять весь рабочий диапазон обоих анализов. Если результаты модифицированной процедуры и изначально валидированного метода признаны сопоставимыми в эксперименте на основании ранее определенного критерия, модифицированный анализ остается валидным для своего предусмотренного назначения. Описание экспериментов, которые подходят для тестирования сопоставимости см. в главе 3.6.8. *Руководства по наземным животным МЭБ* [сноска ⁹], а информацию о панелях проб см. в главе 3.6.6. *Руководства по наземным животным МЭБ* [сноска ⁸].

2.5.2.2. Биологические модификации и оценка сопоставимости

Могут возникать ситуации, когда необходимо или оправданно внести изменения в какие-либо биологические препараты, используемые в анализе. Это может включать изменения в саму образец исследования (напр. изменения в ткани, подлежащей

исследованию или, возможно, исследование в параллели другого вида животного). Это может включать изменения в реагенты (напр. замена рекомбинантного антигена на антиген, полученный из культуры ткани или замена одного конъюгата антитела на другой со схожей иммунологической специфичностью в ИФА). Трудности при осуществлении любого изменения лежат в определении того, требует ли изменение полной ревалидации анализа, как на документарном уровне, так и на уровне применения. По меньшей мере, любое изменение требует оценки соответствующих «аналитических предпосылок» Этапа 1. Более трудное решение относится к «диагностическим характеристикам» Этапа 2. Чтобы помочь в этом, исходный (эталонный) анализ следует сначала сравнить с модифицированным (кандидатным) анализом в контролируемом исследовании с использованием определенной панели положительных и отрицательных диагностических образцов. См. Раздел 3.6.8 *Руководства по наземным животным* (сноска ⁹) для описания оценки сопоставимости. Если оценка сопоставимости не предполагает изменения диагностических характеристик, модифицированный анализ можно ввести в повседневное использование. Если, с другой стороны, наблюдаются различия в диагностической чувствительности и диагностической специфичности, для применения модифицированного анализа потребуется дополнительный Этап 2 или проверка в полевых условиях.

2.5.2.3. Замена истраченных реактивов

Если реактив, например контрольный образец или рабочий стандарт, подходит к концу, нужно приготовить и повторно протестировать замену до того, как этот контрольный образец закончится. Будущий контрольный образец необходимо включить в несколько циклов анализа параллельно с оригинальным контролем для определения их пропорциональной связи. По возможности нужно менять только один реактив за раз, чтобы избежать сложностей с оценкой более чем одной переменной.

2.5.3. Укрепление уверенности в критериях валидации

Так как большинство переменных, связанных с хозяевами, оказывают влияние на диагностические характеристики анализов, желательно со временем увеличивать количество эталонных проб или проб подходящих для анализа латентных классов. План пробоотбора, сбор, транспортировка и среда проведения теста в отношении новых проб должны быть теми же самыми, что и использованные для первоначальной валидации. Увеличение количества проб повышает прецизионность общих показателей диагностической чувствительности и диагностической специфичности и позволяет рассчитывать показатели диагностической чувствительности по таким факторам, как возраст, стадия болезни и нагрузка организма. Ежегодно в соответствующие досье теста нужно включать новые данные.

2.5.4. Верификация существующих анализов (внутренняя валидация)

Если лаборатория рассматривает использование валидированного коммерческого набора или кандидатного анализа на основании опубликованной литературы с данными по валидации то потребуется проведение верификации для определения соответствия анализа заявлениям производителя набора или автора, применительно к критериям валидации Этапа 1, в контексте предусмотренного применения. Это может потребовать проведения ограниченной верификации Asp и ASe с использованием доступных референтных материалов независимо от того, являются ли они внешними и/или приобретенными на местном уровне из целевой популяции. Когда лаборатория уверена, что анализ выполняется в соответствии с описанием

с аналитической точки зрения, то при переходе на ограниченный Этап 2 валидация должна рассматриваться в контексте предусмотренного применения и целевой популяции, перед тем как анализ внедряется в рутинное диагностическое использование.

2.3. ЛИТЕРАТУРНЫЕ ИСТОЧНИКИ

BRANSCUM A.J, GARDNER I.A. & JOHNSON W.O. (2005). Estimation of diagnostic-test sensitivity and specificity through Bayesian modelling. *Prev. Vet. Med.*, **68**, 145–163.

CROWTHER J.R., UNGER H. & VILJOEN G.J. (2006). Aspects of kit validation for tests used for the diagnosis and surveillance of livestock diseases: producer and end-user responsibilities. *Rev. sci. tech. Off. int. Epiz.*, **25** (3), 913– 935.

DEJAEGHER B. & VANDER HEYDEN Y. (2006). Robustness tests. *LCGC Europe*, **19** (7), online at <http://www.lcgceurope.com/lcgceurope/content/printContentPopup.jsp?id=357956>.

ENØE C., GEORGIADIS M.P. & JOHNSON W.O. (2000). Estimating the sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Prev. Vet. Med.*, **45**, 61–81.

FINDLAY J.W.A. & DILLARD R.F. (2007). Appropriate calibration curve fitting in ligand binding assays. *AAPS J.*, **9** (2), E260-E267. (Also on-line as AAPS Journal [2007]; 9 [2], Article 29 [<http://www.aapsj.org>]).

GEORGIADIS M., JOHNSON, W., GARDNER I. & SINGH R. (2003). Correlation-adjusted estimation of sensitivity and specificity of two diagnostic tests. *Appl. Statist.*, **52** (Part 1), 63–76.

GREINER M., SOHR D. & GÖBEL P. (1995). A modified ROC analysis for the selection of cut-off values and the definition of intermediate results of serodiagnostic tests. *J. Immunol. Methods*, **185**, 123–132.

GREINER M., PFEIFFER D. & SMITH R.D. (2000). Principles and practical application of the receiver operating characteristic (ROC) analysis for diagnostic tests. *Vet. Prev. Med.*, **45**, 23–41.

HUI S.L. & WALTER S.D. (1980). Estimating the error rates of diagnostic tests. *Biometrics*, **36**, 167–171.

JACOBSON R.H. (1998). Validation of serological assays for diagnosis of infectious diseases. *Rev. sci. tech. Off. int. Epiz.*, **17**, 469–486.

VESSMAN J., STEFAN R., VAN STADEN J., DANZER K., LINDNER W., BURNS D., FAJGELJ A. & MULLER H. (2001). Selectivity in analytical chemistry. *Pure Appl. Chem.*, **73** (8), 1381–1386.

WORLD ORGANISATION FOR ANIMAL HEALTH (OIE) (2008). OIE Standard for Management and Technical Requirements for Laboratories Conducting Tests for Infectious Diseases. OIE Quality Standard and Guidelines for Veterinary Laboratories: Infectious Diseases. OIE, Paris, France, 1–31.

ZWEIG M.H. & CAMPBELL G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin. Chem.*, **39**, 561–577.

*

* *