

## ВАЛИДАЦИЯ И ВЕРИФИКАЦИЯ МЕТОДА: ЦЕЛИ И СРЕДСТВА ИХ ДОСТИЖЕНИЯ.

Валидация и верификация метода или аналитической системы как часть процесса планирования качества (QP) системы Комплексного Управления Качеством (TQM) медицинских лабораторий.

После выбора метода или аналитической системы следующим этапом в рамках процесса планирования аналитического качества (QP) и на пути его создания в цикле TQM, является их валидация или верификация (Рис.1).

- Выбор метода или аналитической системы.
- Валидация или верификация метода или аналитической системы.
- Планирование качества для аналитического метода или системы.

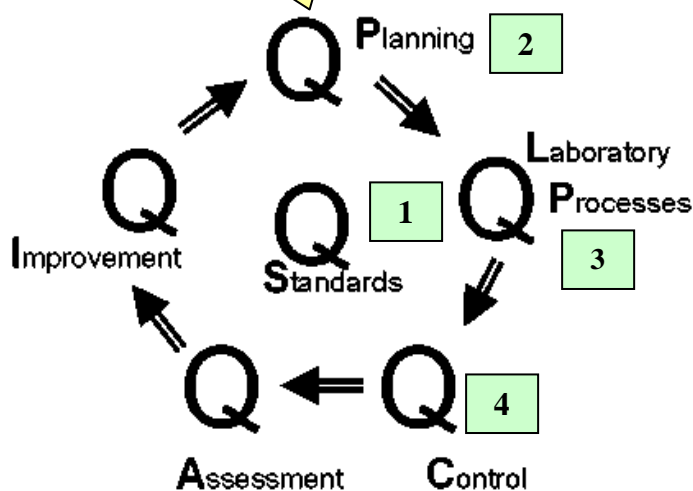


Рисунок 1. Процесс QP и его этапы в системе TQM. Цифрами обозначены этапы цикла TQM, связанные с процессом создания аналитического качества в медицинской лаборатории.

Вообще, для большинства специалистов стран СНГ, понятия валидация и верификация метода являются новыми. В связи с этим, постараюсь дать подробную характеристику этим терминам.

## Валидация и верификация метода: определения терминов и цели.

Согласно документу ISO 15189 Медицинские лаборатории - Требования к качеству и компетенции (ISO 15189 Medical laboratories - Requirements for quality and competence)[1], валидация это - подтверждение путем представления объективных доказательств того, что требования для специфического использования или применения были выполнены. Другими словами, на практике, валидация метода предусматривает выполнения в медицинской лаборатории комплекса мероприятий по оценке аналитической эффективности метода и сравнения её с целями аналитического качества. С другой стороны, согласно этому же документу верификация – это подтверждение путем представления объективных доказательств того, что специфические требования были вы-

полнены. На практике, это означает выполнение в медицинской лаборатории мероприятий по подтверждению спецификаций производителя метода по аналитической эффективности метода.

Как видим, термины валидация и верификация подразумевают выполнение в лаборатории различных мероприятий. Основа различий состоит в том, что мероприятия по валидации метода предусматривают работу по оценке его аналитической эффективности в лаборатории и определении её приемлемости путем сравнения с целями качества, в то время как мероприятия по верификации метода предусматривают работу по доказательству его аналитической эффективности, установленную и заявленную производителем. Далее, если эта эффективность подтвердится (верифицируется), то лаборатория может использовать данные производителя для планирования процедуры Статистического Контроля Качества (СКК). На рисунке 2 в виде блок-схемы представлен обобщенный процесс верификации/валидации метода или аналитической системы (АС) в медицинской лаборатории.

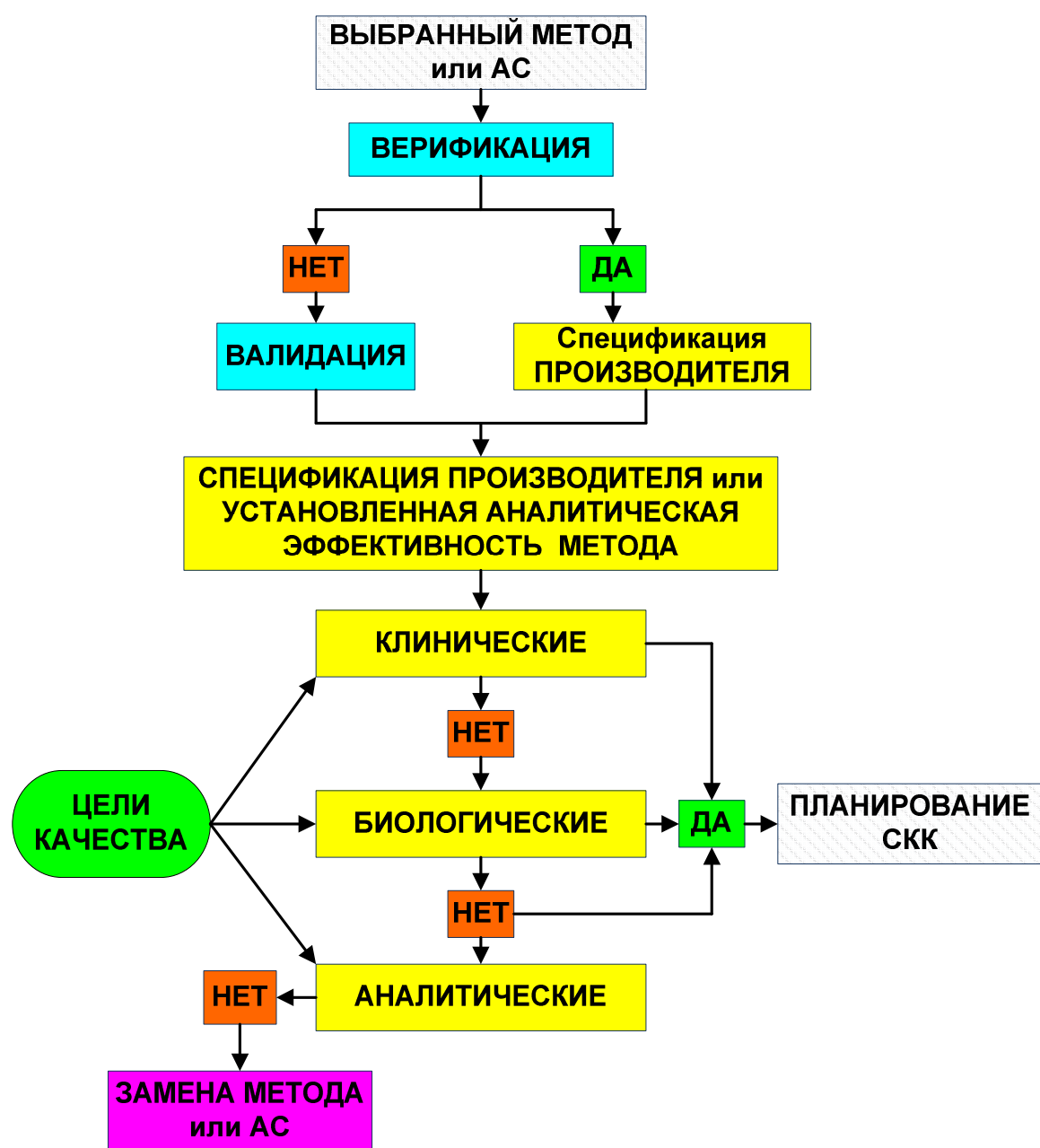


Рисунок 2. Схема процессов верификации/валидации метода/АС в медицинской лаборатории. Как правило, для большинства методов, сначала усилия лабораторных аналитиков направлены на верификацию метода, а валидацию выполняют только тогда, когда его верификация не состоится.

**Все мероприятия по валидации или верификации метода несут в себе тайный смысл, который заключается в определении\доказательстве величин его ошибок и сравнении их с целевыми ошибками требований к качеству.** Действительно, аналитическая эффективность метода определяется величиной его ошибок, а приемлемость для клиники определяется на основании результата сравнения этих ошибок с теми целевыми, которые допустимы для клиники. Тогда, для того, чтобы метод был приемлем для клиники, величины наблюдаемых в нем ошибок должны быть небольшими относительно целевых ошибок, содержащихся в требованиях к качеству и наоборот, если ошибки метода будут велики относительно целевых ошибок, то тогда метод не будет приемлем для клиники. Фокусирование внимания на аналитических ошибках метода и сравнении их с целевыми ошибками требований к качеству, является краеугольным камнем всех мероприятий процесса его верификации или валидации. Таким образом, **цель верификации\валидации метода или АС заключается в определении (или доказательстве) величин ошибок метода\АС, сравнении этих ошибок с ошибками целей качества, и на основании этого сравнения определении их приемлемости для клиники.** Приемлемость для клиники служит гарантией минимальной вероятности помех со стороны ошибок верифицируемого или валидируемого метода на клиническую интерпретацию результатов тестирования. Принимая во внимание важность оценки аналитических ошибок в процессе проведения верификации\валидации метода или АС, остановимся на основных их типах.

## **Аналитические ошибки и их типы.**

В принципе, выделяют два основных типа аналитических ошибок : случайная ошибка (Random Error (RE)) и систематическая ошибка (Systematic Error (SE)). Систематическая ошибка состоит из константной (Constant Systematic Error) и пропорциональной (Proportional Systematic Error) ошибок. Кроме того, на основании эффектов RE и SE, определяют общую ошибку (Total Error (TE)).

- Случайную аналитическую ошибку (Random Error (RE)) (Рис.3) можно определить как результат измерения ( $x_i$ ) минус среднее значение, полученное в результате бесконечного количества измерений того же самого мезюранда ( $\bar{x}$ ). Мезюранд - это величина, которую предполагают измерить (VIM)[8] или, другими словами, любой аналит, который может быть измерен в медицинской лаборатории. Математически, RE можно выразить как:  $\Delta RE = x_i - \bar{x}$ . Направление изменения и величину  $\Delta RE$  невозможно предсказать, а её значение (по модулю) всегда больше нуля ( $\Delta RE > 0$ ). Величины  $\Delta RE$  отдельных измерений формируют величину неprecизионности повторных измерений. Сама же величина неprecизионности измерений определяется дисперсией результатов отдельных измерений вокруг среднего значения мезюранда ( $\bar{x}$ ). Чем выше величина  $\Delta RE$  отдельных измерений, тем больше величина дисперсии их результатов и, следовательно, тем больше величина неprecизионности всех измерений. В принципе, величину неprecизионности измерений аналита можно рассматривать как “среднюю” величину  $\Delta RE$  его отдельных измерений. Как правило, в медицинской лаборатории, величину неprecизионности и, соответственно, “среднюю”  $\Delta RE$  отдельных измерений, определяют путем расчета величины стандартного отклонения (Standard Deviation (SD)) результатов повторных измерений, которая равна корню квадратному из величины дисперсии. Часто, величина SD возрастает по мере увеличения концентрации аналита. В связи с этим, часто на практике вместе с величиной SD рассчитывают ве-

личину коэффициента вариации (Coefficient of Variation (CV)), тем самым ”привязывая” величину SD к величине Среднего Значения.

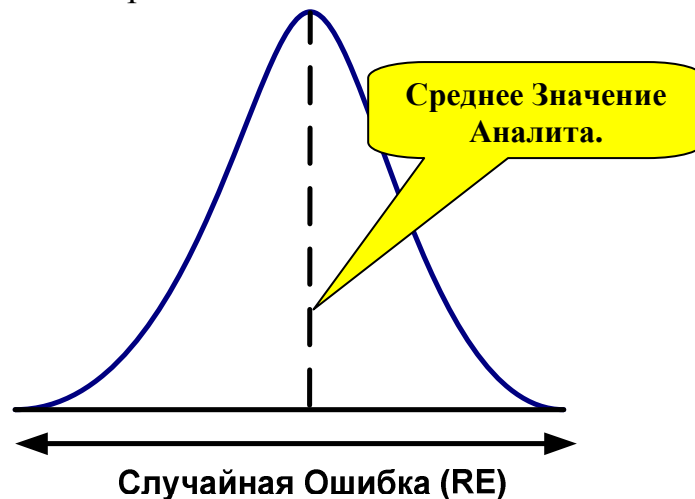


Рисунок 3. Случайная ошибка (Random Error (RE)) повторных измерений аналита или непрецизионность измерения. Формируется в результате распределения результатов повторных измерений аналита вокруг его среднего значения. Количественно, ширина этого распределения и, следовательно, величина непрецизионности измерения аналита, определяется величиной дисперсии, стандартного отклонения или коэффициента вариации.

- Систематическую аналитическую ошибку (Systematic Error (SE)) или аналитическое смещение (Bias) определяют как разницу между средним значением мезюранда ( $\bar{x}$ ), полученным в результате его бесконечного количества измерений, выполненных в условиях повторяемости и ”истинным” значением мезюранда ( $\mu$ ). Математически, SE можно выразить как:  $\Delta SE = \bar{x} - \mu$ . В отличие от случайной ошибки, систематическая ошибка всегда имеет одно направление, которое можно предсказать (Рис.4) и её величина может равняться нулю ( $\Delta SE \geq 0$ ). Следовательно, SE является причиной завышения, либо занижения результатов тестирования. Как правило, в медицинской лаборатории, величину SE определяют либо на основании результатов ВОК, либо на основании сравнительного анализа методов или аналитических систем, используя для этого статистику линейной регрессии, либо статистику парного t-теста.

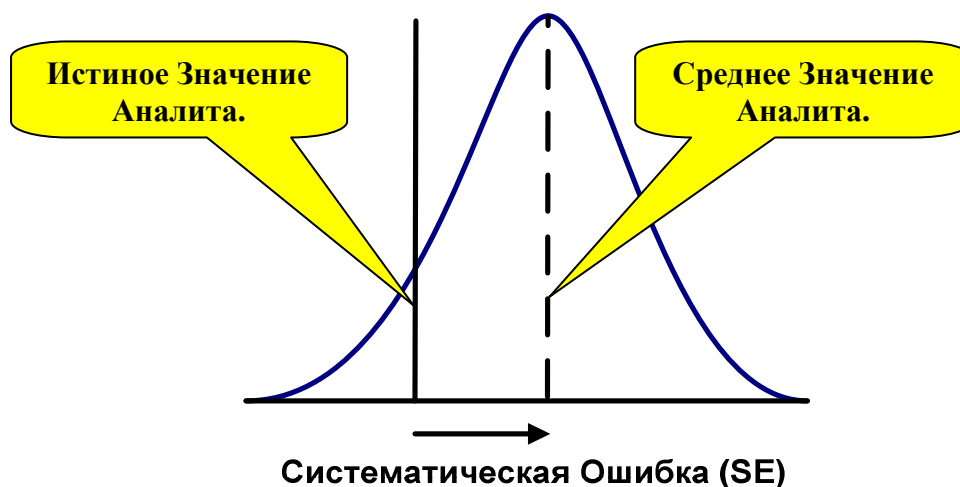


Рисунок 4. Систематическая Ошибка (Systematic Error (SE)). Формируется как разность между наблюдаемым средним значением аналита и его истинным значением. В связи с тем, что многие аналиты не имеют полной цепочки метрологической прослеживаемости к единицам SI, для них в качестве ”истинного” значения могут выступать результаты метода сравнения.

- ◆ Константная систематическая ошибка (Constant Systematic Error) обусловлена влиянием на концентрацию исследуемого аналита интерферентов (см. ниже) и её величина не меняется на всем протяжении рабочего диапазона метода.
- ◆ Пропорциональная систематическая ошибка (Proportional Systematic Error) обусловлена влиянием на концентрацию исследуемого аналита матрикса пробы, её величина возрастает по мере возрастания концентрации исследуемого аналита (Рис.5).

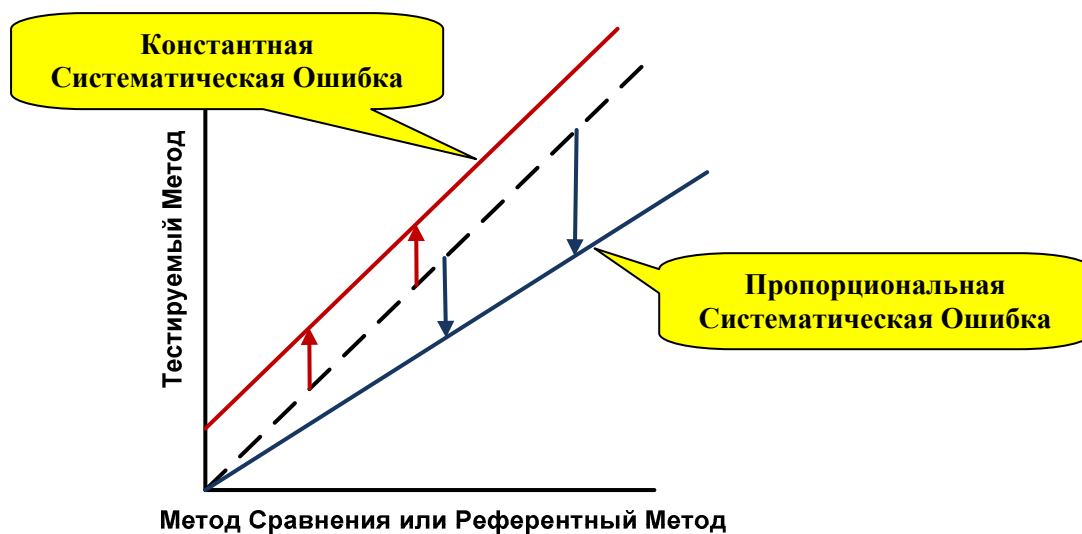


Рисунок 5. Константная и Пропорциональная Систематические Ошибки.

- Общая ошибка определения аналита (Total Error (TE)) определяется чистым или комбинированным эффектом SE и RE на результат измерения аналита. Математически, TE можно выразить так :  $TE = \Delta SE + \Delta RE$ . Так как величина  $\Delta RE$  всегда больше нуля, то, следовательно, и величина TE также будет всегда больше нуля ( $TE > 0$ ). В связи с этим, величина TE отдельного результата измерения аналита не должна превышать установленной её величины. Эта установленная величина TE называется "допустимая ошибка определения аналита" ("allowable total analytical error" (aTE)). Математически, TEa можно выразить так:  $TEa = \Delta SE + 1,65 \times \Delta RE$ , где 1,65 - это величина Z-значения, необходимого для формирования 95%-ного одностороннего доверительного интервала для величины  $\Delta RE$  (Рис.6).

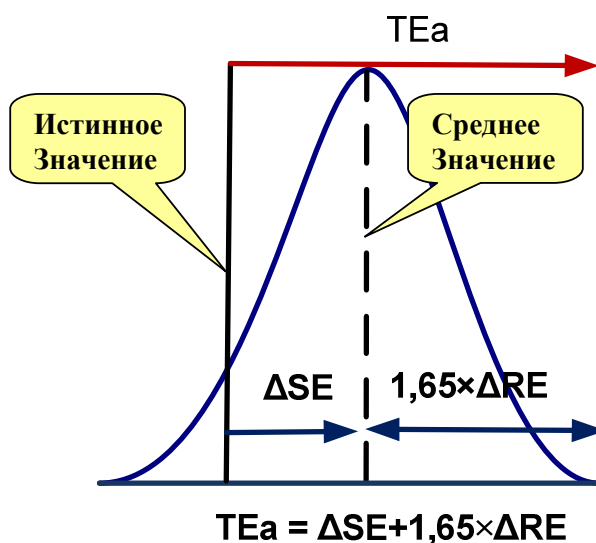


Рисунок 6. Концепция допустимой Общей Ошибки определения аналита (TEa). TEa определяется как сумма величин допустимой систематической ошибки ( $\Delta SE$ ) и допустимой случайной ошибки ( $\Delta RE$ ), умноженной на 1.65, то есть на величину Z-значения, формирующего 95%-ный односторонний доверительный интервал для величины  $\Delta RE$ .

После описания типов аналитических ошибок, остановимся на механизмах или экспериментах, при помощи которых можно определить величины этих ошибок.

## **Верификационные эксперименты по подтверждению величин аналитических ошибок.**

Смысл верификационных экспериментов заключается в подтверждении спецификаций производителя по прецизионности (случайной ошибке) и правильности (систематической ошибке), метода или аналитической системы, то есть их аналитической эффективности. Методология проведения этих экспериментов прекрасно изложена в документе Института Клинических Лабораторных Стандартов (Clinical and Laboratory Standards Institute) EP 15-A3[2]. Следует отметить, что документ предлагает проводить верификацию систематической ошибки метода на основании оценки величины его правильности (trueness), нового термина, рекомендуемого документом ISO 15189 (см.ниже).

## **Валидационные эксперименты по определению величин аналитических ошибок.**

Как продемонстрировано в таблице 1 для определения величин различных типов аналитических ошибок существуют различные валидационные эксперименты. В первом столбце таблицы представлены типы ошибок. Второй столбец таблицы назван “предварительные”, так как указанные в нем валидационные эксперименты легче выполнять и их выполнение занимает меньше времени и усилий, чем “финальных” экспериментов. Финальные эксперименты требовательней, их выполняют только тогда, когда результаты предварительных экспериментов продемонстрируют приемлемую аналитическую эффективность метода. В то же время, “плохие” результаты предварительных экспериментов являются объективным основанием для прекращения процесса валидации метода и его отклонения.

*Таблица 1 : Валидационные эксперименты для определения типов аналитических ошибок.*

<b>Тип Аналитической Ошибки</b>	<b>Валидационные Эксперименты</b>	
	<b>Предварительные</b>	<b>Финальные</b>
Случайная Ошибка.	Внутрисерийный Репликационный эксперимент.	Межсерийный Репликационный эксперимент.
Константная Ошибка.	Интерференция.	Сравнительный Анализ Методов.
Пропорциональная Ошибка.	Эксперимент на открытие.	

- **Репликационные эксперименты** дают информацию о величине случайной ошибки метода. Их выполняют путем повторного определения концентрации аналита(ов) одного и того же образца в течение определённого периода времени, как правило в течение аналитической серии, рабочего дня или в течение рабочего месяца. Цель предварительного репликационного эксперимента заключается в определении величины внутрисерийной случайной ошибки, путем определения величины внутрисерийной непрецизионности. Как правило, его выполняют в течение одного рабочего дня. Цель же финального репликационного эксперимента состоит в определении величины общей случайной ошибки, путем определения величины общей (промежуточной) непрецизионности, которая включает в себя её внутрисерийный и межсерийный компоненты. Как правило, его выполняют в течение двадцати рабочих дней.

**Эксперимент по определению интерференции** дает информацию о величине константной систематической ошибки, возникающей вследствие недостаточной аналитической специфичности метода, а также о вызывающих её причинах, то есть интерферентах. В процессе постановки эксперимента, один экспериментальный образец готовят при помощи добавления подозреваемого в интерференции вещества (билирубин, липиды, гемоглобин) в пробу, содержащую аналит. Вторую аликвоту, оригинальную пробу, разводят в том же соотношении дилюентом, после чего анализируют обе пробы при помощи тестируемого метода и определяют разницу в полученных результатах.

- **Эксперимент на открытие (Recovery experiment)** дает информацию о величине пропорциональной систематической ошибки, возникающей вследствие конкурентной реакции элементов матрикса пробы с аналитом. Тестируемую пробу готовят при помощи добавления стандартного раствора тестируемого аналита к аликвоте пробы пациента. Базовую пробу готовят путем добавления эквивалентного количества дилюента, используемого для приготовления стандартного раствора аналита, во вторую аликвоту того же самого образца пациента. Затем обе пробы анализируют при помощи тестируемого метода и сравнивают концентрации открытого аналита в тестируемой пробе с его концентрацией в базовой пробе.
- **Эксперимент по сравнительному анализу методов** дает информацию о средней величине систематической ошибки, наблюдаемой при исследовании проб пациентов. Кроме того, при помощи этого метода можно определить природу этой ошибки, то есть среднюю величину константного или пропорционального её компонентов. В ходе проведения эксперимента собирают серию проб пациентов и анализируют их при помощи тестируемого метода и метода сравнения, после чего сравнивают полученные результаты с целью определения различий между ними и, следовательно, между методами или аналитическими системами. Величина этих различий определяет величину систематической ошибки между ними.

Подробную методологию выполнения валидационных экспериментов можно найти в соответствующих документах Института Клинических Лабораторных Стандартов (Clinical and Laboratory Standards Institute)[3-5].

## **Текущие и будущие тенденции: Неопределённость или Общая Ошибка?**

Вследствие усилий по внедрению международной стандартизации в практику медицинских лабораторий изменяются традиционные термины и концепции. В своем документе **15189 – Медицинские Лаборатории – Частные требования к качеству и компетентности (15189 - Medical Laboratories -Particular requirements for quality and competence)**[1] Международная Организация по Стандартизации (ISO (International Standards Organization)), дает рекомендации по организации работы медицинских лабораторий. В этом документе в разделе 5.5.1.3. Валидация аналитических процедур (Validation of examination procedures) дается перечень аналитических характеристик метода или аналитической системы, валидацию которых должна выполнить медицинская лаборатория. В этом списке есть новые термины **правильность (trueness)** и **неопределённость (uncertainty)**. В разделе документа 5.3.1.4 Калибровка оборудования и метрологическая прослеживаемость (Equipment calibration and metrological traceability) есть новый термин **метрологическая прослеживаемость (metrological traceability)**. Все эти термины ранее



использовались только в Калибровочных Лабораториях и, следовательно, в соответствующем документе ISO приводилось их разъяснение (ISO 17025 Общие Требования к Компетенции Испытательных и Калибровочных Лабораторий (ISO 17025 General Requirements for the Competence of Testing and Calibration Laboratories)). С разработкой документа ISO 15189 концепции и терминология, которые применялись для калибровочных лабораторий, то есть метрологических лабораторий, стали применяться и для медицинских лабораторий. Иными словами, документ ISO 15189 вводит медицинские лаборатории в мир метрологии, где заказчикам (врачам, пациентам) предлагается продукт (результат теста) с приписанными целевыми значениями вместе с неопределённостью (“сомнением”) в отношении этого целевого значения. Действительно, сегодня медицинские лаборатории все больше и больше начинают работать в координатах правильности (trueness) и неопределённости измерений (measurement uncertainty). Это подтверждают даже названия некоторых, относительно недавно разработанных документов CLSI [2], [6-7], а также введённая в 2007 году Федеральным Медицинским Советом (РiлВÄК) новая метрика оценки аналитического качества (RMSD) для немецких лабораторий (см.ниже). В связи с этим, попробуем описать концепцию неопределённости и сравнить её с концепцией общей ошибки.

### ❖ ISO, VIM и GUM: Основные метрологические термины и их определение.

Начнем с определений новых терминов. Итак, в соответствии с Международным Словарем Основных и Общих Терминов в Метрологии (International Vocabulary of Basic and General Terms in Metrology (VIM)) [8] и Руководству по Выражению Неопределённости измерения (Guide to the expression of Uncertainty in Measurement(GUM)) [9]:

- **Правильность (Trueness) измерений:** степень близости между средним значением, полученным на основании большой серии измерений и **истинным значением**.
- **Точность (Accuracy) измерений:** степень близости между результатом измерения мезюранда и его истинным значением.
- **Неопределенность измерения (Uncertainty of measurement(UM)):** параметр, ассоциированный с результатом измерения, характеризующий дисперсию величин, которую можно было бы обоснованно приписывать к мезюранду;
  - **Стандартная неопределённость (Standard uncertainty(SU)):** неопределённость результатов измерения, выраженная в виде стандартного отклонения.
  - **Комбинированная стандартная неопределённость (Combined standard uncertainty(CSU)):** стандартная неопределённость результата измерения, используется тогда, когда этот результат получен на основании значений ряда других величин и равна положительному квадратному корню суммы членов, которые представляют собой дисперсии или ковариации этих других величин, взвешенных в соответствии с тем, как результат измерения изменяется с изменением этих величин.
  - **Расширенная неопределённость (Expanded Combined Uncertainty(ECU)):** величина, определяющая интервал вокруг ожидаемого результата измерения, охватывает большую часть (фракцию) распределения значений величины, которые могли бы быть обоснованно приписаны к мезюранду.
- **Метрологическая прослеживаемость (Metrological traceability):** свойство результатов измерения соотноситься с результатом установленного метрологического эталона



через непрерывную цепь калибровок измерительных систем или сравнений, в каждом случае с установленной неопределённостью измерения.

- **Цепочка метрологической прослеживаемости (metrological traceability chain):** цепочка различных измерительных систем, связанных с процедурами измерения и стандартами, от измеренных результатов к установленному метрологическому эталону.
- **Цепочка метрологической прослеживаемости** определяется иерархией калибровок от измеренного результата к установленному метрологическому эталону.

#### ❖ ISO, VIM и GUM: Объяснение терминов и концепций.

Фундаментальный принцип заключается в том, что “правильность” можно определить только через “прослеживаемость”, каркасом которой являются референтные методы и референтные\калибровочные материалы. При помощи цепочки метрологической прослеживаемости результат измерения, выполненный при помощи рутинного лабораторного метода, связывается с его истинным значением (Рис.7). Таким образом, цепочка метрологической прослеживаемости обеспечивает стандартизацию результатов исследования анализа, выполненных при помощи разных методов или аналитических систем. Каждая процедура приписывания значений в цепочке метрологической прослеживаемости ассоциирована с неопределённостью измерения, при этом величина комбинированной и расширенной неопределённости возрастает.

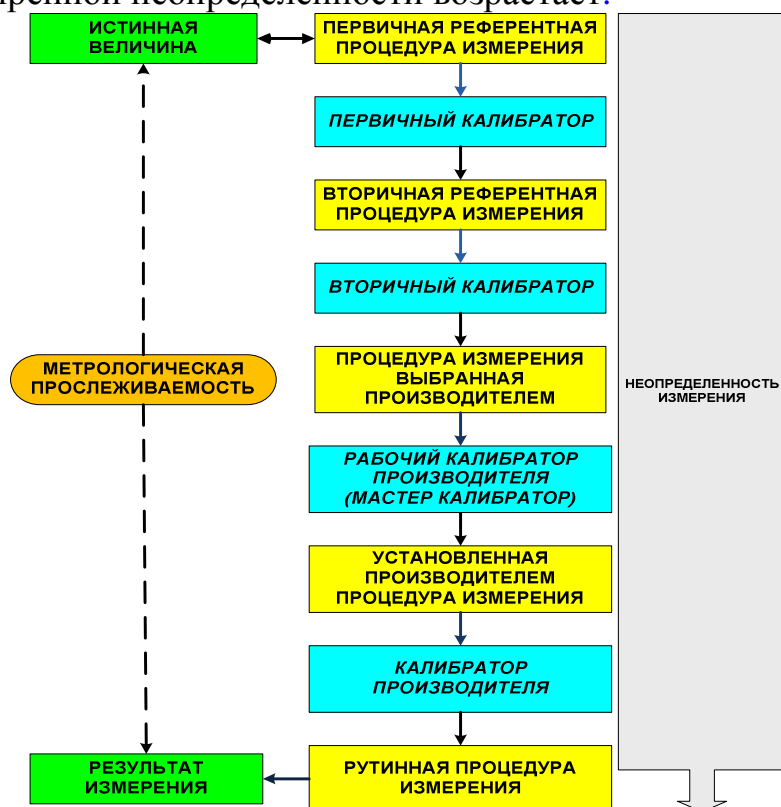


Рисунок 7. Упрощенная схема цепочки метрологической прослеживаемости. Каркасом цепочки являются референтные процедуры и материалы\калибраторы. Калибровки выделены черными стрелками, а процедуры приписывания значений – синими. Связывает истинную концентрацию мезюранда с его концентрацией, измеренной при помощи рутинных лабораторных методов. Каждая процедура приписывания значений в цепочке метрологической прослеживаемости ассоциирована с неопределённостью соответствующего измерения.

Как показано на рисунке 8, используя такую систему референтных методов и материалов, можно определить “прослеживаемую величину”(“traceable value”). В настоящее

время, термин “прослеживаемая величина” (“traceable value”) заменяет более ранний термин “истинная величина” (“true value”). Её величина не может быть точно известна и её правильность должна описываться в терминах “неопределённости измерения”.

Ранее, концепция точности была связана только с систематической ошибкой. Сейчас, в новом мире ISO и GUM, точность отражает ошибку отдельного результата измерения мезюранда (аналита) и включает в себя как систематический, так и случайный её компоненты, а с систематической ошибкой ассоциирован термин “правильность”.

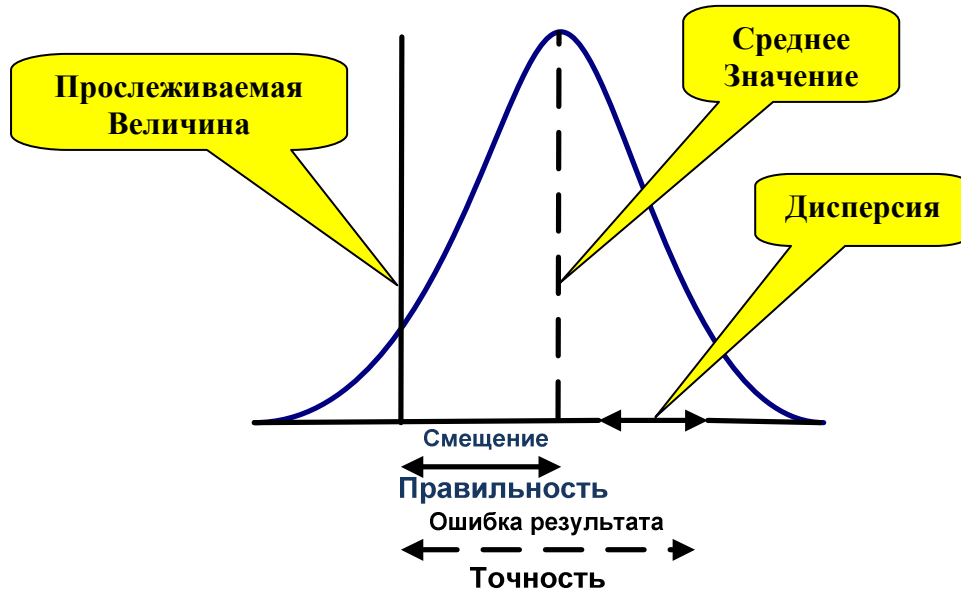


Рисунок 8. Правильность и Точность в соответствии с ISO и GUM. В отличие от традиционной (старой) концепции, когда точность была связана только с систематической ошибкой, сейчас, концепция точности включает в себя как систематический, так и случайный компонент ошибки измерения, а с систематической ошибкой ассоциирован термин “правильность”.

На рисунках 9 и 11 продемонстрирована концепция и терминология, связанная с неопределённостью измерения (UM). Оценку величины неопределённости какого-либо одного компонента измерительного процесса выполняют при помощи определения его стандартного отклонения (Стандартная Неопределённость). Оценку нескольких компонентов процесса измерения определяют посредством суммирования их дисперсий и извлечения квадратного корня из этой объединенной (комбинированной) дисперсии (Комбинированная Стандартная Неопределённость). Эти компоненты дисперсии можно оценить экспериментально (так называемая неопределённость типа А) или теоретически (так называемая неопределённость типа В). Наконец, неопределённость можно выразить в виде доверительного интервала с установленным фактором покрытия (Расширенная Неопределённость или Расширенная Комбинированная Неопределённость с фактором покрытия 2 (95% доверительный интервал), либо 3 (99% доверительный интервал)). Именно величина Расширенной Неопределённости измерения приписывается к мезюранду.

Концепция Неопределённости Измерения сфокусирована на определении дисперсии результата измерения концентрации аналита, которая могла бы быть получена в ситуации, когда проводили бы его многократные измерения. Для этого используют доступные данные по дисперсии различных источников неопределённости измерения связанных с данной измерительной системой для определения интервала значений, в пределах которого с установленным уровнем достоверности будет находиться его прослеживаемое значение. Например, если результат измерения концентрации глюкозы в пробе сыворотки крови будет равен 5,5 ммоль/л, соответствующая информация по ECU может обеспечить

95% достоверность в том, что прослеживаемое значение концентрации глюкозы в этой пробе будет находиться в интервале от 5,3 до 5,7 ммоль\л.

Обобщая, UM не оценивает ошибку, а обеспечивает с установленным уровнем достоверности интервал, где будет находиться прослеживаемое значение концентрации исследуемого аналита. Таким образом, UM - это количественный показатель уровня доверия лаборатории относительно качества результата. Иными словами, UM представляет собой существенный параметр достоверности результатов изменения концентрации аналита. Основой количественной оценки UM является SD. И еще: UM является свойством результата измерения, но не процедуры измерения, при помощи которой он был получен.

Теоретически, в связи с тем, что при расчете UM работает метрологическая прослеживаемость, величина аналитического смещения должна быть элиминирована, например, при помощи повторной калибровки. Тем не менее, на практике это не всегда имеет место. Тогда возникает вопрос: как при расчете UM выполнить коррекцию аналитического смещения? Ответ в том, что подход UM допускает определение неопределённости величины аналитического смещения как комбинации неопределённости приписанной самому референтному значению и стандартной ошибки среднего значения референтного материала, полученной на основании повторных измерений этого референтного материала в лаборатории при помощи данной процедуры измерения. В связи с этим величина неопределённости смещения выражается в виде SD. На первый взгляд это непонятно, поэтому приведу практический пример. Допустим производитель тест-системы ассигновал величину своего калибратора используя для этого сертифицированный вторичный (матрикс-соответствующий) референтный материал (CRM), который является метрологически прослеживаемым до единиц SI (моль) и предоставил свою спецификацию в которой указал, что калибратор коммутабелен с CRM. Медицинская лаборатория купила CRM и выполнила в нем десять репликативных измерений концентрации аналита и получила следующие данные:

	Сертифицированное Значение	1SD	Повторяемость (n=10)	1SD	Аналитическое смещение (Bias)
CRM	5,0 ± 0,1 ммоль\л (95% ДИ)	0,05 ммоль\л	Среднее = 4,5 ± 0,16 ммоль\л (95% ДИ)	0,08 ммоль\л	0,25 ммоль\л

○ **Определение неопределённости смещения ( $U_{Bias}$ )** проводят в два этапа.

➤ Сначала рассчитывают неопределённость среднего значения CRM, полученное в лаборатории, которая равна величине стандартной ошибки среднего:

$$UMV = 0,08/\sqrt{10} = 0,025 \text{ ммоль\л.}$$

➤ После этого рассчитывают величину неопределённости смещения:

$$U_{Bias} = \sqrt{(0,05)^2 + (0,025)^2}; U_{Bias} = 0,0559 \text{ ммоль\л.}$$

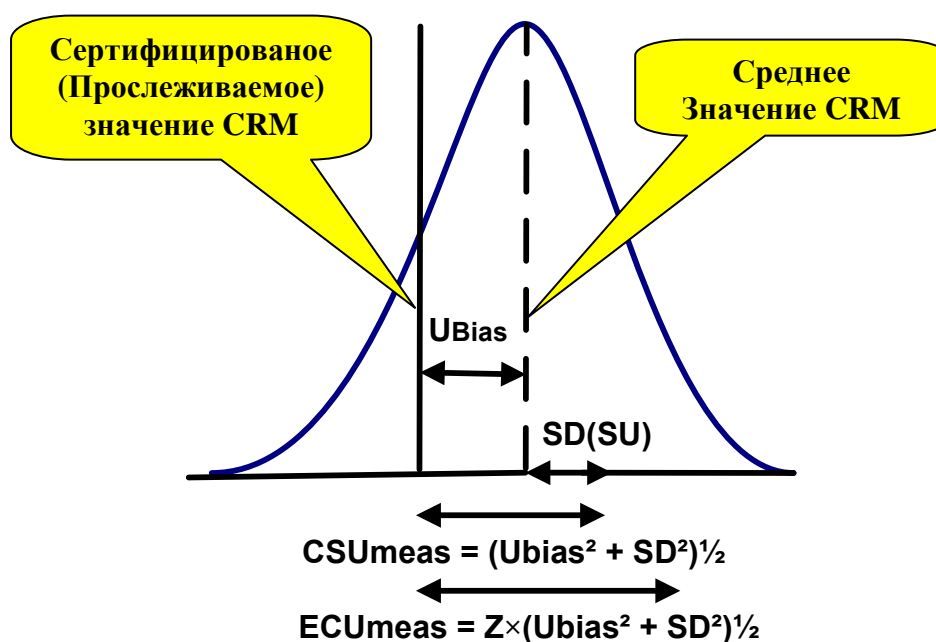
○ **Определение Комбинированной Стандартной Неопределённости измерения ( $CSU_{meas}$ )** проводят как комбинацию величин  $U_{Bias}$  и промежуточной неприцизионности измерения ( $SU$ ). Здесь следует отметить, что медицинские лаборатории часто игнорируют величину  $U_{Bias}$ , если она не превышает величину произвольной точки отсечки, равной 20-30% величины промежуточной неприцизионности. В нашем примере, величина промежуточной неприцизионности измерения составила 0,09 ммоль\л для первого уровня контрольного материала ( $(0,0559/0,09) \times 100 = 62\%$ ) и 0,07 для второго уровня контрольного материала ( $(0,0559/0,07) \times 100 = 79,8\%$ ). То есть, величины  $U_{Bias}$  превышают произвольную точку отсечки, поэтому:

- $CSU_{meas} = (0,0559^2 + 0,09^2)^{1/2} = 0,1$  ммоль\л; для первого уровня КМ;
- $CSU_{meas} = (0,0559^2 + 0,07^2)^{1/2} = 0,09$  ммоль\л; для второго уровня КМ.

○ **Определение Расширенной Комбинированной Неопределённости измерения ( $ECU_{meas}$ )** определяют путем умножения  $CSU_{meas}$  на фактор покрытия, чаще всего 2, обеспечивая её 95%-ный доверительный интервал. Тогда:

- $ECU_{meas} = 0,1$  ммоль\л  $\times 2 = 0,2$  ммоль\л ; для первого уровня КМ;
- $ECU_{meas} = 0,09$  ммоль\л  $\times 2 = 0,18$  ммоль\л; для второго уровня КМ.

Тогда, например, если результат измерения концентрации аналита составит 5 ммоль\л, то есть его концентрация будет близка к области концентраций первого уровня контрольного материала, можно с 95% вероятностью сделать заключение, что прослеживаемое значение концентрации аналита в этой пробе будет находиться в интервале от 4,8 до 5,2 ммоль\л.



*Рисунок 9. Концепция неопределенности измерения. Основой количественной оценки неопределённости измерения является стандартное отклонение (SD), обеспечивающее расчет стандартной неопределённости (SU) измерения. Комбинирование величин стандартных неопределённостей выполняют путем сложения дисперсий, то есть суммирования квадратов стандартных отклонений, получая величину Комбинированной Стандартной Неопределённости измерения (CSU). Расширенную Комбинированную Неопределённость (ECU) определяют путем умножения комбинированной стандартной неопределённости на фактор покрытия (Z), величина которого может обеспечить либо 95%, либо 99% доверительный интервал.*

Подход, описанный выше, называется “движение сверху вниз”. В соответствии с ним, известная величина смещения должна быть элиминирована или минимизирована, а её остаточную величину, оцененную в терминах неопределённости величины смещения, используют для повторной калибровки или коррекции результатов. Часто, величина неопределённости смещения минимальна по сравнению с величиной неопределённости и, поэтому, игнорируется. В связи с этим, общая (промежуточная) величина неопределённости, полученная в результате проведения КК, отражает общую неопределённость результатов измерения. Этот подход можно использовать в медицинских лабораториях, но

он в силу объективных причин сопряжен с определёнными практическими трудностями (см.ниже).

Подход “движение снизу вверх” основан на оценке неопределённостей, ассоциированных с отдельными компонентами измерительной системы и комбинировании их в модель, отражающую их воздействие на всю систему измерения. Этот подход лучше всего подходит для нужд производителей изделий для IVD, а именно для валидации новой процедуры измерения или поиска технических этапов, где можно снизить величину UN. Подходу “движение снизу вверх” посвящен недавно вышедший документ CLSI [6].

### ❖ **RMSD - новая метрика оценки внутрилабораторного качества от RiliBÄK.**

В 2007 году, RiliBÄK (Руководство ("Rili") Федерального Медицинского Совета Германии (BÄK)) представил новую метрику оценки внутрилабораторного качества для немецких лабораторий: Корень Среднеквадратичного Отклонения (Root Mean Square Deviation), которая выражается в виде процента к целевому значению (%RMSD). По сути, эта метрика выражения Общей Ошибки определения аналита способом, схожим с тем, который рекомендован ISO для расчета величины неопределённости:

$$\%RMSD = \frac{\sqrt{k^2 (SD_{cc}^2) + Bias^2}}{TV}$$

где:

$SD_{cc}$  = стандартное отклонение.

$Bias$  = разница, наблюдаемая между Средним Значением и Целевым Значением ( $TV$ ).

$k$  = фактор покрытия.

$TV$  = Целевое Значение (*Target Value*) для контрольной пробы (от производителя).

Как видно из формулы определения RMSD в качестве ”опорной” величины для расчета внутрилабораторного аналитического смещения RiliBÄK рекомендует использовать Целевое Значение контрольного материала, приписанное производителем (*Target Value*). Это говорит о том, что RiliBÄK налагает достаточно жесткие требования производителям контрольных и калибровочных материалов в плане обеспечения метрологической прослеживаемости их продукции. Коррекцию аналитического смещения RiliBÄK рекомендует проводить, возводя его значение в квадрат, после чего, суммируя его с величиной дисперсии. Методически это не совсем правильно, поскольку смещение не является стандартным отклонением и поэтому, возведя его в квадрат мы не получим дисперсию. Более правильно использовать подход со стандартной ошибкой среднего и неопределённостью сертифицированного референтного материала (CRM), определяя неопределённость аналитического смещения (см.выше), но для медицинских лабораторий CRM не всегда может быть доступен.

Как указывалось выше, внутрилабораторная оценка качества медицинских лабораторий Германии основывается на величинах RMSD, которое, либо указано в соответствующем списке аналитов, либо его нужно определить самостоятельно в лаборатории, используя фактор покрытия 3 (Спецификация, Наблюдаемая в Лаборатории (*Lab Observed Specification*)). Этот расчет лаборатория должна проводить на основании данных одного внутрилабораторного контрольного цикла, который продолжается как минимум до выполнения 15 контрольных измерений, в целом около месяца, но не более 3 месяцев.



Как только немецкая лаборатория определяет цель в виде %RMSD, она начинает проводить внутрилабораторную оценку качества, создавая контрольные карты на основании величины Предельной Ошибки (Error Margin, ЦЗ ± Целевой %RMSD). Эта оценка проводится в течение двух временных периодов. Во-первых, в лаборатории должна выполняться ежедневная оценка контрольных результатов, значения которых не должны превышать величину соответствующей Предельной Ошибки. Во-вторых, лаборатории следует выполнить ретроспективную оценку результатов Контрольного Цикла. Для этого, для каждого уровня контрольного материала, на основании данных ежедневного контроля качества, полученных на протяжении всего Контрольного Цикла, рассчитывают среднее значение, стандартное отклонение (SD), аналитическое смещение (bias) и % RMSD. Рассчитанные % RMSD не должны превышать соответствующие целевые величины % RMSD. Если результаты ежедневного контроля качества и особенно ретроспективной оценки его результатов будут превышать целевые % RMSD, то тогда немецкая лаборатория должна выполнить корректирующие мероприятия, вплоть до повторной валидации метода или аналитической системы.

## ❖ Неопределённость или Общая Ошибка: Что лучше?

Исторически, в лабораторной медицине, концепция Общей Ошибки была предшественницей концепции неопределённости измерения. В связи с этим, теоретические аспекты концепции общей ошибки уступают таковым концепции неопределённости измерения, но практические аспекты общей ошибки, несомненно, лучше. Итак:

### ○ Теория:

Теоретически, концепция Общей Ошибки определения аналита (“Total Error”(TE)) построена таким образом, что:

1. Не может эффективно применяться к отдельному результату измерения аналита, поскольку строительные блоки концепции TE не могут точно отражать ошибку отдельного результата измерения, поскольку (Рис.10):
  - Невозможно точно определить величину аналитического смещения. В самом деле, значение мезюранда, приписанное к сертифицированному референтному материалу являются неточными и выражаются в форме единиц  $x \pm y$ . Кроме того, среднее значение, рассчитанное для референтного материала на основании данных репликативных измерений при помощи рутинного метода, имеет неизбежную непрецизионность (стандартную ошибку среднего), поэтому среднее значение, используемое для рутинной процедуры измерения, также выражается в форме единиц  $x \pm y$ . **В связи с тем, что обе величины, используемые для расчета аналитического смещения, имеют неопределённость, величина аналитического смещения не может быть точно известна.** Аналогично, ведет себя и случайная ошибка, поэтому, отдельный результат измерения не может быть точно известен. **В связи с этим, и общая ошибка результата измерения не может быть точно известна.** Это неизбежно даже тогда, когда применяется лучшая (state-of-the-art) референтная процедура измерения. Именно поэтому, значения мезюранда, приписанные к сертифицированному референтному материалу выражают в виде единиц  $x \pm y$ .

- Для определения верхней вероятной величины ошибки измерительной системы, при расчете величины  $aTE$  к величине смещения (Bias) прибавляют одностороннюю 95% вероятность распределения случайной ошибки  $\Delta RE$  ( $Z = 1,65$ ). Во-первых, с точки зрения математической статистики, это, примерно, то же самое, что суммировать груши и яблоки. Во-вторых, как нам известно из анализа результатов внутрилабораторного КК, отдельные результаты одного и того же контрольного материала содержат различные величины случайной ошибки ( $\Delta RE$ ), поэтому мы и фиксируем дисперсию этих результатов. Вот поэтому концепция  $TE$  не может применяться для отдельного результата тестирования, поскольку она не допускает вероятности того, что у какого-либо отдельного результата пациента будет непрецизионность отличная от  $1,65 \times SD$ .

2. Поскольку модель допустимой Общей Ошибки не связана с концепцией метрологической прослеживаемости, теоретически, при помощи её невозможно обеспечить стандартизацию результатов исследования какого-либо аналита, а только гармонизацию.

В итоге, хотя ошибка измерения не может быть точно известна, Концепция Общей Ошибки представляет собой такую теоретическую величину, которая на практике полезна в ситуациях, когда требуется установить приемлемую верхнюю границу общей ошибки для измеряемых результатов, то есть, например, измерении концентрации холестерина в плазме крови. Этот подход не подходит, если рассматривают неопределённость отдельных результатов измерения.

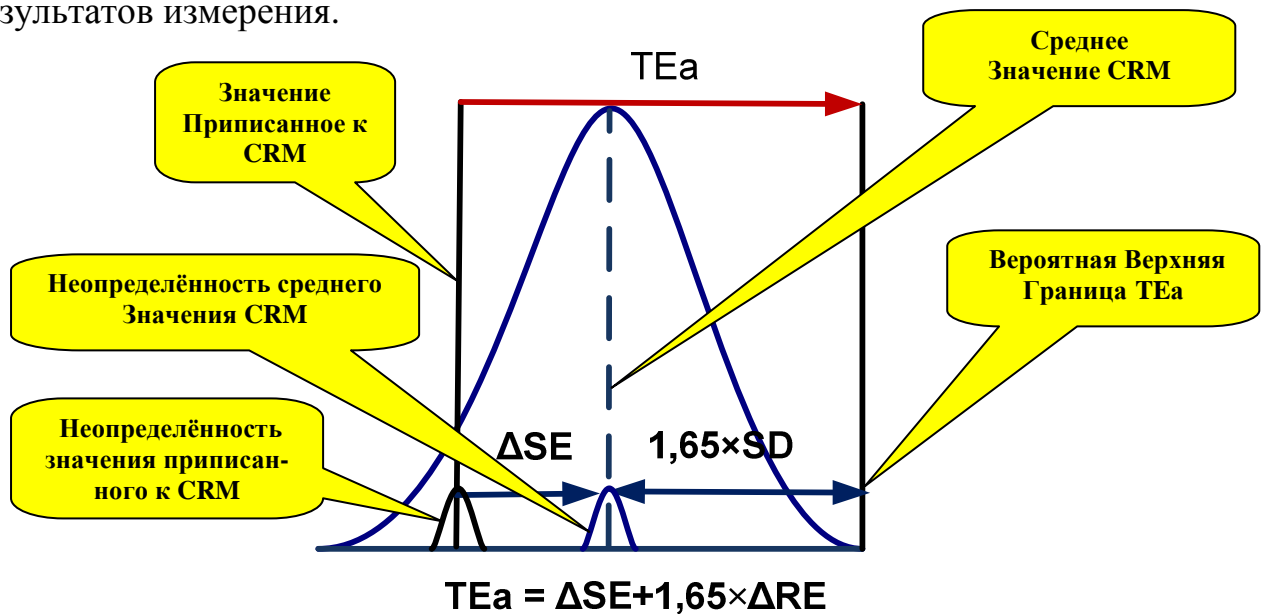


Рисунок 10. Ограничения Концепции допустимой Общей Ошибки измерения аналита. В связи с тем, что строительные блоки концепции не могут быть точно определены, она не может эффективно применяться для характеристики отдельного результата измерения.

Как указывалось выше, в отличие от концепции Общей Ошибки, теория концепции неопределённости измерения аналита (UM), а также связанная с ней концепция метрологической прослеживаемости не рассматривают оценку ошибки измерения аналита. Концепция UM сфокусирована на определении суммарной дисперсии результата, которая могла бы быть получена, если бы вместо однократного проводили бы многократные измерения концентрации аналита. Для определения этой суммарной дисперсии используют дисперсии различных источников неопределённости измерения связанных с конкретной измерительной системой для определения интервала значений, в пределах которого с установленным уровнем достоверности будет находиться его прослеживаемое значение.



Опять же, УМ не оценивает ошибку, а обеспечивает с установленным уровнем достоверности интервал, где будет находиться прослеживаемое значение концентрации исследуемого анализа (Рис.11). Таким образом, в действительности, УМ - это количественный показатель уровня доверия или убеждения лаборатории относительно качества результата. Иными словами, концепция УМ является вероятностным параметром достоверности отдельного результата измерения концентрации анализа и поэтому, в отличие от концепции ТЕа, она может использоваться для оценки неопределённости отдельного результата исследования его концентрации анализа.

УМ является свойством результата измерения, но не свойством самой аналитической процедуры измерения, при помощи которой он был получен, поскольку теоретически, считается, что неопределённость всех процедур измерения данного анализа будет стандартизирована. Опять же, теоретически считается, что такая стандартизация должна обеспечиваться стандартной цепочкой метрологической прослеживаемости. Как видим, теоретически, концепция неопределённости и прослеживаемости измерения гораздо эффективнее, чем концепция Общей Ошибки определения анализа, поскольку может применяться к отдельным результатам измерения анализа, а также позволяет стандартизировать результаты измерения посредством цепочки метрологической прослеживаемости.

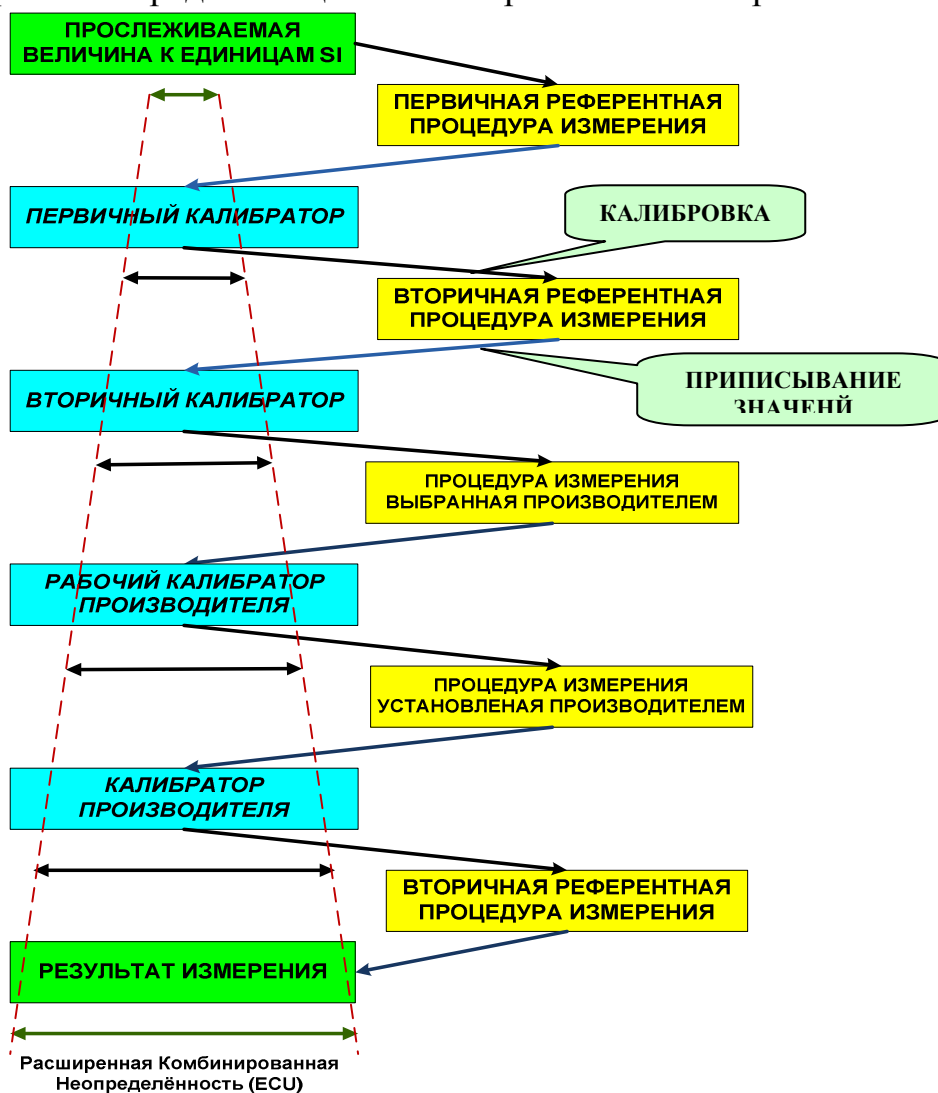


Рисунок 11. Неопределённость измерения и ее фокусировка на определении суммарной дисперсии результата измерения и как следствие этого на ECU. По мере приписывания значений референтным материалам\калибраторам возрастает величина неопределенности измерения, поскольку прибавляется её источник. Метрологическая прослеживаемость обеспечивает стандартизацию результатов тестирования.

## ○ Практика

Сравнить практическую ценность концепций Общей Ошибки определения аналита (“Total Error”(TE)) и Неопределённости Измерения (Uncertainty of measurement (UM)) можно в ракурсе возможности их использования в создании и управлении аналитическим качеством в рамках системы TQM и в частности в процессе планирования качества (QP).

➤ **Стандарты (цели) качества (QS).** На сегодняшний день, все цели качества разработаны на основе концепции допустимой Общей Ошибки определения аналита (“allowable total analytical error” (aTE)). Достаточно вспомнить действующую иерархию спецификаций целей аналитического качества, принятую на международной Стокгольмской консенсусной конференции 1999 года. Относительно концепции UM, можно сказать, что ее сторонники рекомендуют установить целевые значения величин неопределённости измерения аналитов на основе их приемлемости для клиники в самой медицинской лаборатории, что совсем не просто. В противном случае, практическое применение этой концепции не имеет смысла.

### ➤ Планирование Качества (QP).

◆ **Выбор метода или аналитической системы.** В материале “Выбор метода или аналитической системы” мы говорили, что одним из ключевых шагов процесса выбора метода (аналитической системы) является *сравнение желаемых требований к аналитическим характеристикам и заявленных в спецификации производителей методов или аналитических систем*. Сегодня, производители составляют свои спецификации только на основании концепции TE, указывая систематическую ошибку (SE, Bias) в виде уравнения регрессии и случайную ошибку (RE) в виде стандартного отклонения (SD) и\или CV. Иногда, некоторые производители указывают в своих спецификациях прослеживаемость своих калибраторов к референтным методам и\или материалам (единицам SI) и этим все ограничивается. А где же неопределённость измерения?

◆ **Валидация и верификация метода или аналитической системы.** Как указывалось здесь выше в соответствии с требованиями документа ISO 15189, *валидация это - подтверждение путем представления объективных доказательств того, что требования для специфического использования или применения были выполнены*. Мы говорили, что применительно к концепции TE, это означает **определение аналитических ошибок метода и сравнении их с целями качества, которые тоже сформированы на основании TEa, или допустимых SE и RE**. А как выполнить валидацию метода или аналитической системы руководствуясь концепцией UM? Допустим, что мы сможем в лаборатории определить UM какого-либо аналита при помощи нашего валидируемого метода, но тогда с чем ее сравнить, чтобы определить её приемлемость для клиники? Ведь Цели Качества (QS) для UM не сформированы. Теоретически, можно допустить, что лаборатория может заранее сформировать клинические требования к качеству метода на основании мнения клиницистов и UM, но это непростая и достаточно трудоемкая работа, которая не по силам большинству медицинских лабораторий. Опять же, согласно документу ISO 15189, *верификация – это подтверждение путем представления объективных доказательств того, что специфические требования были выполнены*. С точки зрения концепции TE это означает **доказательство заявленных производителем величин ошибок метода, и сравнении их с целями качества, которые тоже**

**сформированы на основании TEa, или допустимых SE и RE.** Тогда снова возникает резонный вопрос: как выполнить верификацию метода или аналитической системы, работая с концепцией UM, если спецификации производителей не содержат UM, а Цели Качества (QS) для UM не сформированы? На мой взгляд, этот вопрос все еще остаётся открытым.

- ◆ **Планирование аналитического качества для метода или аналитической системы.** Ключевым мероприятием планирования аналитического качества является разработка оптимальной, то есть наиболее эффективной процедуры КК, соответствующей спецификации теста и аналитическим характеристикам процесса тестирования метода. В настоящее время, процесс планирования качества детально разработан для модели TE, но не для концепции UM. В самом деле, модель TE можно использовать для создания как аналитической, так и клинической модели планирования качества. [10,11].

В документе ISO 15189 в разделе 5.6.2.1 подчеркивается, что "лаборатория должна разработать процедуры КК, необходимые для верификации достижения желаемого качества результатов". Однако, модель UM не дает каких-либо рекомендаций или инструментов для выполнения этой задачи. Действительно, как нам разработать процедуры КК, если мы не знаем целей качества (QS) в виде целевых UM? Что же касается документа ISO 15189, то с одной стороны он требует выполнять валидацию и верификацию методов и аналитических систем, разрабатывать процедуры КК, а с другой не указывает, как это сделать на практике. Это не удивительно, поскольку документ ISO 15189 является стандартом "высокого уровня" в котором даны только общие рекомендации того, что следует сделать в медицинской лаборатории, но в нем мало говорится о том, как это сделать на практике.

Итак, мы видим, что насколько сильна и гармонична концепция UM теоретически, настолько же она слаба в плане практического применения. На мой взгляд, сегодня, с точки зрения практического применения в медицинских лабораториях, концепция UM сыра и не доработана. В то же время, не смотря на свои теоретические недостатки, концепция TE очень практична, проверена временем, логически встраивается в процесс создания аналитического качества в рамках TQM. В связи с этим не удивительно, что, например, аккредитация медицинских лабораторий США до сих пор основана, прежде всего, на требованиях CLIA, хотя, в то же время, интерес к аккредитации в соответствии с документом ISO 15189 нарастает.

## **Выводы.**

- Валидация и Верификация методов или аналитических систем являются частью процесса планирования качества (QP) системы TQM медицинских лабораторий. Цель валидации и верификации метода или аналитической системы заключается, соответственно, в определении или доказательстве величин ошибок метода или аналитической системы, сравнении этих ошибок с ошибками целей качества, и на основании этого определении их приемлемости для клиники.
- Существует два основных типа аналитических ошибок: случайная ошибка (Random Error (RE)) и систематическая ошибка (Systematic Error (SE)). Систематическая ошибка состоит из константной (Constant Systematic Error) и пропорциональной (Proportional Systematic Error) ошибок. Кроме того, на основании эффектов RE и SE, определяют общую ошибку (Total Error (TE)).

- Определение и доказательство (подтверждение) аналитических ошибок выполняют при помощи соответствующих валидационных и верификационных экспериментов.
- Усилия по внедрению глобальной стандартизации в практику медицинских лабораторий привели к пересмотру традиционных терминов и концепций. Тем не менее, хотя теоретически, новая концепция неопределённости измерения может обеспечить индивидуализацию оценки результатов тестирования, с практической точки зрения она все еще не доработана и её использование в создании и управлении аналитическим качеством остается спорным.
- Сегодня разумно, чтобы новые концепции правильности и метрологической прослеживаемости были в зоне ответственности производителей калибраторов и аналитических систем, поскольку немногие медицинские лаборатории имеют соответствующие ресурсы и время для выполнения этой работы.
- В настоящее время, на мой взгляд, наиболее грамотный подход практической реализации концепций метрологической прослеживаемости и неопределённости в медицинских лабораториях обеспечивает Руководство Федерального Медицинского Совета Германии (RiBÄK), используя для этого новую метрику внутрилабораторной оценки качества RMSD. Поскольку с одной стороны, руководство дает целевые значения величин RMSD, а с другой подробно описывает, как нужно использовать их для внутрилабораторного контроля качества.

## Литература.

1. ISO15189 Medical laboratories - Particular requirements for quality and competence. 2003. International Organization for Standards, Geneva Switz. Thrid edition 2012-11-01.
2. CLSI. EP15-A3. User Verification of Performance for Precision and Trueness. Clinical Laboratory Standards Institute, Wayne, PA, 2014.
3. CLSI EP5-A3. Evaluation of Precision Performance of Quantitative Measurement Methods. Clinical and Laboratory Standards Institute, Wayne, PA 2014.
4. CLSI EP9-A3. Method Comparison and Bias Estimation using Patient Samples. Clinical and Laboratory Standards Institute, Wayne, PA, 2014.
5. CLSI EP7-A2. Interference Testing in Clinical Chemistry. Clinical and Laboratory Standards Institute, Wayne, PA, 2005.
6. CLSI C51-A. Expression of Measurement Uncertainty in Laboratory Medicine. Clinical and Laboratory Standards Institute, Wayne, PA, 2012.
7. CLSI X5-R. Metrological Traceability and its Implementation; A Report. Clinical and Laboratory Standards Institute, Wayne, PA, 2006.
8. International Vocabulary of basic and general terms in metrology (VIM). ISO, Geneva, 2004
9. GUM. Guide to the expression of uncertainty in measurement. ISO, Geneva, 1995.
10. Westgard JO. Internal quality control: planning and implementation strategies. Ann.Clin.Diochem. 2003;40:593-611.
11. CLSI C24-A3. Statistical Quality Control for Quantitative Measurement Procedures: Principles and Definitions; Approved Guideline - Third Edition. Clinical Laboratory Standards Institute, Wayne, PA, 2006.