

Специфика и стратегия использования тестов для определения антител к SARS-CoV-2

М.И. Прищепа, к.т.н., ЗАО «АНАЛИТИКА»

Преамбула. Прежде всего, хотелось обратить внимание читателей этого обзора, особенно тех, кто профессионально занимается лабораторными исследованиями или использует их результаты в своей работе, что вследствие природы измерений, нет и не может быть таких лабораторных тестов, с помощью которых можно было бы получать результаты со 100% точностью или со 100% надежностью. И дело здесь не только в том, что полезные измерительные сигналы всегда сопровождаются шумами, искажающими в конечном итоге результаты измерений любых физических величин. Что приводит, например, к отклонению результатов измерений уровней аналитов от их истинных значений, и тем больше, чем больше уровень шума по отношению к сигналу и/или чем ближе измеряемый уровень аналита к порогу аналитической чувствительности. Дело также в том, что для большинства количественных и полуколичественных тестов отсутствует адекватное метрологическое обеспечение. В том числе отсутствуют референтные и рабочие эталоны как в виде высокостабильных и высокоточных стандартных образцов состава и свойств биопроб, необходимых для метрологической аттестации аналитических систем в местах их эксплуатации, так и в виде мер аналитов, необходимых для калибровки этих систем. В этой связи, метрологическая прослеживаемость (и сопоставимость) результатов, получаемых количественными тестами, не может быть в настоящее время обеспечена стандартными способами, изложенными в документах ИСО или в национальных стандартах по обеспечению единства измерений, не только на 100% уровне точности, но порой и на уровне, требуемом для решения типовых клинических задач.

В основном по тем же причинам, изложенным выше, невозможно обеспечить 100% достоверность отрицательных или положительных результатов, получаемых полуколичественными тестами. Среди результатов анализа проб пациентов, получаемых при использовании любого полуколичественного теста, всегда будет какое-то количество и ложноположительных, и ложноотрицательных результатов. При прочих равных условиях доля таких результатов будет меньше у тех тестов, у которых реальные (истинные) значения для их чувствительности и специфичности (см. ниже) ближе к 100%. На практике, из-за отсутствия у испытателей времени и должного по структуре и объему пула биопроб, вместо истинных значений определяют оценочные значения для этих диагностических показателей. Сокращая до разумных пределов сроки испытаний, количество тестируемых биопроб, а также упрощая их структуру как по количеству проб с аналитом и без него, так и количеству проб, содержащих возможные вещества-интерференты. Реальные (истинные) значения для чувствительности и специфичности любых тестов не могут быть в принципе равны 100%. В силу природы измерений, как об этом уже упоминалось ранее. Тем не менее, в результате проведения сокращенных испытаний полуколичественного теста можно получать оценочные значения для его чувствительности и/или для специфичности, равные 100%. Например, когда испытатели не ставят перед собой задачу определить максимально точно реальные значения для диагностических показателей исследуемого теста, а ограничиваются предварительной оценкой его диагностических возможностей, оценивая чувствительность и специфичность по результатам, полученным на выборках с небольшим объемом биопроб с упрощенной структурой пула. Обычно в силу нехватки времени и/или средств. В таких случаях, чтобы не вводить в заблуждение будущих пользователей теста, его изготовители и/или испытатели, помимо полученных оценок для чувствительности и специфичности, должны также вычислять и границы доверительных интервалов для их истинных значений. Используя для этого полученные оценочные значения и данные о структуре и количестве протестированных в ходе испытаний биопроб.

Видимо, именно с этой целью в списке [1] серологических тестов для определения антител к SARS-CoV-2, допущенных FDA - Food and Drug Administration к лабораторной практике США, помимо оценочных значений чувствительности и специфичности тестов, полученных в ходе их испытаний, приводятся также границы 95% доверительных интервалов для истинных значений этих диагностических показателей. Такая информация позволяет будущим пользователям тестов узнать,

какова реальная надежность и точность полученных оценочных значений, показывая в каких границах с 95% доверительной вероятностью могут фактически находиться истинные значения тех или иных диагностических показателей теста. Эти границы вычисляются в документе [1] в соответствии с рекомендациями CLSI, изложенными в руководстве EP12-A2 [4]. На базе данных об оценочных значениях для диагностических показателей тестов и о количествах биопроб, использованных для определения этих значений. Следует отметить, что в силу ряда причин авторами этого руководства для вычисления границ доверительных интервалов рекомендуется использовать именно алгоритм Вильсона (Wilson Score), который наиболее надежно определяет такие границы при небольших объемах пула (количествах тестируемых биопроб) и при граничных (близких к 100%) величинах оценочных значений для чувствительности и специфичности. Соответствующие формулы и примеры расчета верхних и нижних границ 95% доверительных интервалов для истинных значений диагностических показателей типовых серологических тестов на антитела к SARS-CoV-2, будут приведены во второй половине этого обзора.

Как видно из описаний эксплуатационных характеристик тестов, изложенных в документе [1], у всех серологических тестов, допущенных FDA к лабораторной практике, истинные значения для чувствительности и специфичности не равны в точности 100% даже в случаях, когда оценочные значения для некоторых из перечисленных тестов оказались равными 100%. Это лишнее показывает, что по результатам тестирования пула биопроб, ограниченного по их количеству и структуре, в принципе, можно получать оценочные значения для чувствительности и/или специфичности равные 100%. Но при произвольной структуре и увеличении количества исследуемых биопроб это становится всё менее и менее возможным. Этот факт очень важно иметь в виду будущим пользователям тестов, поскольку на практике реализуется как раз ситуация, когда лаборатория, используя тот или иной тест, проводит анализ произвольных проб пациентов во всё большем объеме. Вследствие чего всегда следует ожидать, что среди получаемых лабораторией результатов будут встречаться как ложноположительные, так и ложноотрицательные.

Диагностические показатели полуколичественных тестов. В отличие от диагностических возможностей количественных лабораторных тестов, которые описываются аналитическими характеристиками, определяющими в конечном итоге возможные величины отклонений результатов измерений уровней аналитов от их фактических значений, диагностические возможности полуколичественных лабораторных тестов, к которым в том числе относятся серологические тесты, предназначенные для установления наличия или отсутствия антител к SARS-CoV-2 в пробах пациентов, обычно описываются двумя показателями – диагностическими чувствительностью и специфичностью. Величины чувствительности **Sen** (Sensitivity) и специфичности **Spe** (Specificity) целиком определяются устройством самих тестов и характеризуют на вероятностном языке их диагностические возможности по выявлению в пробах пациентов искомого аналита. Помимо этих двух показателей для описания диагностических возможностей полуколичественных тестов используют также прогностические показатели **PPV** (Positive Predictive Value) и **NPV** (Negative Predictive Value), которые характеризуют степень надежности соответственно положительных и отрицательных результатов, получаемых с использованием этих тестов. Величины последних двух показателей, как будет показано ниже, зависят не только от диагностических возможностей используемого теста, но и от доли пациентов тестируемой популяции, в пробах которых искомым аналит реально содержится.

Как было упомянуто ранее в Преамбуле, результаты полуколичественного тестирования биопроб на выявление в них антител к SARS-CoV-2 могут оказаться (см. Таблицу 1) либо истинно положительными (TP – True Positive), либо ложноположительными (FP – False Positive), либо истинно отрицательными (TN – True Negative), либо ложноотрицательными (FN – False Negative). Других вариантов результатов тут быть не может. Конкретная картина распределения результатов тестирования зависит в общем случае от фактического наличия антител в тестируемых пробах, от их уровней и от значений диагностических показателей используемого теста (см. Таблицу 2).

Таблица 1. Возможные результаты тестирования биопроб по выявлению в них антител к SARS-CoV-2 типовым полуколичественным тестом.

Антитела в биопробах ЕСТЬ (+)	Антител в биопробах НЕТ (-)
TP - True Positive (+)	FP - False Positive (+)
FN - False Negative (-)	TN - True Negative (-)

Показатель чувствительности Sen характеризует способность полуколичественных тестов идентифицировать биопробы, которые содержат искомый анализ, в данном случае пробы пациентов с антителами к SARS-CoV-2. Для каждого теста он определяется в процентах путем деления числа проб, которые реально содержат искомый анализ и были выявлены данным тестом в ходе тестирования биопроб выбранного для испытаний пула, на истинное число протестированных проб, реально содержащих искомый анализ. Оценочное значение этого диагностического показателя для каждого теста рассчитывают по следующей формуле:

$$Sen, \% = 100 * TP / (TP + FN), \quad (1)$$

где величина TP (True Positive) – число проб, реально содержащих искомый анализ, в которых его присутствие было выявлено данным тестом, а величина FN (False Negative) - число проб, реально содержащих искомый анализ, но для которых присутствие искомого анализа в них не было выявлено данным тестом. Для типового серологического теста, результаты тестирования которым 1000 биопроб по выявлению в них антител приведены в Таблице 2, показатель чувствительности будет равен:

$$Sen, \% = 100 * 93 / (93 + 7) = 93\%$$

Таблица 2. Характерный пример результатов тестирования 1000 биопроб по выявлению в них антител к SARS-CoV-2 типовым полуколичественным серологическим тестом в ходе его испытаний.

Истинное число биопроб с антителами -100	Истинное число биопроб без антител -900
True Positive (TP) - 93	False Positive (FP) - 10
False Negative (FN) - 7	True Negative (TN) - 890

Показатель специфичности Spe характеризует способность полуколичественных тестов идентифицировать биопробы, которые не содержат искомый анализ, в данном случае пробы пациентов без антител к SARS-CoV-2. Для каждого теста он вычисляется в процентах как отношение числа проб, реально не содержащих искомый анализ, для которых данный тест в ходе тестирования выбранного пула биопроб, показал отсутствие в них искомого анализа, к истинному числу протестированных проб, реально не содержащих искомый анализ. Оценочное значение данного диагностического показателя для каждого теста рассчитывается по следующей формуле:

$$Spe, \% = 100 * TN / (TN + FP), \quad (2)$$

где величина TN (True Negative) – число проб, реально не содержащих искомый анализ, для которых данный тест показал отсутствие в них искомого анализа, а величина FP (False Positive) - число проб, реально не содержащих искомый анализ, но для которых данный тест показал присутствие в них искомого анализа. Для типового серологического теста, результаты тестирования которым 1000 биопроб на наличие в них антител приведены в Таблице 2, показатель специфичности будет равен:

$$Spe, \% = 100 * 890 / (890 + 10) = 98,89\%$$

Любой полуколичественный тест, как было упомянуто выше, всегда будет выдавать какое-то количество ложноположительных и ложноотрицательных результатов в зависимости от его истинных значений для чувствительности Sen и специфичности Spe , которые, в свою очередь, не могут быть равными 100%. Именно значения показателей PPV и NPV для данного теста позволяют тем, кто интерпретирует получаемые результаты тестирования, понимать какова вероятность того, что в пробе

пациента, для которой данный тест дал положительный результат, реально будет содержаться искомый анализ, в нашем случае антитела к SARS-CoV-2, и какова вероятность того, что в пробе пациента, для которой данный тест дал отрицательный результат, реально не будут содержаться искомый анализ, в нашем случае антитела к SARS-CoV-2. Значения показателей PPV и NPV (см. ниже) рассчитываются на основе данных о чувствительности и специфичности тестов, а также данных о доле (Prevalence) представителей тестируемой популяции, которые уже имеют антитела к SARS-CoV-2.

Следует упомянуть, что в документе FDA [1] особо отмечается тот факт, что значения показателей PPV и NPV для конкретного теста достаточно сильно могут зависеть от величины доли представителей обследуемой популяции, которые уже имеют антитела к SARS-CoV-2. В Таблице 3 и на Рисунке 1 представлена динамика изменения значений PPV и NPV для двух типовых серологических тестов на антитела к SARS-CoV-2 – при их отдельном и совместном использовании - по мере увеличения доли (Prevalence) тех, кто в данной популяции уже имеет антитела к SARS-CoV-2. Оценочные значения для чувствительности и специфичности этих тестов приведены в заголовке к Таблице 3. Значения для PPV и NPV были получены с использованием xls-калькулятора FDA [3], ссылка для скачивания которого приведена во введении (About this page) документа [1] в третьей строке снизу. Воспользоваться этим калькулятором (файл Excel) для вычисления конкретных значений PPV и NPV как для оценки работы одиночных тестов, так и для оценки совместной работы двух независимых тестов, может каждый, поскольку и сам документ [1] и калькулятор находятся в свободном доступе. Калькулятор позволяет задавать любые значения для чувствительности и специфичности для каждого из пары используемых тестов, а также задавать любые значения для доли (Prevalence) представителей популяции, которые уже имеют антитела к SARS-CoV-2.

Таблица 3. Динамика изменений показателей PPV и NPV двух типовых серологических тестов, предназначенных для выявления антител к SARS-CoV-2, в зависимости от величины Prevalence (доли популяции, которая уже имеет антитела). Первый тест имеет: чувствительность Sen1 – 97,0%, специфичность Spe1 – 93,2%. Второй тест имеет: чувствительность Sen2 – 93,0%, специфичность Spe2 – 95,0%. Показатель PPV(1+2) характеризует случай положительных результатов и для первого и для повторного теста. Показатель NPV(1+2) характеризует случай положительного результата для первого теста и отрицательного – для повторного теста.

Prevalence,%	1	2	3	5	7	10	20	35	55	75
PPV1,%	12,6	22,5	30,6	42,9	51,8	61,3	78,1	88,5	94,6	97,7
NPV1,%	100,0	99,9	99,9	99,8	99,8	99,6	99,2	98,3	96,2	91,2
PPV2,%	15,8	27,5	36,5	49,5	58,3	67,4	82,3	90,9	95,8	98,2
NPV2,%	99,9	99,8	99,8	99,6	99,4	99,2	98,2	96,2	91,7	81,9
PPV(1+2),%	72,8	84,4	89,1	93,3	95,2	96,7	98,5	99,3	99,7	99,9
NPV(1+2),%	98,9	97,9	96,9	94,8	92,7	89,5	79,2	63,9	43,8	24,1

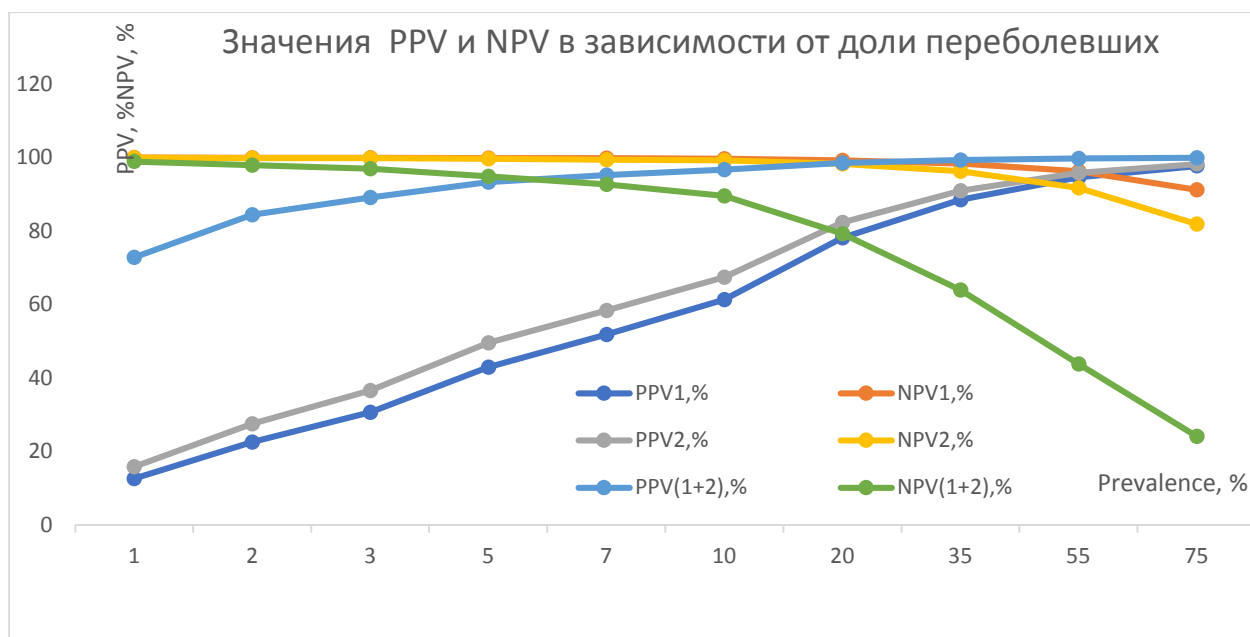


Рисунок 1. Динамика изменения значений PPV и NPV для двух типовых серологических тестов на антитела к SARS-CoV-2 - при их отдельном и парном использовании - по мере увеличения доли (Prevalence) тех, кто в данной популяции уже имеют антитела к SARS-CoV-2. Первый тест имеет: чувствительность Sen1 – 97,0%, специфичность Spe1 – 93,2%. Второй тест имеет: чувствительность Sen2 – 93,0%, специфичность Spe2 – 95,0%.

Важно отметить, что из-за разной скорости распространения инфекции доля Prevalence может широко варьироваться между территориями и между различными группами людей, например даже среди групп разных медицинских работников. При низкой (до 20%) величине доли тех, кто в данной популяции уже имеют антитела к SARS-CoV-2, вероятность фактического наличия антител у пациента при положительном результате однократного тестирования типовым тестом с диагностическими показателями, описанными в Таблице 3, будет составлять менее 80%, что вряд ли можно считать достаточно высокой, чтобы можно было бы сделать надежное заключение о том, что обследуемый пациент был ранее инфицирован или о том, что он действительно имеет антитела к данному коронавирусу. Поэтому в общем случае для увеличения суммарной надежности результата тестирования в документе FDA [1] рекомендуется **повторное тестирование пробы** с положительным результатом на предмет наличия в ней антител к SARS-CoV-2, причем желательно, чтобы альтернативный тест выявлял антитела к другому антигену нового коронавируса. Такая же стратегия двойного тестирования проб с первым положительным результатом рекомендуется и во временных рекомендациях CDC - Centers for Disease Control and Prevention [2]. Ввиду отсутствия надежных данных о текущей доле представителей популяции разных территорий США, которые имеют антитела к SARS-CoV-2, все расчеты значений показателей PPV и NPV для тестов, временно допущенных FDA для использования в практике лабораторий США [1], были проведены в предположении, что эта доля составляет 5%, что скорее всего соответствовало реальной ситуации на май-июнь 2020 года.

Оценочные значения показателей PPV и NPV для конкретного полуколичественного теста рассчитываются по следующим формулам:

$$\text{PPV, \%} = 100 \cdot \text{TP} / (\text{TP} + \text{FP}), \quad (3)$$

где величина TP (True Positive) – число проб, содержащих искомый анализит, в которых его присутствие было выявлено данным тестом, а величина FP (False Positive) - число проб, реально не содержащих искомый анализит, но для которых данный тест показал присутствие анализита.

$$\text{NPV, \%} = 100 \cdot \text{TN} / (\text{TN} + \text{FN}), \quad (4)$$

где величина TN (True Negative) – число проб, реально не содержащих искомый анализ, для которых данный тест показал отсутствие искомого анализа, а величина FN (False Negative) - число проб, реально содержащих искомый анализ, но для которых присутствие искомого анализа не было выявлено данным тестом.

Примеры расчета значений показателей PPV и NPV для типового серологического теста приведены в Таблицах 4 и 5. В Таблице 4 доля представителей популяции, которые имеют антитела к SARS-CoV-2 (Prevalence), предполагается равной 5%, а в Таблице 5 - равной 15%.

Таблица 4. Примеры расчета значений PPV и NPV для типового серологического теста на антитела к SARS-CoV-2. Оценочное значение для чувствительности теста Sen – 97,0%, для специфичности теста Spe – 93,2%. **Prevalence – 5%**. Обследуемая популяция – 10000 пациентов.

Число пациентов с антителами -500 (5%)	Число пациентов без антител -9500 (95%)
True Positive (TP) - 485	False Positive (FP) – 646
False Negative (FN) -15	True Negative (TN) - 8854
PPV, % = 42,9%	NPV, % = 99,8%

Таблица 5. Примеры расчета значений PPV и NPV для типового ИФА теста на антитела к SARS-CoV-2. Оценочное значение для чувствительности теста Sen – 97,0%, для специфичности теста Spe – 93,2%. **Prevalence – 15%**. Обследуемая популяция – 10000 пациентов.

Число пациентов с антителами -1500 (15%)	Число пациентов без антител -8500 (85%)
True Positive (TP) - 1455	False Positive (FP) – 578
False Negative (FN) -45	True Negative (TN) – 7922
PPV, % = 71,6%	NPV, % = 99,4%

Значения показателей PPV(1+2) и NPV(1+2) для алгоритма двойного тестирования, когда положительно протестированные пациенты тестируются повторно независимым тестом, рассчитываются по следующим формулам:

$$\begin{aligned} \text{PPV}(1+2), \% &= 100 \cdot \left\{ \frac{\text{TP1} \cdot \text{Sen2} / 100}{\text{TP1} \cdot \text{Sen2} / 100 + \text{FP1} \cdot (1 - \text{Spe2} / 100)} \right\} = \\ &= \frac{\text{TP1} \cdot \text{Sen2}}{\text{TP1} \cdot \text{Sen2} / 100 + \text{FP1} \cdot (1 - \text{Spe2} / 100)}, \end{aligned} \quad (5)$$

где показатель PPV(1+2) характеризует положительную прогностическую величину для случая, когда оба результата положительные, величина TP1 – число проб, содержащих искомый анализ, в которых его присутствие было выявлено первым тестом, величина FP1 - число проб, реально не содержащих искомый анализ, но для которых первый тест показал присутствие анализа, величины Sen2 и Spe2 – соответственно чувствительность и специфичность второго теста.

$$\begin{aligned} \text{NPV}(1+2), \% &= 100 \cdot \left\{ \frac{\text{FP1} \cdot \text{Spe2} / 100}{\text{FP1} \cdot \text{Spe2} / 100 + \text{TP1} \cdot (1 - \text{Sen2} / 100)} \right\} = \\ &= \frac{\text{FP1} \cdot \text{Spe2}}{\text{FP1} \cdot \text{Spe2} / 100 + \text{TP1} \cdot (1 - \text{Sen2} / 100)}, \end{aligned} \quad (6)$$

где показатель NPV(1+2) характеризует отрицательную прогностическую величину для случая, когда первый результат положительный, а второй – отрицательный, величина FP1 - число проб, реально не содержащих искомый анализ, но для которых первый тест показал присутствие анализа, величина TP1– число проб, содержащих искомый анализ, в которых его присутствие было выявлено первым тестом величины, Sen2 и Spe2 – соответственно чувствительность и специфичность второго теста.

Примеры расчета значений показателей PPV(1+2) и NPV(1+2) для типовых полуколичественных тестов и их динамика в зависимости от доли переболевших (Prevalence) приведены ранее в Таблице 3 и на Рисунке 1. Следует отметить, что значения PPV(1+2) имеют тенденцию к росту по мере увеличения доли переболевших. Чего нельзя сказать о показателе NPV(1+2). Значение NPV(1+2) при доле переболевших равной 10% становится менее 90% и затем по мере ее увеличения показатель NPV(1+2) драматически снижается, то есть вероятность фактического

отсутствия антител в пробе начинает эффективно уменьшаться, причем при доле переболевших, равной 55%, она становится менее 44%.

Примечание 1. Доля переболевших в среднем по официальным данным на конец июня-начало июля с.г. по США составляла около 0,8%, а в среднем по РФ - 0,5%, в Москве -1,8%. Тогда вероятность фактического наличия антител в пробе, которая дала положительный результат, при использовании типового серологического теста с чувствительностью 93% и специфичностью 95% составит соответственно 16% - в среднем по США, около 9% - в среднем по РФ и около 28% - для Москвы (см. данные в Таблице 3 для показателя PPV2,%). Значение показателя PPV для используемого теста, например, равное 28%, означает, что из 100 положительно протестированных пациентов только у 28 фактически будут антитела к SARS-CoV-2. Скорее всего, реальная доля переболевших в несколько раз больше официальных статистических данных. Так что в реальных условиях для того же теста при наличии положительного результата и доле "переболевших" в 2% и 3% та же вероятность фактического наличия антител будет составлять соответственно уже почти 28% и 37%. При доле "переболевших" в 5%, 10% и 20% - для того же теста эта же вероятность будет уже равна соответственно 50%, 67% и 82%. При доле "переболевших" в 35%, 55% и 75% - для того же теста эта вероятность будет равна соответственно 91%, 96% и 98%. При этом повторный положительный результат, полученный независимым тестом, будет свидетельствовать о фактическом наличии антител с вероятностью уже свыше 93% при доле «переболевших», равной 5%. Таким образом, в настоящих условиях пандемии, когда доля переболевших низкая (менее 20%) видимо просто необходимо повторять тестирование пациента еще раз, если первый результат оказался положительным.

Примечание 2. Следует отметить, что «надежность» результатов при использовании даже «суперточных» тестов на антитела тоже очень сильно зависит от доли переболевших. В условиях 1% доли переболевших при использовании суперточного теста с чувствительностью 99% и специфичностью 99% вероятность фактического наличия антител при положительном результате составит только 50%. Это означает, что из 100 положительно протестированных пациентов только у 50 пациентов фактически будут антитела к SARS-CoV-2. При доле в 2% - уже 67%. При доле в 3% - уже 75%. При доле в 5% - 84%. При доле в 10% - 93%. И только при доле переболевших в 15-16% вероятность фактического наличия антител при положительном результате составит 95%. Так что в Санкт-Петербурге, в Москве и на любой другой территории РФ в текущих условиях пандемии надо, видимо, делать повторный анализ на антитела еще одним, но независимым тестом, если первый тест дал положительный результат. Даже если первый тест является «суперточным» тестом и по чувствительности, и по специфичности.

Доверительные интервалы для диагностических показателей полуколичественных тестов. Следует отметить, что на практике, ввиду ограниченности средств и времени, вместо определения истинных значений для диагностических показателей полуколичественных тестов, в том числе для чувствительности и специфичности, их оценки делаются с использованием статистически значимого, но ограниченного количества биопроб изначально здоровых людей, заболевших или тех, которые уже выздоровели. Точность такой оценки, в виду того, что объем исследований при проведении клинических испытаний тестов всегда ограничен, зависит, в основном, от количества проанализированных биопроб. Чем больше биопроб использовано при проведении испытаний, тем с более высокой точностью проведена оценка. В таких случаях принято на основе полученных оценочных значений для чувствительности и специфичности приводить доверительные интервалы, границы которых зависят не только от числа проанализированных биопроб, но и от выбранной доверительной вероятности. В лабораторной медицине обычно используют 95% доверительные интервалы, в том числе и для истинных значений диагностических показателей полуколичественных серологических тестов.

Как было упомянуто в начале обзора, верхнюю и нижнюю границы доверительных интервалов для истинных значений чувствительности и специфичности рекомендуется вычислять по оценочным значениям для этих показателей и объему использованных для этого биопроб в соответствии с методологией, изложенной в руководстве CLSI EP12-A2 [4]. Авторы этого руководства

для вычисления границ доверительных интервалов рекомендуют использовать алгоритм Вильсона (Wilson Score), наилучшим образом определяющий такие границы при небольших объемах выборки (количестве исследуемых биопроб) и при граничных (близких к 100%) величинах оценочных значений для чувствительности и специфичности.

Верхнюю и нижнюю границы доверительных интервалов для истинных значений чувствительности и специфичности вычисляют по формуле:

$$(\text{Sen},\%)_{\text{вг/нг}} = 100 * \{ (2np + z^2) \pm z [(z^2 + 4np(1-p))^{1/2}] / 2(n + z^2), \quad (7)$$

где n – количество проб, протестированных для определения оценочных значений чувствительности или специфичности, p – оценочные значения для $\text{Sen},\%$ или $\text{Spe},\%$ в процентах, деленные на 100, z – коэффициент, равный 1,96 для 95% двустороннего доверительного интервала ($z=1,65$ – для 90% доверительного интервала и $z=2,58$ – для 99% доверительного интервала).

Ранее в Таблице 2 были приведены результаты тестирования 1000 биопроб типовым серологическим тестом по выявлению в них антител к SARS-CoV-2, на основе которых было вычислено оценочное значение для чувствительности, равное $\text{Sen},\% = 93,00\%$. Истинное количество проб с антителами было равно $n=100$, а без антител – соответственно 900. Воспользуемся этими данными о чувствительности ($\text{Sen},\% = 93\%$) и о количестве проб, протестированных для определения оценочных значений чувствительности ($n=100$), и вычислим по формуле (7) значение для нижней границы $(\text{Sen},\%)_{\text{нг}}$ 95% интервала для истинного значения чувствительности. Это значение будет равно:

$$(\text{Sen},\%)_{\text{нг}} = 100 * \{ [(2 * 100 * 0,93 + 1,96 * 1,96) - 1,96 * [(1,96 * 1,96 + 4 * 100 * 0,93 * (1 - 0,93))]^{1/2}] / 2 * (100 + 1,96 * 1,96) = 86,3\%$$

Соответственно верхняя граница $(\text{Sen},\%)_{\text{вг}}$ 95% интервала для истинного значения чувствительности будет равна:

$$(\text{Sen},\%)_{\text{вг}} = 100 * \{ [(2 * 100 * 0,93 + 1,96 * 1,96) + 1,96 * [(1,96 * 1,96 + 4 * 100 * 0,93 * (1 - 0,93))]^{1/2}] / 2 * (100 + 1,96 * 1,96) = 96,6\%$$

Для этого же серологического теста по результатам тестирования 1000 биопроб, приведенным в той же Таблице 2, было определено оценочное значение для специфичности, равное $\text{Spe},\% = 98,9\%$. Количество проб, протестированных для определения оценочных значений специфичности (истинное число проб без антител), было равно $n=900$. Воспользуемся этими данными о специфичности теста $\text{Spe},\% = 98,9\%$ и о количестве проб, протестированных для определения оценочных значений специфичности $n=900$, и вычислим по формуле (7) значение для нижней границы $(\text{Spe},\%)_{\text{нг}}$ 95% интервала для истинного значения специфичности. Это значение будет равно:

$$(\text{Spe},\%)_{\text{нг}} = 100 * \{ [(2 * 900 * 0,989 + 1,96 * 1,96) - 1,96 * [(1,96 * 1,96 + 4 * 900 * 0,989 * (1 - 0,989))]^{1/2}] / 2 * (900 + 1,96 * 1,96) = 98,0\%$$

Соответственно верхняя граница $(\text{Spe},\%)_{\text{вг}}$ 95% интервала для истинного значения специфичности будет равна:

$$(\text{Spe},\%)_{\text{вг}} = 100 * \{ [(2 * 900 * 0,989 + 1,96 * 1,96) + 1,96 * [(1,96 * 1,96 + 4 * 900 * 0,989 * (1 - 0,989))]^{1/2}] / 2 * (900 + 1,96 * 1,96) = 99,4\%$$

В результате будем иметь следующую информацию о диагностических показателях рассматриваемого типового теста:

- оценочное значение для чувствительности Sen равно 93, а нижняя и верхняя границы 95% доверительного интервала для ее истинного значения равны соответственно 86,3% и 96,6%;

- оценочное значение для специфичности Spe равно 98,9%, а нижняя и верхняя границы 95% доверительного интервала для ее истинного значения равны соответственно 98,0% и 99,4%.

Нижние и верхние границы приведенных 95% доверительных интервалов информируют о том, что истинное значение для чувствительности этого теста с вероятностью 95% лежит где-то в пределах границ указанного для нее доверительного интервала, то есть оно не ниже 86,3%, но и не выше 96,6%. Соответственно истинное значение для специфичности этого теста с вероятностью 95% лежит где-то в пределах границ указанного для нее доверительного интервала, то есть оно не ниже 98,0%, но и не выше 99,4%.

Следует обратить внимание, что ширина доверительного интервала для истинного значения специфичности теста оказалась почти в 7 раз меньше, чем для его чувствительности. Это объясняется в первую очередь разными объемами пулов протестированных биопроб – для оценки чувствительности тестировались 100 биопроб, а для оценки специфичности тестировались 900 биопроб. Иначе говоря, чем больше протестировано биопроб для оценки диагностического показателя, тем ближе его оценочное значение к истинному.

Примечание 3. Обычно в ходе клинических испытаний полуколичественных тестов ограничиваются тестированием менее 100 биопроб для получения оценочных значений для их чувствительности, что не так много для ее надежной оценки. Из-за небольшого объема тестируемых биопроб полученные в ходе испытаний оценочные значения для чувствительности могут значительно отличаться от истинных. Именно поэтому в документе FDA [1] помимо оценочных значений, полученных в ходе клинических испытаний серологических тестов на антитела к SARS-CoV-2, временно допущенных FDA в лабораторную практику США, дополнительно приведены 95% доверительные интервалы для истинных значений не только чувствительности и специфичности, но и для показателей PPV и NPV.

Нижние и верхние границы 95% доверительных интервалов для истинных значений показателей PPV и NPV для испытываемого теста определяют по формулам (3) и (4), в которые подставляют соответствующие значения TP, FP, TN и FN, вычисляемые по алгоритму, изложенному в Таблице 2, с учетом выбранной доли «переболевших» (Prevalence). Примеры расчета этих границ приведены в Таблицах 6 и 7. При расчете нижних (верхних) границ 95% доверительных интервалов для истинных значений показателей PPV и NPV надо использовать соответственно нижние (верхние) границы 95% доверительных интервалов для истинных значений чувствительности и специфичности. Напомним, что в документе FDA [1] при вычислении оценочных значений для показателей PPV и NPV, и также при вычислении нижних и верхних границ 95% доверительных интервалов для истинных значений показателей PPV и NPV, доля переболевших выбрана равной 5%. Калькулятор, предлагаемый в этом документе для свободного доступа, позволяет легко проводить расчеты оценочных значений для PPV и NPV для любых долей «переболевших» и оценочных значений для чувствительности и специфичности тестов.

Таблица 6. Примеры расчета нижних границ доверительных интервалов для истинных значений PPV и NPV для типового серологического теста выявления антител к SARS-CoV-2. Оценочное значение для чувствительности теста – 93,0%, для специфичности – 98,9%, для PPV – 81,7%, для NPV – 99,6%. Обследуемая популяция – всего 10.000 пациентов. **Prevalence – 5%**. Нижняя граница 95% доверительного интервала для истинного значения чувствительности (**Sen,%**)**нг равна 86,3%**, и соответственно для истинного значения специфичности (**Spe,%**)**нг равна 98,0%**.

Число пациентов с антителами -500 (5%)	Число пациентов без антител -9500 (95%)
True Positive (TP) – 432	False Positive (FP) – 190
False Negative (FN) – 68	True Negative (TN) – 9310
PPVнг = 69,5%	NPVнг = 99,3%

$$PPV_{нг}, \% = 100 * 432 / (432 + 190) = 69,5\%,$$

$$NPV_{нг}, \% = 100 * 9310 / (9310 + 68) = 99,3\%$$

Таблица 7. Примеры расчета верхних границ PPV_{нг} и NPV_{нг} доверительных интервалов для истинных значений PPV и NPV для типового серологического теста выявления антител к SARS-CoV-2. Оценочное значение для чувствительности теста – 93,0%, для специфичности – 98,9%, для PPV – 81,7%, для NPV – 99,6%. Обследуемая популяция – всего 10.000 пациентов. **Prevalence – 5%**. Верхняя граница 95% доверительного интервала для истинного значения чувствительности (**Sen,%**)**вг равна 96,6%** и соответственно для истинного значения специфичности (**Spe,%**)**вг равна 99,4%**.

Число пациентов с антителами -500 (5%)	Число пациентов без антител -9500 (95%)
True Positive (TP) - 483	False Positive (FP) – 57
False Negative (FN) -17	True Negative (TN) – 9443
PPV _{нг} = 89,44%	NPV _{нг} = 99,8%

$$PPV_{вг}, \% = 100 * 483 / (483 + 57) = 89,44\%,$$

$$NPV_{вг}, \% = 100 * 9443 / (9443 + 17) = 99,8\%$$

В результате для диагностических показателей PPV и NPV рассматриваемого теста получим следующее. Оценочное значение для показателя PPV равно 81,7%. Нижняя и верхняя границы 95% доверительного интервала для истинного значения PPV равны соответственно 69,5% и 89,44%. Оценочное значение для показателя NPV равно 99,6%. Нижняя и верхняя границы 95% доверительного интервала для истинного значения NPV равны соответственно 99,3% и 99,8%.

Стратегия использования полуколичественных ИФА-тестов в лабораторной практике.

Варианты стратегии использования в лабораторной практике США полуколичественных тестов, предназначенных для выявления антител к SARS-CoV-2, изложены во временных рекомендация CDC [2]. В этом документе отмечается, что все тесты, используемые в практике коммерческих лабораторий США, должны иметь временное разрешение FDA. Перечень тестов, временно допущенных FDA к использованию в лабораторной практике США, опубликован в документе [1]. В документе [2] также отмечается, что в текущих условиях пандемии, связанной с распространением нового коронавируса SARS-CoV-2, когда доля переболевших COVID-19 еще достаточно низкая (варьируется от 1% до 25%), в большинстве случаев преимущественно обладают тесты с максимальной специфичностью, обеспечивающие наибольшие значения для положительного прогностического показателя PPV. Например, в условиях популяции, где доля переболевших равна 5%, тест с 90% чувствительностью и с 95% специфичностью будет обеспечивать для прогностического показателя PPV значение, равное 49%. Другими словами, меньше половины протестированных таким тестом, для которых были получены положительные результаты по наличию у них антител, будут на самом деле иметь антитела. С другой стороны, тот же самый тест при его использовании условиях популяции, где доля переболевших будет составлять уже свыше 52%, будет обеспечивать для прогностического показателя PPV значение, превышающее 95%, означая, что менее, чем у 1 из 20 положительно протестированных пациентов будет ложноположительный результат.

В этой связи в документе [2] для текущих условий пандемии с низкими долями переболевшего COVID-19 населения предлагаются следующие три стратегии для увеличения значений положительного прогностического показателя:

1) выбирать тест с очень большой специфичностью, по возможности с 99,5% специфичностью и выше, что будет давать максимально высокие значения (более 90%) для положительного прогностического показателя при его использовании для тестирования популяции с долей переболевших, равной 5% и более процентов;

2) концентрировать тестирование на пациентах, имеющих априори высокую вероятность того, что они уже имеют антитела к SARS-CoV-2, например, на пациентах, имеющих симптомы, сходные с симптомами инфицированных новым коронавирусом;

3) применять так называемый алгоритм двойного тестирования, когда положительно протестированные пациенты тестируются повторно еще одним тестом. Эффективность такого алгоритма базируется на тестировании пробы пациента двумя «независимыми» тестами, каждый из которых обладает уникальными технологическими характеристиками, например нацеленностью на разные эпитопы коронавируса (S-белок или N-капсидный антиген), или различаются своими форматами/типами.

Применение третьей стратегии позволяет в условиях тестирования представителей популяции, где доля переболевших составляет около 5%, обеспечивать для прогностического показателя PPV значение равное 95% даже при использовании второго теста, который обладает сходными с первым тестом диагностическими показателями - имеет 90% чувствительность и 95% специфичность. Для проведения расчетов значений прогностического показателя PPV для алгоритма двойного тестирования с использованием тестов с разными чувствительностью и специфичностью в рекомендациях CDC [2] предлагается использовать калькулятор, разработанный FDA и упомянутый ранее в данном обзоре [3].

В данных рекомендациях CDC также отмечается, что использование в клинической практике результатов тестирования на наличие антител к SARS-CoV-2 дополнительно имеют некоторые ограничения. Так, до сих пор еще не найдены ответы на вопросы, касательно уровня антител, который обеспечивает защиту от повторного инфицирования, касательно длительности действия приобретенной иммунной защиты, и касательно факторов, ассоциируемых с наработкой человеческим организмом антител.

В документе [2] также отмечается, что некоторые тесты на anti-SARS-CoV-2 могут реагировать и на другие коронавирусы, например на коронавирусы, которые вызывают обычные простудные заболевания. Это может приводить к увеличению ложноположительных результатов при тестировании такими тестами. Некоторые пациенты могут после инфицирования не развивать у себя в организме антитела до детектируемого уровня. Для других пациентов считается возможным, что уровень антител может со временем снижаться до недетектируемого уровня. И наконец, антитела IgM, IgA и IgG не появляются сразу же после инфицирования. Таким образом, результаты серологических тестов на антитела к SARS-CoV-2 не могут со 100% надежностью указывать на присутствие или отсутствие факта текущего или предыдущего инфицирования новым коронавирусом.

Заключение. Цель данного обзора состояла в том, чтобы, во-первых, проинформировать пользователей серологических тестов на антитела к SARS-CoV-2 о специфике и возможной стратегии их применения в настоящих условиях, и во-вторых, рассказать о технологиях определения значений для диагностических показателей полуколичественных тестов и алгоритмах оценки их надежности, что позволяет оценивать реальные диагностические возможности тестов и правильно интерпретировать получаемые результаты. В обзоре приведены также ссылки на специализированный калькулятор FDA, формулы и примеры расчетов как оценочных значений, так и границ доверительных интервалов для истинных значений диагностических показателей серологических тестов, предназначенных для выявления антител к SARS-CoV-2. Это дает возможность любому пользователю таких тестов самостоятельно проводить вычисления необходимых параметров и оценивать надежность результатов, получаемых с их помощью, с учетом региональных особенностей (доли Prevalence, размера обследуемой популяции и других).

Литература.

1. Диагностические характеристики серологических тестов для выявления антител к новому коронавирусу, временно допущенных FDA для использования в лабораторной практике США www.fda.gov/medical-devices/emergency-situations-medical-devices/eua-authorized-serology-test-performance
2. Interim Guidelines for COVID-19 Antibody Testing in USA Clinical and Public Health Settings www.cdc.gov/coronavirus/2019-ncov/lab/resources/antibody-tests-guidelines.html
3. Calculator for Positive Predictive Value (PPV) and Negative Predictive Value (NPV) for individual tests and combined <https://www.fda.gov/media/137612/download>
4. Clinical and Laboratory Standards Institute. Approved Guideline CLSI EP12-A2. User Protocol for Evaluation of Qualitative Test Performance. Second Edition, Volume 28, Number 3, January 2008