

Достоверность диагностических тестов

- Validity – is an expression of the degree to which a measurement measures what it purports to measure.
- Достоверность – это выражение способности инструмента измерять то, что он должен измерять.

- «Золотой стандарт» - общепринятый наилучший из всех имеющихся метод измерения/диагностики определенного признака (заболевания, фактора риска, биологического явления и т.п.), используемый для сравнения с новым методом измерения или диагностики.

Представление результатов оценки достоверности дихотомического диагностического теста

Золотой стандарт

		З(+)	З(-)
Результаты теста	T(+)	Истинно- положительный (ИП)	Ложно- положительный (ЛП)
	T(-)	Ложно- отрицательный (ЛО)	Истинно- отрицательный (ИО)

Показатели достоверности теста

- Чувствительность (sensitivity)
- Специфичность (specificity)
- Прогностическая ценность положительного результата теста
- Прогностическая ценность отрицательного результата теста

Показатели достоверности теста

- Чувствительность (sensitivity) – способность теста определять болезнь у больных (доля лиц с положительным результатом теста у больных, вероятность получения положительного результата теста у больных)

Показатели достоверности теста

- Специфичность (specificity) – способность теста определять отсутствие болезни у здоровых (доля лиц с отрицательным результатом теста среди здоровых, вероятность получить отрицательный результат теста среди здоровых людей)

Чувствительность (Se)

Золотой стандарт

		З(+)	З(-)
Результаты теста	T(+)	ИП	ЛП
	T(-)	ЛО	ИО
		ИП+ЛО	ЛП+ИО

$$Se = \frac{ИП}{ИП+ЛО} \times 100\%$$

На самом деле
больные

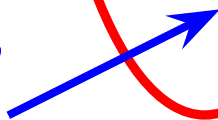
Специфичность (Sp)

Золотой стандарт

		Золотой стандарт	
		З(+)	З(-)
Результаты теста	T(+)	ИП	ЛП
	T(-)	ЛО	ИО
		ИП+ЛО	ЛП+ИО

$$Sp = \frac{ИО}{ЛП+ИО} \times 100\%$$

На самом деле
здоровые



Задача

- Предположим, что мы имеем гипотетическую популяцию размером 1000 человек, 100 из которых имеют заболевание и 900 не имеют. Мы располагаем определенным тестом, который может разделить людей на больных ИБС и здоровых людей. В данном случае это тест ЭКГ. Насколько хорош оказался примененный нами тест, если он выявил 80 больных и 800 здоровых людей?
- Во-первых, насколько хорош был наш тест в выявлении больных людей?
- Во-вторых, насколько хорош был наш тест в выявлении здоровых людей?

Чувствительность (**Se**)

ИБС

З(+)

З(-)

ЭКГ

T(+)

80

100

T(-)

20

800

100

900

1000

$$\mathbf{Se} = \frac{80}{100} \times 100\% = 80\%$$

Специфичность (**Sp**)

ИБС

З(+)

З(-)

ЭКГ

T(+)

80

100

T(-)

20

800

100

900

1000

$$\mathbf{Sp} = \frac{800}{900} \times 100\% = 89\%$$

Высоко-чувствительный тест

- Редко пропускает болезнь, т. е. редко дает ложно-отрицательные результаты
- Высокая ценность отрицательного результата теста: врач с большей уверенностью может сказать, что пациент с отрицательным результатом здоров.

Высоко-специфичный тест

- Редко классифицирует здоровых как больных, т.е. редко дает ложно-положительные результаты
- Высокая ценность положительного результата теста: врач с большей уверенностью может сказать, что пациент с положительным результатом болен.

- Обычно не бывает диагностических тестов с очень высокой чувствительностью и очень высокой специфичностью одновременно.
- Чувствительность и специфичность как бы «торгуются» друг с другом: если увеличивается чувствительность, то снижается специфичность и наоборот.

Высокая чувствительность или высокая специфичность, что мне нужно?

- Число ЛП и ЛО результатов, которые может «выдержать» система здравоохранения
- Последствия, связанные с получением ЛП и ЛО результатов
- Злокачественность или серьезность заболевания или исхода
- Заразность заболевания
- Стоимость диагностического теста
- Риски для здоровья, связанные с диагностическими тестами и ранним лечением

Прогностическая ценность теста

- Прогностическая ценность положительного результата теста (positive predictive value) – показывает вероятность наличия болезни при положительном результате теста
- Прогностическая ценность отрицательного результата теста (negative predictive value) – показывает вероятность отсутствия болезни при отрицательном результате теста

(+PV)

Золотой стандарт

З(+)

З(-)

Результаты теста

T(+)	ИП	ЛП	ИП+ЛП
T(-)	ЛО	ИО	ЛО+ИО

The table is a 2x2 contingency table. The top row is circled in red. A blue arrow points from the bottom right towards the 'ИП+ЛП' cell in the top row.

$$+PV = \frac{\text{ИП}}{\text{ИП} + \text{ЛП}} \times 100\%$$

Положительный результат

(-PV)

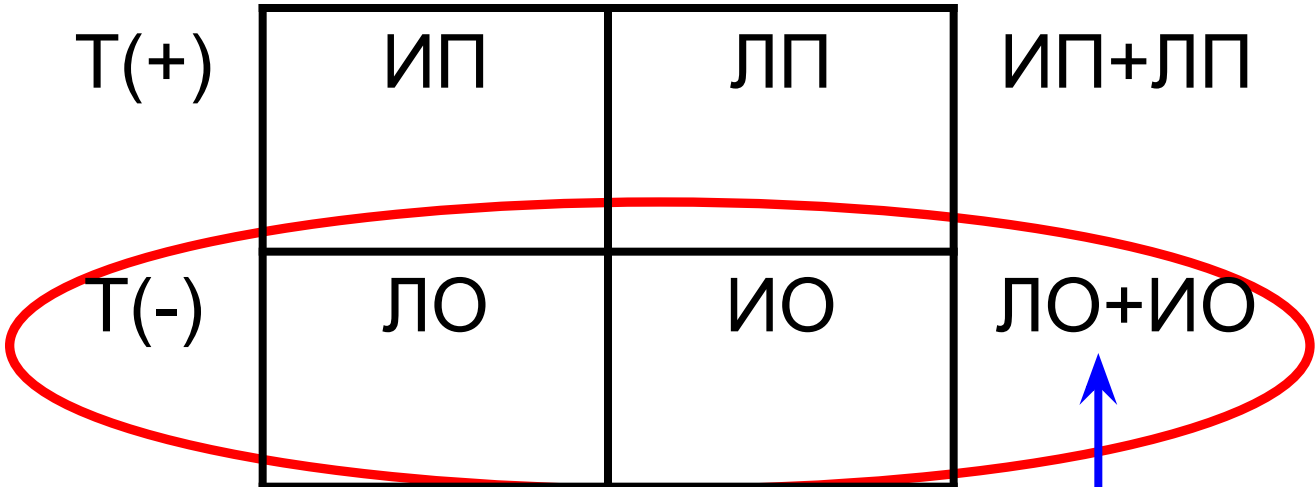
Золотой стандарт

З(+)

З(-)

Результаты теста

Т(+)	ИП	ЛП	ИП+ЛП
Т(-)	ЛО	ИО	ЛО+ИО



$$-PV = \frac{ИО}{ЛО+ИО} \times 100\%$$

Отрицательный результат

Задача

- Предположим, что мы имеем гипотетическую популяцию размером 1000 человек, 100 из которых имеют заболевание и 900 не имеют. Мы располагаем определенным тестом, который может разделить людей на больных ИБС и здоровых людей. В данном случае это тест ЭКГ. Насколько хорош оказался примененный нами тест, если он выявил 80 больных и 800 здоровых людей?
- Какова вероятность наличия болезни при положительном результате теста?
- Какова вероятность отсутствия болезни при отрицательном результате теста?

(+PV)

ИБС

З(+)

З(-)

ЭКГ

Т(+)

80

100

180

Т(-)

20

800

1000

$$+PV = \frac{80}{180} \times 100\% = 44,4\%$$

(-PV)

ИБС

З(+)

З(-)

ЭКГ

T(+)	80	100	180
T(-)	20	800	820
			1000

$$-PV = \frac{800}{820} \times 100\% = 97,6\%$$

Задача

- Достоверность теста А на предмет выявления ревматоидного артрита была измерена на группе людей, 100 из которых были больны этим заболеванием и 400 из которых не имели этого заболевания. В первой группе 80 человек получили положительный результат теста и во второй группе 8 человек получили положительный результат. Определите чувствительность и специфичность теста А.

		Ревматоидный артрит		
		3(+)	3(-)	
Тест А	T(+)	80	8	
	T(-)		392	
		100	400	

$$Se = \frac{80}{100} \times 100\% = 80\%$$

$$Sp = 98\%$$

Хотели бы вы что-нибудь еще знать о ходе испытания теста перед тем как внедрить его в практику?

- Как была сформирована выборка для испытания теста?
 - Многие диагностические тесты положительны только на стадии ярко выраженных случаев заболеваний, чем примененные на доклинической стадии.
- Доп. инф.: Тест испытывался на госпитализированных пациентах. Как этот факт повлияет на чувствительность теста?
 - 80% чувствительность может не отражать реальную чувствительность теста А. Скорее всего чувствительность в данном случае завышена.

Тест испытывался на госпитализированных пациентах.
Что можно сказать о специфичности данного теста?

- Специфичность – это способность теста определять отсутствие заболевания среди людей без заболевания.
- Специфичность может быть занижена, когда тест испытывается на госпитализированных пациентах, так как другие заболевания могут иметь схожие с ревматоидным артритом проявления.

Если инструмент (тест) используется для определения распространенности заболевания, но имеет низкую чувствительность, как это повлияет на показатель распространенности?

- Низкая чувствительность означает, что люди с заболеванием будут ошибочно отнесены к не имеющим данного заболевания.
- Следовательно показатель распространенности будет занижен.

Если тест имеет низкую специфичность, то как это повлияет на показатель распространенности заболевания?

- Низкая специфичность означает, что будет много людей, которые ошибочно будут классифицированы как больные.
- Следовательно показатель распространенности заболевания будет завышен.
- Ошибка мисклассификации (классификационная ошибка) – разновидность информационной ошибки.

Ошибка мисклассификации

- Это ложное отнесение индивидуума, значения или какого-либо признака в категорию, к которой он/оно не принадлежит
- Вероятность классификационной ошибки может быть одинаковой в разных изучаемых подгруппах (недифференцированная мисклассификация)
- Когда вероятность возникновения классификационной ошибки разная в изучаемых подгруппах, то мы говорим о дифференцированной мисклассификации.

- Рассчитайте показатель распространенности ревматоидного артрита, используя тест А, в двух небольших населенных пунктах - гор.А и гор.Б
- Заранее известно, что согласно «Золотому стандарту» в гор.А распространенность ревматоидного артрита = 21%, а в гор.Б 7%
- Население каждого пункта – 10 000 жителей

Гор.А

		РА		
		3(+)	3(-)	
Тест А	T(+)	1680	158	1838
	T(-)	420	7742	
		2100	7900	10 000

***Se = 80% Sp = 98%**

Распространенность = $1838/10\ 000 \times 100 = 18,4\%$

Гор.Б

		РА		
		3(+)	3(-)	
Тест А	T(+)	560	186	746
	T(-)	140	9114	
		700	9300	10 000

***Se = 80% Sp = 98%**

Распространенность = $746/10\ 000 \times 100 = 7,5\%$

- Согласно реальной распространенности РА в гор.А и гор.Б, мы видим, что показатель распространенности в гор.А в 3 раза больше чем в гор.Б ($21 / 7 = 3$)
- Какое соотношение мы получим при использовании теста А?
- $18,4 / 7,5 = 2,5$ вместо 3
- Таким образом, когда мы сравниваем 2 группы людей, используя тест с одинаковыми для двух групп чувствительностью и специфичностью, любая ошибка мисклассификации будет **всегда уменьшать** разницу между сравниваемыми группами

Задача

- Недовольный тестом А, врач Б разработал новый тест для определения ревматоидного артрита. Этот тест, назанный в честь доктора тестом Б, имеет чувствительность 99% и специфичность 86%. Тест Б стали использовать для выявления распространенности заболевания в гор.Б, а результат стали сравнивать с распространенностью в гор.А (где используется тест А). Как мы помним, в гор.А распространенность ревматоидного артрита была 18,4% согласно тесту А, а реальная распространенность превышала таковую в гор.Б в 3 раза.
- Не производя никаких расчетов, можете ли вы сказать будет ли распространенность ревматоидного артрита в гор.А выше или ниже чем в гор.Б и во сколько раз?

- Нет, мы не можем без расчетов ответить на вопрос
- Если ошибка мисклассификации различна в двух сравниваемых группах, т.е. если имеется разница в чувствительности и/или специфичности, то возможно возникновение ошибки **в любом направлении**
- Истинное различие в показателе распространенности может быть стерто, занижено или завышено. Можно получить различие в показателях, хотя это различие на самом деле отсутствует

- В данном примере использовались разные тесты с различной достоверностью
- Но ошибка мисклассификации может так же быть различной в разных сравниваемых группах даже при использовании одного и того же теста

- Чувствительность и специфичность не являются просто свойством теста или инструмента измерения, скорее они являются свойством теста или инструмента, примененного к конкретной группе или популяции

Гор.Б

		Ревматоидный артрит		
		З(+)	З(-)	
Тест А	T(+)	693	1302	1995
	T(-)	7	7998	
		700	9300	10 000

***Se = 99% Sp = 86%**

Распространенность = $1995/10\ 000 \times 100 = 19,9\%$

- Во сколько раз распространенность ревматоидного артрита в гор.Б выше чем в гор.А?
- Таким образом, распространенность РА в гор. Б на 0.92 раза выше чем в гор.А (18,4/19,9). Следовательно, вследствие дифференцированной мисклассификации соотношение исказилось и РА оказался более распространенным в гор.Б!

Домашнее задание

Задача №1

- Физикальное исследование (пальпацию) использовали в качестве скрининга рака молочной железы у 2500 женщин, у которых рак был подтвержден биопсией (случаи) и у 5000 контрольных женщин (у которых рак отсутствовал). Результаты пальпации были позитивны у 1800 случаев и у 800 контролей.
- Охарактеризуйте достоверность пальпации, используя показатель чувствительности, специфичности и предсказательной ценности положительного и отрицательного результата теста.

Домашнее задание

Задача №2

- Физикальное обследование и аудиометрический тест были назначены 500 пациентам, у которых подозревалось заболевание слуха.
- Из 500 человек 300 (как определил золотой стандарт) на самом деле имели заболевание слухового аппарата

Домашнее задание

Задача №2

- В сравнении с физикальным обследованием, аудиометрический тест является:
- а) одинаково чувствительным и специфичным
- в) менее чувствительным и специфичным
- с) менее чувствительным, но более специфичным
- d) более чувствительным, но менее специфичным
- е) более чувствительным и более специфичным