

Медицинская биостатистика

Лекция 1

Статистика в диагностике Тишков Артем Валерьевич

Кафедра физики, математики и информатики
ПСПбГМУ им. акад. И.П. Павлова

Автор слайдов: Хромов-Борисов Никита
Николаевич

Nikita.KhromovBorisov@gmail.com

<http://independent.academia.edu/NikitaKhromovBorisov>



Статистика

- **Статистику можно определить как науку об изучении статистических данных**

ИЛИ

- **как науку об изменчивости результатов подсчетов и измерений в разных областях научной и практической деятельности человека.**
- **В силу неизбежной изменчивости результатов в биомедицине, фактически все ее выводы и доказательства являются вероятностно-статистическими.**



Данные, пример: каковы признаки этой собаки?



- **Качественные:**
- **Ее окрас - коричневый с черным**
- **У нее длинная шерсть**
- **Она энергичная**
- **Количественные:**
- **сЧЕТНЫЕ:**
 - **У нее 4 ноги**
 - **У нее два брата**
- **мерные:**
 - **Ее вес – 25,5 кг**
 - **Ее рост (в холке) 56,5 см**



Популяция и выборка

- В медицине и в эпидемиологии термин *Популяция* является калькой с английского *population* - синоним статистического понятия «генеральная совокупность», т.е. совокупность индивидуумов, из которой отбирается **выборка**.
- Поскольку вероятностной моделью популяции является определенная **случайная величина** (вероятностная переменная), которая характеризуется определенным **распределением**, то с этой точки зрения удобно говорить о **выборке** из данного **распределения**.



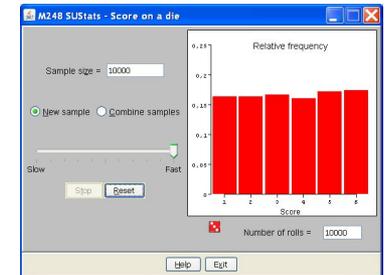
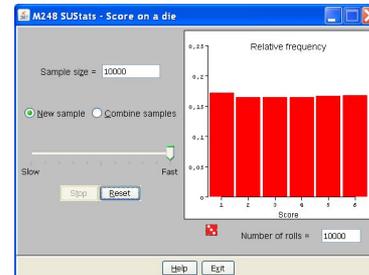
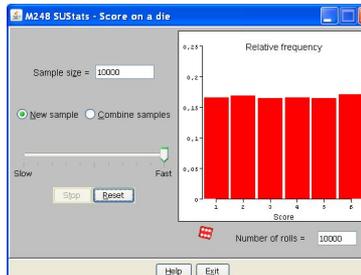
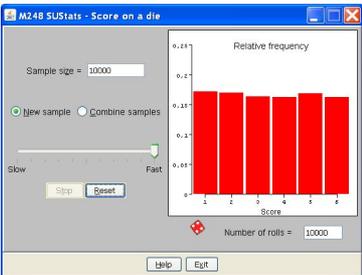
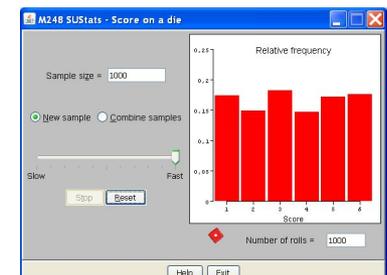
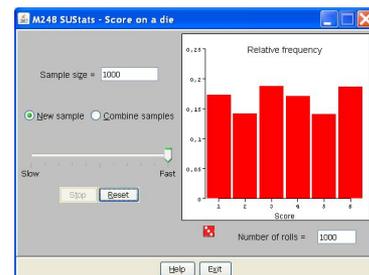
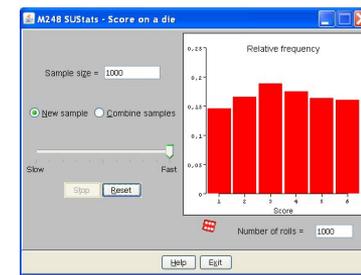
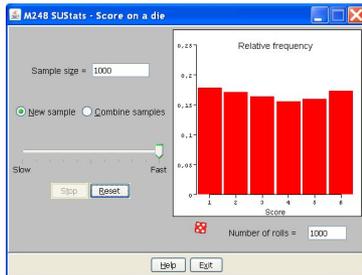
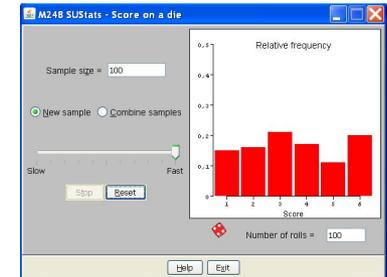
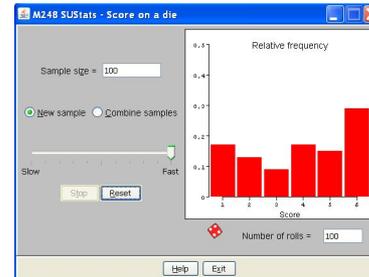
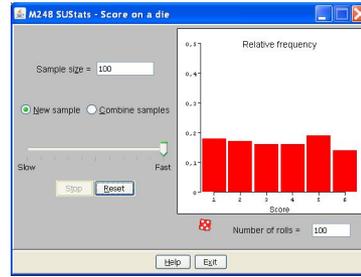
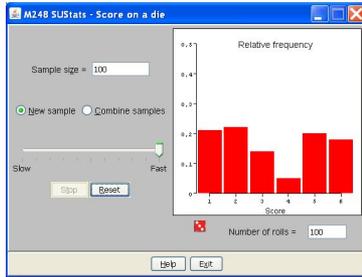
Случайность выборки обуславливает ее репрезентативность

- Один из важнейших вопросов, решаемых статистикой:
- Действительно ли данная выборка отражает основные свойства (параметры) распределения, из которого она извлечена?
- Другими словами, *представительна ли (репрезентативна ли)* имеющаяся выборка?
- Если выборка случайная, то закон больших чисел (ЗБЧ) гарантирует ее репрезентативность.



Имитация многократного вбрасывания игральной кости: программа SUStats.jar

<http://www.jsc.nildram.co.uk/examples/sustats/SUStats.html>





Простейшая форма закона больших чисел - теорема Бернулли

- Она утверждает, что если вероятность события $P(A)$ неизменна во всех испытаниях, то с увеличением числа испытаний частота события $f(A)$ все ближе сходится к его вероятности, т.е. частота становится все более устойчивой и все меньше отклоняется от вероятности.

$$f(A) \xrightarrow{P} P(A)$$
$$n \rightarrow \infty$$

- *Следствие:*
- Мы можем не знать значение вероятности события, но проведя по возможности как можно больше испытаний, мы можем использовать наблюдаемую частоту в качестве надежной статистической оценки этой неизвестной вероятности.

Статистика в диагностике

Статистический контроль
качества диагностических
тестов с бинарными исходами



Что такое «Золотой стандарт» в диагностике?

«Золотой стандарт» (или эталон сравнения) – такой диагностический тест, который максимально точно (практически безошибочно) определяет наличие или отсутствие определенной болезни у пациента.

- **Например, при диагнозе рака одними из наиболее убедительных принято считать результаты гистологической диагностики.**
- **Эталонный метод диагностики совсем не обязательно должен быть единственным. Чаще всего это есть результат применения нескольких диагностических методов.**



Всё познается в сравнении

- Золотой стандарт бывает дорог, болезнен, рискован и потому его нецелесообразно (неэкономично) использовать в массовых обследованиях.
- Очевидно, что и врач и пациент воздержатся от его применения, если им будет предложен другой более простой, щадящий, но не менее надежный метод диагностики.
- **Вопрос:** как убедиться в надежности нового метода?
- **Ответ:** надо сравнить результаты, получаемые с использованием предлагаемого теста, с результатами золотого стандарта.



Статистический анализ – фактически единственный надежный путь оценки качества и сравнения диагностических тестов

- **Диагностические тесты далеко не всегда дают верные ответы, и поэтому возникает необходимость выразить количественно, насколько **точен, надежен и информативен** данный тест.**
- **Очевидно, что лучший и фактически единственный путь работать с несовершенной информацией – использовать *статистические методы* и выражать результаты диагноза и прогноза в терминах *вероятности*.**
- ***Поэтому следует осознавать, что диагностические тесты определяют не наличие или отсутствие болезни, но лишь их вероятности.***



Основные схемы формирования выборок для статистического контроля качества диагностических тестов

- **1. Одновыборочное, когортное (cohort), общепопуляционное - population-based) или поперечно-срезовое одномоментное (cross-sectional) формирование и исследование выборки.** При одновыборочной схеме из данной популяции отбирается **одна** случайная выборка (когорты) субъектов без использования информации о наличии или отсутствии у них данной болезни.
- **2. Двухвыборочное, при котором формируются две независимые случайные выборки по схеме «случаи – контроли» (case-control).**

**Пример одновыборочного
исследования.**

**Ранняя диагностика рака
молочной железы -**

Маммография



Массовый маммографический скрининг

- Banks E., Reeves G., Beral V., Bull D., Crossley B., Simmond M., Hilton E., Bailey S., Barret N., Briers P., Englis R., Jackson A., Kutt E., Lavelle J., Rockal L., Wallis M.G., Wilson M., Patnick J.
- **Influence of personal characteristics of individual women on sensitivity and specificity of mammography in the Million Women Study: cohort study**
- *BMJ, 2004; Vol. 329; No 7464; p. 477*



Результаты массового маммографического скрининга

Таблица сопряженности (перекрестной классификации) 2 × 2

Результат маммографии	Рак молочной железы (золотой стандарт: биопсия)		Всего
	Есть [+]	Нет [-]	
Положительный [+]	629	3 885	4 514
Отрицательный [-]	97	117 744	117 841
Всего	726	121 629	122 355

Таблица 2 × 2 – удобная, наглядная и компактная форма представления результатов сравнения диагностических методов

Проверяемый Диагностический Тест	Эталон сравнения («Золотой стандарт»)		Всего
	[+]	[-]	
[+]	<i>a</i>	<i>b</i>	<i>a + b</i>
[-]	<i>c</i>	<i>d</i>	<i>c + d</i>
Всего	<i>a + c</i>	<i>b + d</i>	<i>n</i>

- Обычно (по умолчанию):
 - В столбцах такой таблицы представляют результаты применения Эталона сравнения («Золотого стандарта»).
 - В строках таблицы представляют результаты применения изучаемого метода.
- Иногда, однако, строки и столбцы в таблице меняют местами.



Основные обозначения для исходов диагностического исследования

- **D** = результаты, полученные с использованием эталона сравнения (от англ. Disease).
- **D+** = наличие болезни по результатам, полученным с использованием эталона сравнения.
- **D-** = отсутствие болезни по результатам, полученным с использованием эталона сравнения.
- **T** = результаты, полученные с использованием проверяемого диагностического средства (от англ. Test).
- **T+** = «**ПОЗИТИВ**», т.е. положительный результат проверяемого теста.
- **T-** = «**НЕГАТИВ**», т.е. отрицательный результат проверяемого теста.

Таблица 2 x 2: возможные исходы диагностического исследования и их наблюдаемые численности

Проверяемый метод диагностики	Эталон сравнения		Всего
	[+]	[-]	
[+]	 (T+, D+) a_{TP}	 (T+, D-) b_{FP}	(T+) $a + b$
[-]	 (T-, D+) c_{FN}	 (T-, D-) d_{TN}	(T-) $c + d$
Всего	(D+) $a + c$	(D-) $b + d$	n

- **Наблюдаемые численности:**
- a_{TP} – число истинных «Позитивов», т.е. верно положительных результатов тестов (T+, D+),
- b_{FP} – число ложных «Позитивов», т.е. ошибочно положительных результатов теста (T+, D-),
- c_{FN} – число ложных «Негативов», т.е. ошибочно отрицательных результатов теста (T-, D+),
- d_{TN} – число истинных «Негативов», т.е. верно отрицательных результатов тестов (T-, D-)



Вероятности исходов. Они нам неизвестны

		Эталон		Всего 0
		[+]	[-]	
Тест	[+]	 $P(T+, D+)$	 $P(T+, D-)$	$P(T+)$
	[-]	 $P(T-, D+)$	 $P(T-, D-)$	$P(T-)$
Всего		$P(D+)$	$P(D-)$	1



Основная логика статистического оценивания: точечные оценки

- Обычно вероятности P (или иные параметры) нам неизвестны.
- Мы их **оцениваем** (благодаря ЗБЧ) по наблюдаемым численностям: a , b , c и d .
- **Например**, точечной оценкой для распространенности болезни по данным одновыборочного исследования $P(D+)$ может служить наблюдаемая частота
 - $f(D+) = (a + c) / n$,
- т.е. доля больных от общего числа n обследованных.



Наблюдаем частоты $f(A)$ (разделим все элементы таблицы на количество пациентов). Их мы признаем за статистические оценки неизвестных нам вероятностей $P(A)$

		Эталон		Всего
		[+]	[-]	
Тест	[+]	 $f(T+, D+)$	 $f(T+, D-)$	$f(T+)$
	[-]	 $f(T-, D+)$	 $f(T-, D-)$	$f(T-)$
Всего		$f(D+)$	$f(D-)$	1



Два основных типа статистических оценок

- Принципиальной особенностью выборочных данных является их изменчивость (варьирование).
- Поэтому при оценивании на их основе параметров распределения, из которого они извлечены, **нельзя ограничиваться одним числом.**
- Обязательно нужны еще оценки их варьирования. Поэтому различают два типа статистических оценок:
- **Точечное оценивание** – оценка одним числом.
- **Интервальное оценивание** – оценка интервалом.
- В статистике для интервального оценивания используются **Доверительные Интервалы.**
- **Доверительный интервал** — это такой интервал, который содержит (накрывает) неизвестный оцениваемый параметр с заданной вероятностью.
- Такая вероятность называется **Доверительной Вероятностью** или **Уровнем Доверия.**
- **Уровень Доверия** выбирается исследователем.



Точечные оценки доли

- Частотническая:

$$\varphi_{\text{freq}} = \frac{y_{\text{obs}}}{n}$$

- Бейзовская:

$$\varphi_{\text{Bayes}} = \frac{y_{\text{obs}} + 1}{n + 2}$$

- если *a priori*

$$\tilde{\varphi} \sim \text{Beta}(a = 1; b = 1)$$



Основная логика статистического оценивания: интервальные оценки

- Понятно, что если мы многократно повторим сравнения данного диагностического теста с эталоном сравнения, то наблюдаемые частоты $f(X_i)$ неизбежно будут варьировать.
- Поэтому задача математиков – вывести математический закон (вероятностное распределение), которому подчиняется варьирование этой частоты.
- Если такой закон найден, то тогда можно получить доверительные интервалы (ДИ) для оценки вероятностей $P(X_i)$ с заданной доверительной вероятностью $(1 - \alpha)$.



Частотнический доверительный интервал и доверительная вероятность

$$P\{\tilde{\theta}_{\text{нижн.}} \leq \theta \leq \tilde{\theta}_{\text{верхн.}}\} = 1 - \alpha$$

- вероятность того, что истинное, неизвестное нам значение параметра θ_{unkn} расположено внутри данного интервала.

$$P\{\theta \leq \tilde{\theta}_{\text{нижн.}}\} = \frac{\alpha}{2}$$

- вероятность того, что истинное значение оцениваемого параметра θ_{unkn} расположено слева от нижней границы этого интервала

$$P\{\tilde{\theta}_{\text{верхн.}} \leq \theta\} = \frac{\alpha}{2}$$

- вероятность того, что истинное значение оцениваемого параметра θ_{unkn} расположено справа от верхней границы этого интервала слева.



Что такое Доверительный Интервал (ДИ)?

- $100(1 - \alpha)\%$ -й Доверительный Интервал (ДИ) **есть** такой интервал, который с заданной доверительной вероятностью
 - $(1 - \alpha)$
- **накрывает** искомое оцениваемое значение параметра.
- Принципиально важно понимать, что ДИ является **случайным**.
- Это означает, что от опыта к опыту его границы будут колебаться, варьировать.



Надежность доверительных интервалов (ДИ)

Уровень значимости α	Уровень доверия $100(1 - \alpha)\%$	Степень надёжности ДИ
0,05	95%	Низкая
0,01	99%	Средняя
0,001	99,9%	Высокая



Точность и надежность статистических оценок

- Чем уже ДИ, тем оценка точнее.
- Поэтому показателем точности интервальной оценки может служить ширина доверительного интервала $W = \varphi_U - \varphi_L$.
- Чем больше доверительная вероятность ($1 - \alpha$), тем оценка надежнее.
- Однако ДИ с доверительной вероятностью 100% - бессмыслен.
- Например, 100%-й ДИ для доли будет содержать все значения в границах от 0 до 1: $[0, 1]$.



ДИ

- Итак, например, 99%-й ДИ означает, что если мы многократно повторим наши наблюдения, то в 99% случаев получаемые интервалы накроют (неизвестное) значение оцениваемой вероятности P (или иного параметра θ).
- Но в 1% случаев мы можем «промахнуться»: вычисленный интервал не накроет искомое значение.
- И промах этот может случиться как раз с данной конкретной выборкой.
- Отсюда мы приходим к выводу о неизбежной необходимости многократно повторять опыты (или наблюдения).



Распознавание и предсказание

- Не существует единого количественного показателя, с помощью которого можно было бы измерить способность диагностического теста **распознавать и предсказывать** наличие или отсутствие данной болезни.
- Показатели качества диагностического теста удобно разделить на
 - *показатели его **распознавательной способности** и*
 - *показатели его **предсказательной способности**,*
 - Они образуют две пары противоположностей:
 - *показатели и **контр-показатели**.*
- Только сопоставляя эти противоположности можно получить их содержательные интерпретации.



Основные вероятностные показатели распознавательной способности диагностического теста

- Четыре показателя распознавательной способности образуют две пары противоположностей:
- *Чувствительность Se*
- *и контр-чувствительность $coSe = 1 - Se$.*
- *Специфичность $Sр$*
- *и контр-специфичность $coSp = 1 - Sp$.*
- *Они являются условными вероятностями*



Что такое Чувствительность диагностического теста?

- Чувствительность *Se* (от англ. Sensitivity) есть вероятность «позитива» у субъекта с болезнью:
 - $P(T+ | D+) = P(T+, D+) / [P(D+)]$
- Она является вероятностным показателем способности положительных результатов диагностического теста («позитивов») распознавать («чувствовать») наличие болезни.



Что такое Специфичность диагностического теста?

- Специфичность S_p (от англ. Specificity) есть вероятность «негатива» у субъекта без болезни:
 - $P(T- | D-) = P(T-, D-) / [P(D-)]$
- Она является вероятностным показателем способности отрицательных результатов диагностического теста («негативов») распознавать отсутствие болезни.



Сводка определений основных показателей распознавательной способности проверяемого диагностического теста (ПДТ)

Проверяемый тест (ПДТ)		Эталон сравнения	
		Результат:	
		[+]	[-]
Результат:	[+]	Чувствительность – вероятность позитива у субъекта с болезнью $Se = P(T+ D+)$	Контр-специфичность – вероятность позитива у субъекта без болезни $coSp = P(T+ D-) = 1 - Sp$
	[-]	Контр-чувствительность – вероятность негатива у субъекта с болезнью $coSe = P(T- D+) = 1 - Se$	Специфичность – вероятность негатива у субъекта без болезни $Sp = P(T- D-)$



Идеал и бесполезность

Идеальный тест

Тест	Эталон	
	[+]	[-]
[+]	$Se = 1$	$coSp = 0$
[-]	$coSe = 0$	$Sp = 1$

$$Se + Sp = 2$$

ИЛИ

$$AUC = (Se + Sp)/2 = 1$$

Бесполезный тест (пример)

Тест	Эталон	
	[+]	[-]
[+]	$Se = 0,5$	$coSp = 1$
[-]	$coSe = 1$	$Sp = 0,5$

$$Se + Sp = 1$$



Идеальный и бесполезный тесты; значение $AUC_{uninf} = 0,5$ – неинформативное

- Если тест идеальный, то
 - $Se = 1$ и $Sp = 1$,
- Т. е.
 - $Se + Sp = 2$ или $AUC = (Se + Sp) / 2 = 1$
- Если
 - $Se + Sp = 1$ или $AUC = (Se + Sp) / 2 = 0,5$,
- то тест следует признать бесполезным.
- Значение $AUC_{uninf} = 0,5$ – неинформативное.

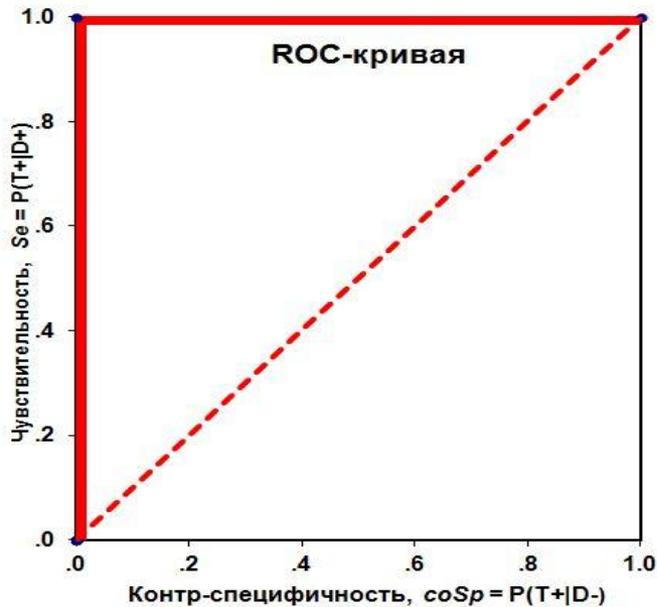


ROC-кривые и AUC

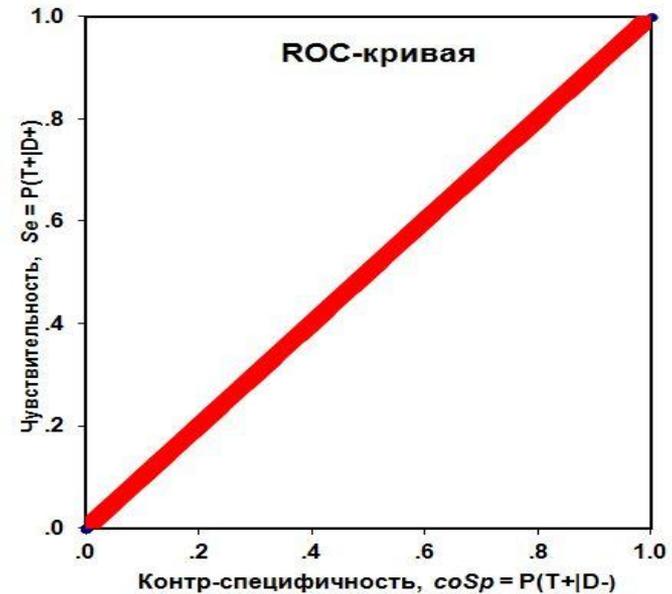
Программа:

edpsychassociates.com/Software/ROCanalysis.xls

Идеальный тест



Бесполезный тест





Бейзовские точечные оценки показателей распознавательной способности проверяемого диагностического теста (при равномерном приоре)

Проверяемый тест (ПДТ)		Эталон сравнения	
		Результат:	
		[+]	[-]
Результат:	[+]	$Se = f(T+ D+) = \frac{a+1}{a+c+2}$	$coSp = f(T+ D-) = \frac{b+1}{b+d+2}$
	[-]	$coSe = f(T- D+) = \frac{c+1}{a+c+2}$	$Sp = f(T- D-) = \frac{d+1}{b+d+2}$
Всего		1	1



Наблюдаемые точечные и интервальные статистические оценки показателей распознавательной способности маммографии

<i>Se</i>		<i>W</i>		<i>coSp = (1 - Sp)</i>			<i>W</i>	Уровень доверия
	0,84	0,89	0,05		0,031	0,033	0,002	95%
0,87	0,83	0,90	0,07	0,032	0,031	0,033	0,002	99%
	0,82	0,90	0,08		0,030	0,034	0,003	99,9%
<i>coSe = (1 - Se)</i>				<i>Sp</i>				
	0,11	0,10	0,05		0,967	0,970	0,003	95%
0,14	0,10	0,17	0,07	0,968	0,967	0,970	0,003	99%
	0,10	0,18	0,08		0,966	0,970	0,004	99,9%
1				1				

Программа DiagStat.xls

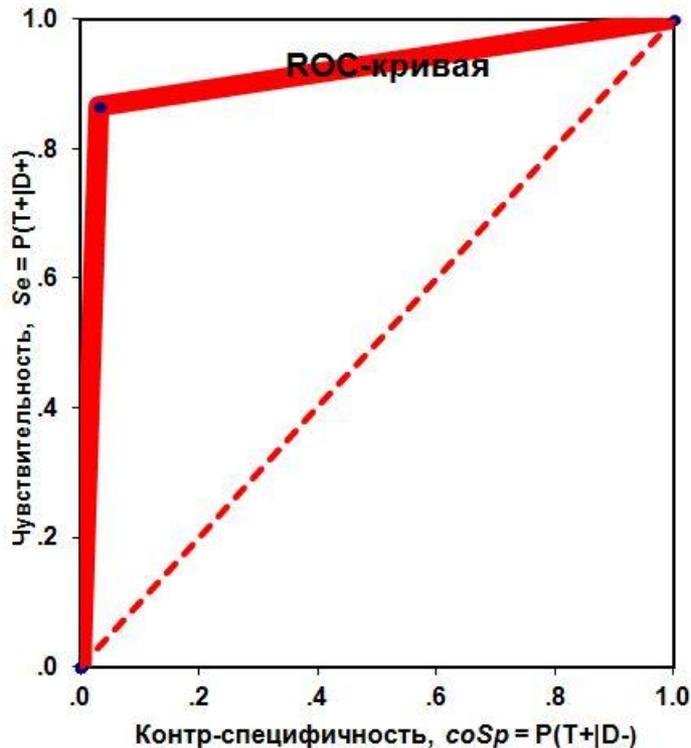


Логика использования доверительных интервалов для оценки *статистической значимости* полученных результатов

- $100(1 - \alpha)\%$ -й ДИ является интервальной оценкой истинного, неизвестного нам значения показателя (параметра) φ_{unkn} с заранее выбранной нами доверительной вероятностью $100(1 - \alpha)\%$.
- Если $100(1 - \alpha)\%$ -й ДИ не покрывает неинформативное значение оцениваемого показателя (параметра), тогда приходят к выводу, что оцениваемое значение φ_{unkn} статистически значимо отличается от неинформативного.
- При этом добавляют, что отличие значимо на выбранном уровне значимости α .
- Когда же $100(1 - \alpha)\%$ -й ДИ покрывает неинформативное значение оцениваемого показателя, то делают вывод, что оцениваемое значение φ_{unkn} статистически не отличается от неинформативного (на данном уровне значимости α).



ROC-кривая и AUC для результатов маммографии



- $AUC = (Se + Sp) / 2 = (0,87 + 0,97) / 2 = 0,92$
- с 99%-м ДИ от 0,88 до 0,95
- Компактная запись:
- $AUC = {}_{0,88}0,92_{0,95}$
- Полученный 99%-й ДИ [0,88;0,95] не содержит неинформативное значение $AUC_{uninf} = 0,5$, следовательно, оцениваемое этим интервалом значение AUC_{unkn} статистически значимо отличается от неинформативного значения 0,5 на уровне значимости $\alpha = 0,01$.



Статистическая значимость и информативность Se и Sp

- Для информативности теста в целом следует анализировать полусумму Se и Sp. Однако в некоторых случаях исследователя может интересовать чувствительность или специфичность в отдельности.
- Например, для скрининга важна чувствительность. Для ROC-анализа, при выборе точки отсечения, необходимо чтобы и Se и Sp были разумно большими.
- Для достаточно распространенных заболеваний, будем считать неинформативными для Sp и Se значение 0,5.
- В нашем примере маммологического скрининга 99,9%-й ДИ для Se, равный $[0,82;0,90]$, не покрывают



Возможные словесные интерпретации для градаций Se

Se	Способность положительных результатов диагностического теста распознавать наличие болезни:
0,0 – 0,5	Практически бесполезная
0,5 – 0,7	Низкая
0,7 – 0,9	Умеренная
0,9 – 1,0	Высокая



Возможные словесные интерпретации для градаций Sp

Sp	Способность отрицательных результатов диагностического теста распознавать отсутствие болезни:
0,0 – 0,5	Практически бесполезная
0,5 – 0,7	Низкая
0,7 – 0,9	Умеренная
0,9 – 1,0	Высокая



- Когда в нашем распоряжении имеются оценки чувствительности и специфичности маммографического обследования, мы можем сообщить пациентке лишь следующее:
- «если у Вас рак, то вероятность того, что результат Вашего маммографического обследования будет положительным, равна 87%.
- Если же у Вас нет рака, то вероятность того, что результат Вашего маммографического обследования будет отрицательным, равна 97%».



Предсказательная способность диагностического теста

- Получив у данного пациента положительный (или отрицательный) результат данного диагностического теста, мы более всего заинтересованы получить ответ на основной прагматический вопрос:
- Какова вероятность того, что у этого пациента действительно имеется (или действительно отсутствует) данная болезнь?
- Статистическая оценка такой вероятности называется **предсказательной вероятностью** положительного (или отрицательного) результата теста.
- Синонимы: предсказательность, прогностичность; предсказательная, прогнозирующая, прогностическая ценность диагностического теста, пост-тестовая вероятность,
- По-английски: **PPV** – Positive Predictive Value (of a test) и **NPV** – Negative Predictive Value (of a test).



Основные вероятностные показатели предсказательной способности диагностического теста

- **Четыре показателя предсказательной способности образуют две пары противоположностей:**
 - *предсказательность «позитивов» PPV*
 - *и их контр-предсказательность $coPPV = 1 - PPV$*
 - *предсказательность «негативов» NPV*
 - *и их контр-предсказательность $coNPV = 1 - NPV$*

Они также являются условными вероятностями



Определения основных показателей предсказательной способности проверяемого диагностического теста

Проверяемый тест (ПДТ)		Эталон сравнения	
		Результат:	
		[+]	[-]
Результат:	[+]	<p>Предсказательность позитива – вероятность наличия болезни у субъекта с позитивом</p> <p>$PPV = P(D+ T+)$</p>	<p>Контр-предсказательность позитива – вероятность отсутствия болезни у субъекта с позитивом</p> <p>$coPPV = P(D- T+) = 1 - PPV$</p>
	[-]	<p>Контр-предсказательность негатива – вероятность наличия болезни у субъекта с негативом</p> <p>$coNPV = P(D+ T-) = 1 - NPV$</p>	<p>Предсказательность негатива – вероятность отсутствия болезни у субъекта с негативом</p> <p>$NPV = P(D- T-)$</p>



Вероятностные показатели распознавательной и предсказательной способностей диагностического теста

$Se = P(T+ D+)$ Чувствительность	$P(T D)$	$coSp = 1 - Sp = P(T+ D-)$ Counter-specificity
$coSe = 1 - Se = P(T- D+)$ Контр-чувствительность		$Sp = P(T- D-)$ Специфичность
$P(T D) \neq P(D T)$		
$PPV = P(D+ T+)$ Предсказательность ПОЗИТИВОВ	$P(D T)$	$coPPV = 1 - PPV = P(D- T+)$ Контр-предсказательность ПОЗИТИВОВ
$coPPV = 1 - NPV = P(D+ T-)$ Контр-предсказательность НЕГАТИВОВ		$NPV = P(D- T-)$ Предсказательность НЕГАТИВОВ



Что такое Распространенность болезни, $Prev = P(D+)$?

Вероятность обнаружить пациента с данной болезнью в данной популяции: $P(D+)$ называется **распространенностью** болезни.

Ее **оценкой** является доля, $f(D+)$, лиц с данной болезнью в выборке объемом n из данной популяции, выявленных с помощью золотого стандарта.

Два варианта оценки распространенности:

1. *Данные демографической (медицинской) статистики.*
2. *Общепопуляционное исследование - исследование, в котором из данной популяции отбирают случайным образом большую группу людей (репрезентативную выборку).*

Оценка распространенности болезни при одновыборочном исследовании

Проверяемый метод диагностики	Эталон сравнения		Всего
	[+]	[-]	
[+]	(T+,D+) a	(T+,D-) b	(T+) a + b
[-]	(T-,D+) c	(T-,D-) d	(T-) c + d
Всего	(D+) a + c	(D-) b + d	n

Вероятность обнаружить пациента с данной болезнью в данной популяции: $P(D+)$.

В одновыборочном исследовании ее оценкой является доля лиц с данной болезнью в выборке объемом n из данной популяции: $f(D+)$, выявленных с помощью золотого стандарта.

В одновыборочном исследовании она вычисляется как отношение количества больных ($a + c$), выявленных с помощью золотого стандарта, в данной выборке, к общему числу обследованных n :

$$Prev = f(D+) = (a + c)/n$$

Или

$$Prev = f(D+) = (a + c + 1)/(n + 2)$$



Бейзовские точечные оценки показателей предсказательной способности ПДТ в одновыборочном исследовании (при равномерном приоре)

Проверяемый тест (ПДТ)		Эталон сравнения		Всего
		Результат:		
		[+]	[-]	
Результат:	[+]	$PPV = f(D+ T+) = \frac{a+1}{a+b+2}$	$coPPV = f(D- T+) = \frac{b+1}{a+b+2}$	1
	[-]	$coNPV = f(D+ T-) = \frac{c+1}{c+d+2}$	$NPV = f(D- T-) = \frac{d+1}{c+d+2}$	1



Наблюдаемые значения точечных и интервальных оценок показателей предсказательной способности маммографии и распространенности рака молочной железы

<i>PPV</i>			<i>W</i>	<i>coPPV = (1 - PPV)</i>			<i>W</i>	Всего	Уровень доверия
	0,13	0,15	0,02		0,85	0,87	0,02		
0,14	0,13	0,15	0,02	0,86	0,85	0,87	0,02	1	99%
	0,12	0,16	0,04		0,84	0,88	0,04		99,9%
<i>coNPV = (1 - NPV)</i>				<i>NPV</i>					
	0,0007	0,0010	0,0003		0,9990	0,9993	0,0003		95%
0,0008	0,0006	0,0011	0,0005	0,9992	0,9989	0,9994	0,0005	1	99%
	0,0006	0,0011	0,0005		0,9990	0,9994	0,0006		99,9%
Распространенность рака молочной железы, <i>Prev</i>				Контр-распространенность рака молочной железы, <i>coPrev</i>					
	0,006	0,007	0,001		0,994	0,995	0,001		95%
0,006	0,005	0,007	0,002	0,994	0,994	0,95	0,001	1	99%
	0,005	0,007	0,002		0,993	0,995	0,02		99,9%

Программа DiagStat.xls



Статистическая значимость отличия *PPV* от *Prev*

- В данном случае 99,9%-го ДИ для *PPV* [0,13;0,15] не пересекается с 99,9%-м ДИ для *Prev* [0,005;0,007], следовательно, оцениваемое значение *PPV* статистически значимо отличается от распространенности рака молочной железы в обследованной популяции на уровне значимости $\alpha = 0,001$.



Идеальный и малопригодный тесты в терминах PPV и NPV

- В терминах предсказательных вероятностей:
- Если тест идеальный, то
 - $PPV = 1$ и $NPV = 1$.
- Если
 - $PPV \leq 0,5$ и $NPV \leq 0,5$,
- то тест следует признать практически малопригодным.



Малоинформативные значения показателей предсказательной способности проверяемого диагностического теста (ПДТ)

Проверяемый диагностический тест (ПДТ)		Эталон сравнения		
		Результат:		Всего
		[+]	[-]	
Результат:	[+]	$PPV_{\text{lowinf}} = 0,5$	$coPPV_{\text{lowinf}} = 0,5$	1
	[-]	$coNPV_{\text{lowinf}} = 0,5$	$NPV_{\text{lowinf}} = 0,5$	1



Возможные словесные интерпретации для градаций *PPV* и *NPV*

$PPV = P(D+ T+)$	Способность «позитивов» предсказывать наличие болезни
0,0 – 0,5	Практически малопригодная
0,5 – 0,7	Низкая
0,7 – 0,9	Средняя
0,9 – 1,0	Высокая
$NPV = P(D- T-)$	Способность «негативов» предсказывать отсутствие болезни



Компактная форма записи для совместного представления точечных и интервальных оценок

- **Пример:**

- Точечная оценка чувствительности
маммографии: **0,87**

- Нижняя граница 99,9%-го доверительного
интервала (ДИ): **0,83**

- Верхняя граница 99,9%-го ДИ: **0,90**

- **Компактная запись:**

- **${}_{0,82}0,87_{0,90}$**



До и после получения позитива

- Статистическая оценка вероятности наличия рака молочной железы **до** маммографического обследования равна оценке его распространенности:

$$\bullet \text{ Prev} = f(D+) = {}_{0,005}0,006_{0,007}$$

- После получения позитива **99,9%-й ДИ** для предсказательной вероятности наличия у пациентки рака

$$\bullet \text{ PPV} = f(D+ | T+) = {}_{0,12}0,14_{0,16}$$

- не перекрывается с **99,9%-м ДИ** для распространенности.
- Следовательно, оцениваемое этим интервалом неизвестное значение **PPV** статистически высоко значимо превышает распространенность болезни.



Контр-прогностичность ПОЗИТИВОВ

- Однако оно статистически значимо **МЕНЬШЕ** неинформативного значения **0,5**, и мы приходим к выводу о **практической непригодности** положительных результатов маммографии для надежного предсказания наличия рака молочной железы.
- Этот вывод становится очевидным, если обратиться к статистической оценке контр-прогностичности позитивов, которая составила:
 - $coPPV = f(D- | T+) = (1 - PPV) = {}_{0,86}0,86_{0,88}$
- Таким образом, получается, что наличие позитива свидетельствует не столько о наличии рака молочной железы, сколько о его отсутствии.



До и после получения негатива

- Статистическая оценка вероятности отсутствия рака молочной железы до маммографического обследования равна оценке его контр-распространенности:
 - $coPrev = f(D-) = \frac{0,994}{0,993} \frac{0,995}{0,995}$ или 99,41%.
- После получения негатива оценка предсказательной вероятности отсутствия у пациентки рака статистически значимо повышается и становится равной:
 - $NPV = f(D- | T-) = \frac{0,9992}{0,9989} \frac{0,9994}{0,9994}$ или 99,92%.
- Соответственно, оценка контр-предсказательности негативов получается очень малой:
 - $coNPV = (1 - NPV) = f(D+ | T-) = \frac{0,0008}{0,0006} \frac{0,0011}{0,0011}$ или 0,08%.



Отношения правдоподобий

- Между показателями распознавательной и предсказательной способностей ПДТ имеет место взаимозависимость, которая определяется четырьмя объединяющими их показателями качества диагностического теста, называемыми *отношениями правдоподобий*.
- Они также образуют две взаимно инвертированные пары:
 - *отношение правдоподобий для «позитивов»* $LR[+]$ и его инверсия $invLR[+] = 1/LR[+]$,
 - *отношение правдоподобий для «негативов»* $LR[-]$ и его инверсия $invLR[-] = 1/LR[-]$.



Отношения правдоподобий для позитивов и негативов и их инверсии

Проверяемый диагностический тест (ПДТ)		Эталон сравнения	
		Результат:	
		[+]	[-]
Результат:	[+]	<p>Отношение правдоподобий для позитивов:</p> $LR[+] = \frac{P(T+ D+)}{P(T+ D-)}$	<p>Инверсия отношения правдоподобий для позитивов:</p> $invLR[+] = \frac{P(T+ D-)}{P(T+ D+)}$
	[-]	<p>Инверсия отношения правдоподобий для негативов:</p> $invLR[-] = \frac{P(T- D+)}{P(T- D-)}$	<p>Отношение правдоподобий для негативов:</p> $LR[-] = \frac{P(T- D-)}{P(T- D+)}$



Идеальный и бесполезный тесты в терминах LR ; значение $LR_{uninf} = 1$ – неинформативное

- Если тест идеальный, то
 - $LR[+] = \infty$ и $LR[-] = \infty$.
- Если
 - $LR[+] \leq 1$ и $LR[-] \leq 1$,
- то тест следует признать бесполезным.
- Значение $LR_{uninf} = 1$ – неинформативное.



Точечные и интервальные статистические оценки отношений правдоподобий для положительных и отрицательных результатов маммографии

LR[+]			W	LR[-]			W	Уровни доверия
	25,9	28,2	2,3		6,0	8,7	2,7	95%
27,1	25,6	28,6	3,0	7,2	5,7	9,3	3,6	99%
	25,2	29,0	3,8		5,4	10,0	4,6	99,9%

В данном случае 99,9%-е ДИ для обоих LR не покрывают неинформативные значения $LR[+]_{uninf} = LR[-]_{uninf} = 1$.

Значит, оцениваемые этими ДИ неизвестные истинные значения обоих LR статистически высоко значимо отличаются от неинформативного значения (на уровне значимости $\alpha = 0,001$).



Две интерпретации LR

- $LR[+]$ показывает, во сколько раз **чаще** позитивы будут получаться у субъектов с болезнью, чем у субъектов без болезни.
- $LR[-]$ показывает, во сколько раз **чаще** негативы будут получаться у субъектов без болезни, чем у субъектов с болезнью.
- В то же время отношение правдоподобий для позитивов $LR[+]$ показывает, во сколько раз различаются претестовые (априорные) и посттестовые (апостериорные) «**одды**» (**шансы**) в пользу наличия болезни против ее отсутствия.
- А отношение правдоподобий для негативов $LR[-]$ показывает, во сколько раз различаются претестовые (априорные) и посттестовые (апостериорные) «**одды**» (**шансы**) в пользу отсутствия болезни против ее наличия.



Что такое «шансы за/против» («одды») для случайного события?

- Два английских слова **chance** и **odds** на русский переводятся фактически одинаково: «шанс» и «шансы».
- Это создает трудности в их понимании.
- Поэтому здесь для термина **odds** предлагается использовать уточняющее словосочетание «шансы за/против».
- **Шансы за/против** события A есть отношение вероятности осуществления этого события $P(A)$ к вероятности его неосуществления $P(\bar{A})$:
 - $O(A) = P(A) / P(\bar{A}) = P(A) / [1 - P(A)]$



LR в терминах шансов за/против («ОДДОВ»)

- **LR[+]**
- **есть отношение посттестовых шансов за/против наличия болезни к претестовым шансам у субъекта с «ПОЗИТИВОМ».**
- **LR[-]**
- **есть отношение посттестовых шансов за/против отсутствия болезни к претестовым шансам у субъекта с «НЕГАТИВОМ».**



$$LR = \frac{P(B | A)}{P(B | \bar{A})} = \frac{P(A | B)}{P(\bar{A} | B)} \div \frac{P(A)}{P(\bar{A})} = \frac{O(A | B)}{O(A)}$$

$$LR[+] = \frac{P(T+ | D+)}{P(T+ | D-)} = \frac{P(D+ | T+)}{P(D- | T+)} \div \frac{P(D+)}{P(D-)}$$

$$invLR[+] = \frac{P(T+ | D-)}{P(T+ | D+)} = \frac{P(D- | T+)}{P(D+ | T+)} \div \frac{P(D-)}{P(D+)}$$

$$invLR[-] = \frac{P(T- | D+)}{P(T- | D-)} = \frac{P(D+ | T-)}{P(D- | T-)} \div \frac{P(D+)}{P(D-)}$$

$$LR[-] = \frac{P(T- | D-)}{P(T- | D+)} = \frac{P(D- | T-)}{P(D+ | T-)} \div \frac{P(D-)}{P(D+)}$$



Принятые словесные интерпретации для градаций $LR[+]$ и $LR[-]$

$LR[+]$	Повышение посттестовых шансов за/против наличия болезни у субъекта с позитивом по сравнению с претестовыми шансами за/против наличия у него болезни
1 – 3	Практически ничтожное
3 – 10	Малое
10 – 33	Среднее
33 – 100	Высокое
100 – 1000	Очень высокое
> 1000	Практически идеальное
$LR[-]$	Повышение посттестовых шансов за/против отсутствия болезни у субъекта с негативом по сравнению с претестовыми шансами за/против отсутствия у него болезни



Сводка основных полученных результатов по статистическим оценкам распознавательной и предсказательной способностей маммографии для скрининга рака молочной железы и их интерпретация

Показатель	Наблюдаемая точечная оценка с 99,9%-м ДИ	Интерпретация качества теста
<i>Se</i>	$0,82$ 0,87 $0,90$	Средняя чувствительность
<i>Sp</i>	$0,9678$ 0,968 $0,9683$	Высокая специфичность
<i>Prev</i>	$0,005$ 0,006 $0,007$	Редкое заболевание
<i>PPV</i>	$0,13$ 0,14 $0,15$	Практически непригодная предсказательность
<i>NPV</i>	$0,9989$ 0,9992 $0,9994$	Высокая предсказательность
<i>LR[+]</i>	$25,5$ 27,1 $28,4$	Умеренная информативность
<i>LR[-]</i>	$5,4$ 7,2 $10,0$	Низкая информативность

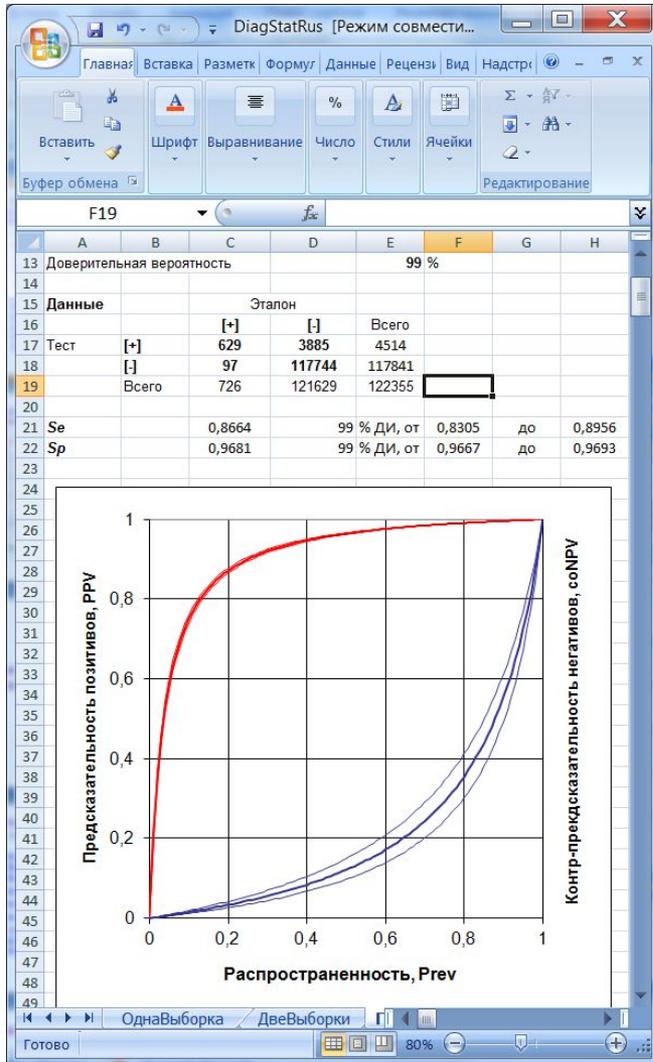


Полезное правило

- **Где и когда только возможно, следует визуализировать результаты статистического анализа, т.е. представлять их наглядно – графически.**



График предсказательностей. Программа DiagStat.xls

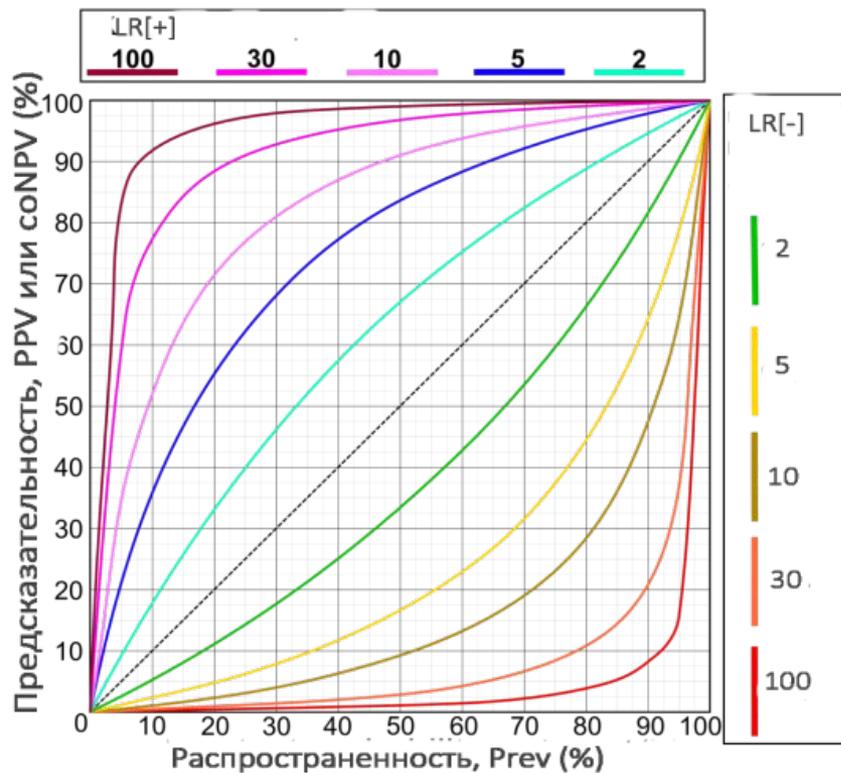
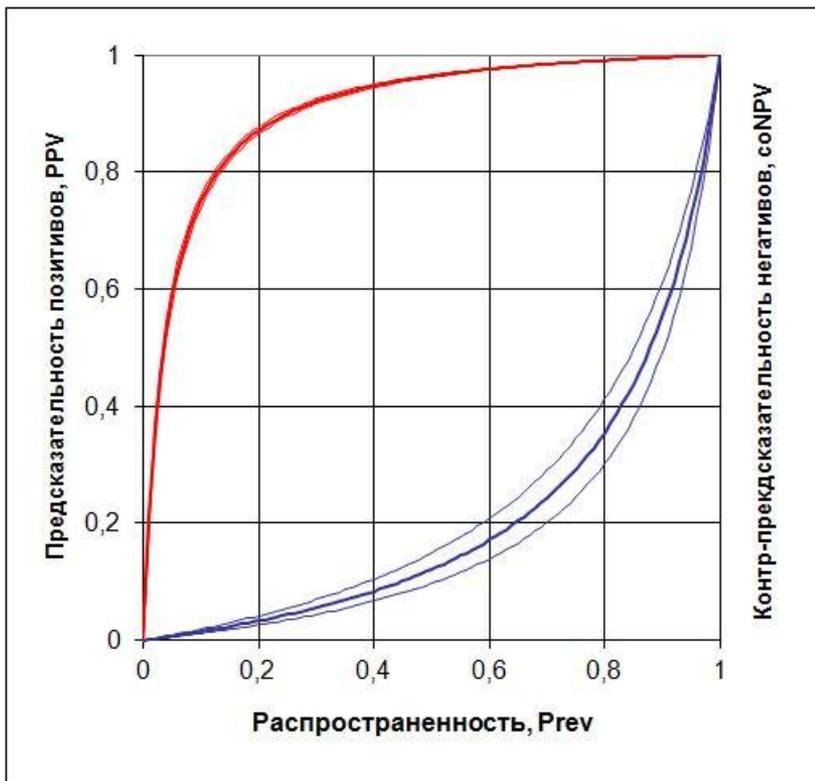


- Чем более выгнута красная кривая для PPV , тем лучше способность положительных результатов теста предсказывать наличие болезни.
- Чем более выгнута синяя кривая для $(1 - NPV)$, тем лучше способность отрицательных результатов теста предсказывать отсутствие болезни.



Сравнение графиков предсказательностей с шаблоном

[http://en.wikipedia.org/wiki/Pre- and post-test probability](http://en.wikipedia.org/wiki/Pre- and_post-test_probability)





Предсказательно-информативная распространенность заболевания PPV_{inf}

- Точечной оценкой предсказательно-информативной распространенности при заданном (фиксированном) значении планируемой предсказательности позитивов является:

$$Prev_{inf} = P(D^+ | PPV_{fix}) = \frac{PPV_{fix}}{PPV_{fix} + coPPV_{fix} \times LR[+]}$$

- Если мы хотим иметь предсказательность хотя бы
- $PPV_{fix} = 0,5$, то:

$$Prev_{inf} = \frac{1}{1 + LR[+]}$$



Предсказательно-информативные распространенности *infPrev* для планируемых значений предсказательности *planPPV*

<i>infPrev</i>	Оценка	Границы ДИ		Ширина ДИ	Уровень доверия
		Нижняя	Верхняя		
Информативная распространенность при <i>planPPV</i> = 0,7					
		0,076	0,083	0,007	95%
<i>infPrev</i> = <i>P(D+ PPV=0,7)</i>	0,079	0,076	0,084	0,008	99%
		0,075	0,085	0,010	99,9%
Информативная распространенность при <i>planPPV</i> = 0,9					
		0,24	0,26	0,02	95%
<i>infPrev</i> = <i>P(D+ PPV=0,9)</i>	0,25	0,24	0,26	0,02	99%
		0,24	0,26	0,02	99,9%

В группе риска женщин с распространенностью РМЖ *Prev* = 25% маммография будет предсказывать наличие РМЖ с вероятностью *PPV* = 90%.

Воспроизводимость результатов маммографии



<http://breastscreening.cancer.gov/data>

США, 2002 - 2006 гг.

Маммо- графия	Рак молочной железы (биопсия)		Всего
	есть, [+]	нет, [-]	
[+]	12 896 <i>a</i>	335 077 <i>b</i>	347 973 <i>a + b</i>
[-]	3 428 <i>c</i>	2 939 938 <i>d</i>	2 943 366 <i>c + d</i>
Всего	<i>a + c</i> 16 324	<i>b + d</i> 3 275 015	<i>n</i> 3 291 339

США, 2007 - 2009 гг.

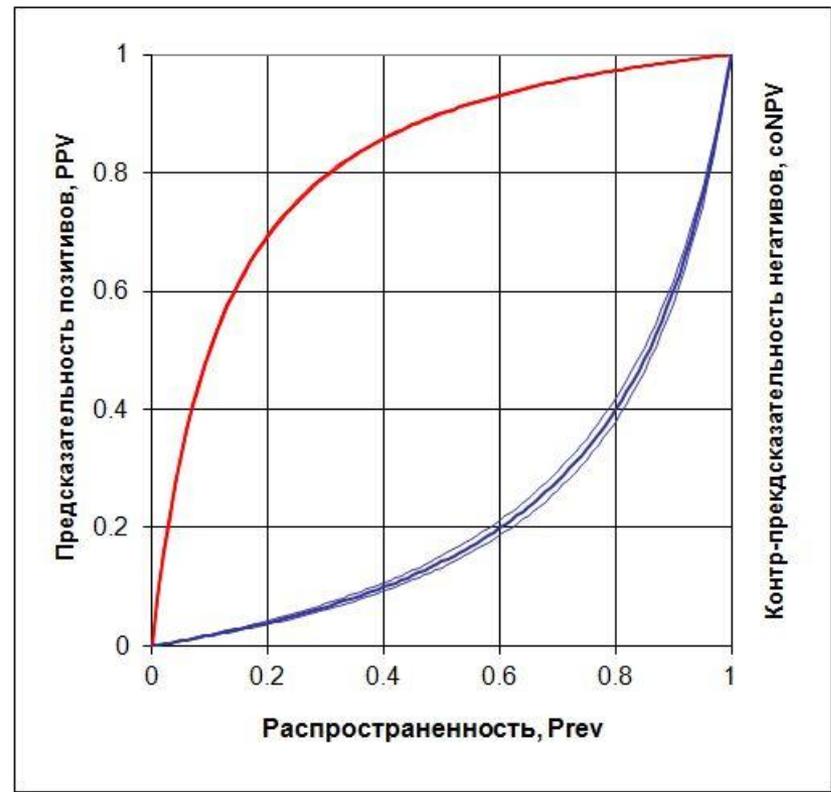
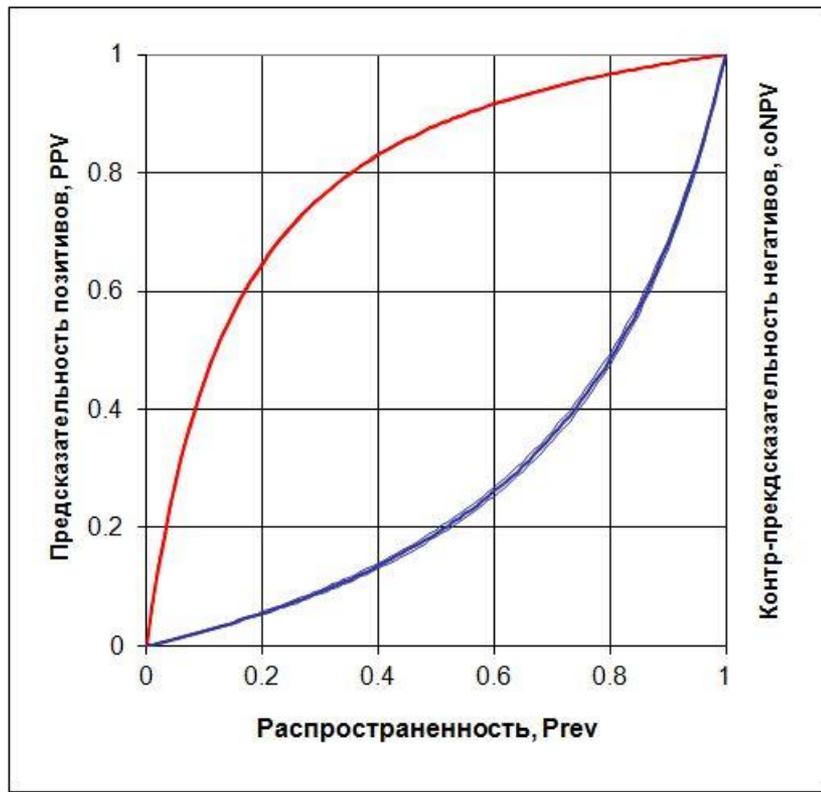
Маммо- графия	Рак молочной железы (биопсия)		Всего
	есть, [+]	нет, [-]	
[+]	5 025 <i>a</i>	177 076 <i>b</i>	182 101 <i>a + b</i>
[-]	889 <i>c</i>	1 711 592 <i>d</i>	1 712 481 <i>c + d</i>
Всего	<i>a + c</i> 5 914	<i>b + d</i> 1 888 668	<i>n</i> 1 894 582



<http://breastscreening.cancer.gov/>

США, 2002 - 2006 гг.

США, 2007 - 2009 гг.



**Двухвыборочное исследование
по
по схеме «случаи – контроли»**



Что такое исследование типа «случаи – контроли»?

- При формировании выборок по схеме **«случаи – контроли»** случайным образом отбираются **две независимые группы** субъектов:
- - группа субъектов с болезнью – **«случаи»** и
- - группа субъектов без болезни – **«контроли»**
- Такие исследования менее затратны, чем одновыборочные.
- Но для них необходимо иметь оценки распространенности заболевания, которые так или иначе получаются на основе одновыборочных исследований.



Целиакия

- **Целиакия** — нарушение пищеварения, которое вызывается некоторыми пищевыми продуктами, содержащими определённые белки — глютен (клейковина) и близкие к нему белки (авенин, гордеин и др.) в таких злаках, как пшеница, рожь, ячмень и овёс.
- У некоторых субъектов, предрасположенных к целиакии, эти белки приводят к повреждению ворсинок тонкой кишки.
- Болезнь имеет смешанный аутоиммунный, аллергический, наследственный генез, наследуется по аутосомно-доминантному типу.
- Синонимы: болезнь Ги — Гертера — Гейбнера, глютенэнтеропатия, кишечный инфантилизм (англ.: *Coeliac disease, non-tropical sprue, c(o)eliac sprue, gluten enteropathy and gluten intolerance*).

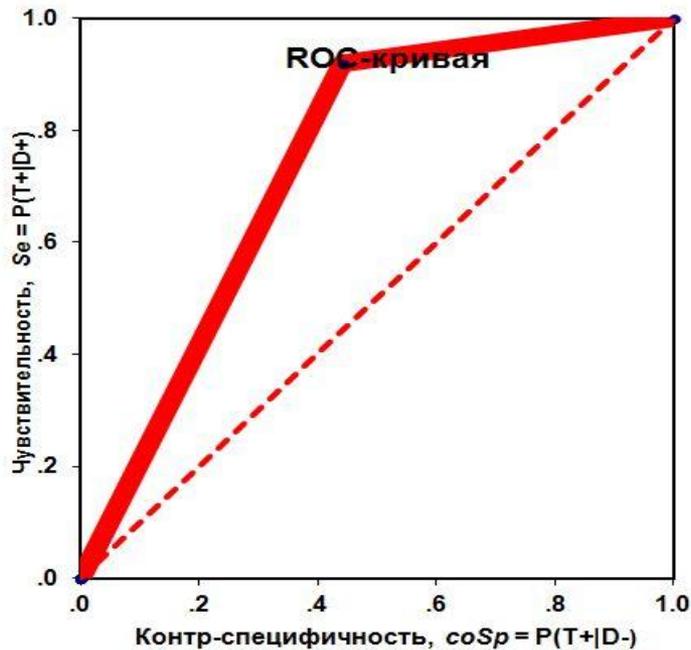


Вохмянина Н.В. Опыт *HLA*-генотипирования больных целиакией. Диагностическая значимость *HLA-DQ2* и *HLA-DQ8*. // Медицинская генетика, 2010. - Том 9, N 7. - С. 33-41.

Гаплотипы	Численности и частоты гаплотипов в группах		
	Больные целиакией II-IV степени	Контрольная группа	Всего
Предрасполагающие: DQ2 + DQ8	191 $Se = \begin{matrix} 0,86 & 0,92 & 0,95 \end{matrix}$	173 $coSp = \begin{matrix} 0,42 & 0,44 & 0,47 \end{matrix}$	364
Остальные	17 $coSe = \begin{matrix} 0,046 & 0,082 & 0,14 \end{matrix}$	219 $Sp = \begin{matrix} 0,53 & 0,56 & 0,58 \end{matrix}$	236
Всего	208	392	600



ROC-кривая и AUC для HLA-гаплотипирования целиакии



- $AUC = (Se + Sp) / 2 = (0,92 + 0,56) / 2 = 0,74$
- с 99%-м ДИ
- от 0,68 до 0,80
- Компактная запись:
- $AUC = {}_{0,68}^{0,74}_{0,80}$



Предсказательность позитива PPV и негатива NPV в двухвыборочном исследовании

- В общем виде согласно теореме Байеса (Бейза)

$$\begin{aligned} P(D+ | T+) &= \frac{P(D+) \times P(T+ | D+)}{P(D+) \times P(T+ | D+) + P(D-) \times P(T+ | D-)} = \\ &= \frac{Prev \times Se}{Prev \times Se + (1 - Prev) \times (1 - Sp)} = PPV \end{aligned}$$

$$\begin{aligned} P(D- | T-) &= \frac{P(D-) \times P(T- | D-)}{P(D-) \times P(T- | D-) + P(D+) \times P(T- | D+)} = \\ &= \frac{(1 - Prev) \times Sp}{(1 - Prev) \times Sp + Prev \times (1 - Se)} = NPV \end{aligned}$$



Распространенность целиакии

- Долгое время целиакия считалась редким заболеванием с распространенностью от 1 : 10 000 до 1 :1000.
- Однако по современным данным она считается довольно распространенным заболеванием.
- В США и в Европе (в России в частности) распространенность целиакии варьирует от 1:300 до 1:100, т.е. от **0,3%** до **1%**.
- Средняя распространенность целиакии в мире оценивается как 1 : 184, т.е. **0,54%**.



Сравнение диагностической и прогностической способностей маммографии и *HLA*-гаплотипирования

• Мамография и рак молочной железы

- $Se = {}_{0,83}^{0,87}_{0,90}$
- $Sp = {}_{0,967}^{0,968}_{0,969}$
- $LR[+] = {}_{25}^{27}_{29}$
- $LR[-] = {}_{0,11}^{0,14}_{0,17}$

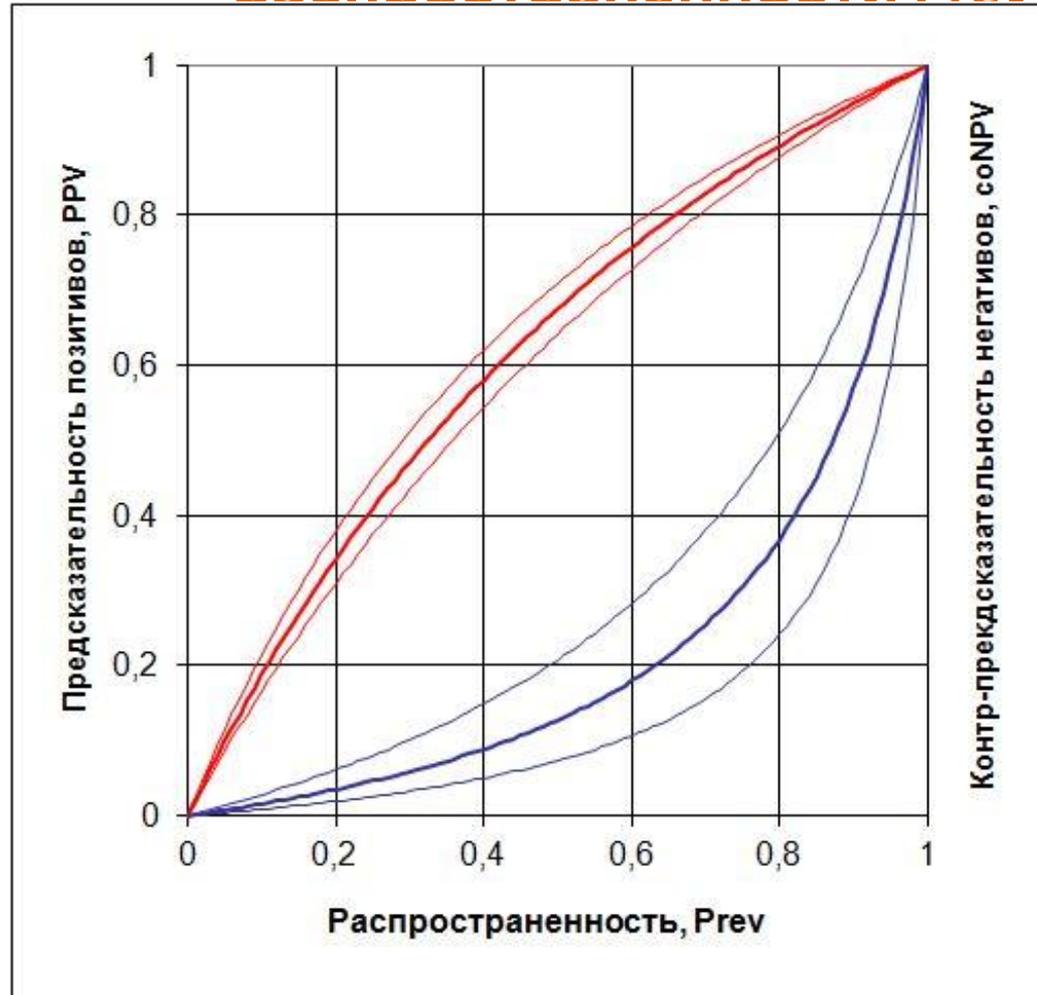
- $f(D+) = {}_{0,0054}^{0,0059}_{0,0065}$
- $PPV = {}_{0,126}^{0,139}_{0,153}$
- $NPV = {}_{0,9989}^{0,9992}_{0,9994}$

• *HLA*-гаплотипирование и целиакия

- $Se = {}_{0,86}^{0,92}_{0,95}$
- $Sp = {}_{0,53}^{0,56}_{0,58}$
- $LR[+] = {}_{1,8}^{2,1}_{2,3}$
- $LR[-] = {}_{0,08}^{0,15}_{0,26}$
- **«Внешние данные»:**
- $f(D+) = 0,0054$
- **Размах: от 0,0033 до 0,010**
- $PPV = {}_{0,007}^{0,011}_{0,18}$
- $NPV = {}_{0,9982}^{0,9992}_{0,9996}$



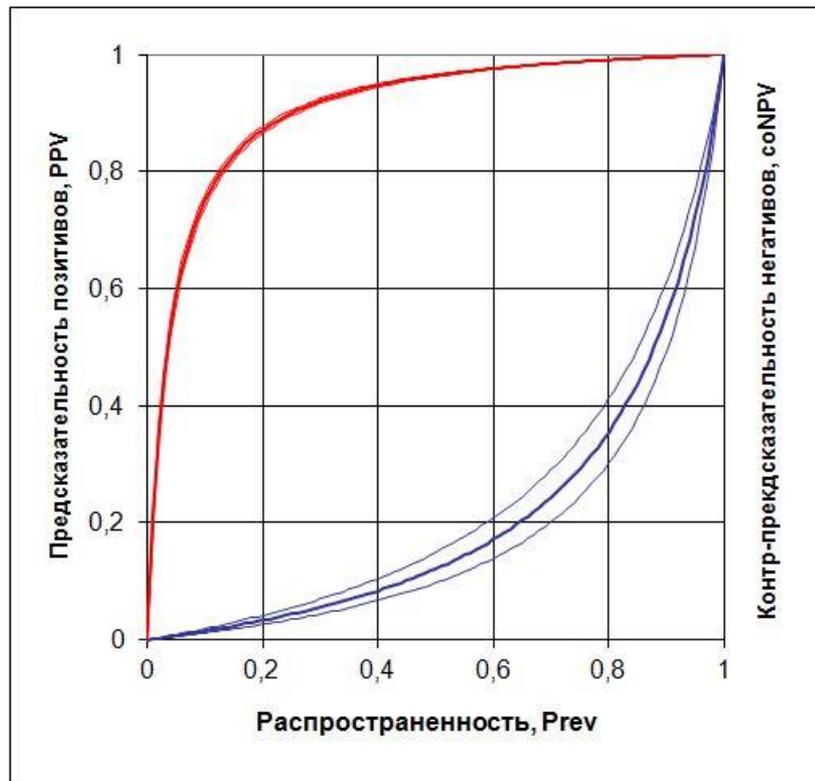
Зависимость предсказательностей PPV и $coNPV$ для HLA-гаплотипирования целиакии от ее распространенности $Prev$



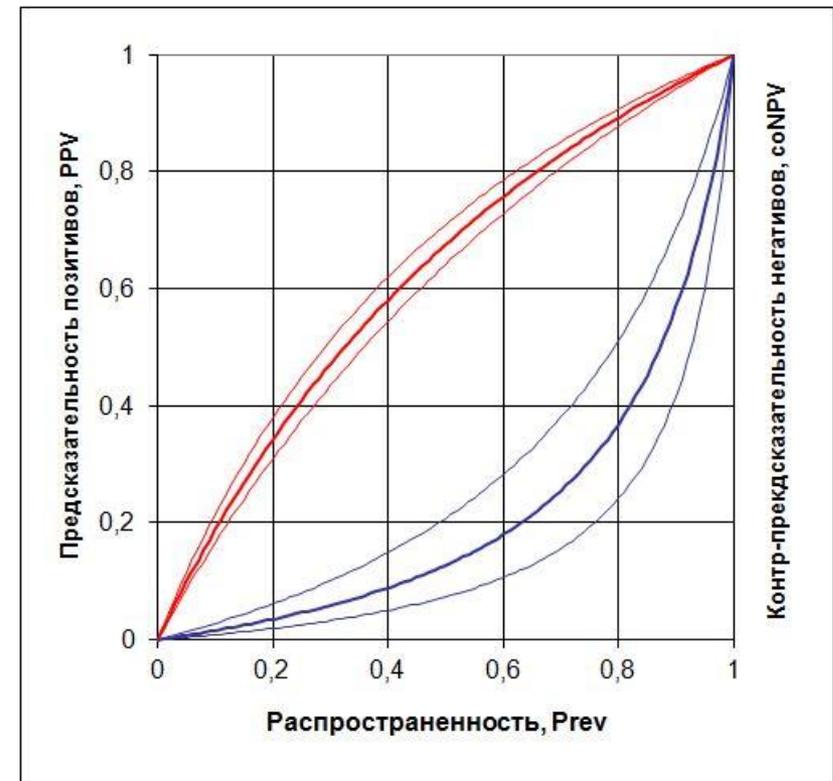


Почувствуйте разницу

Прогностичности
маммографии при раке
молочной железы



Прогностичности *HLA*-
гаплотипирования



**Воспроизводимость
результатов применения
HLA-гаплотипирования для
диагностики целиакии**



Вохмянина Н.В. Опыт HLA-генотипирования больных целиакией. Диагностическая значимость HLA-DQ2 и HLA-DQ8. // Медицинская генетика, 2010. - Том 9, N 7. - С. 33-41.

Mäki M., Mustalahti K., Kokkonen J., Kulmala P., Haapalahti M., Karttunen T., Ilonen J., Laurila K., Dahlbom I., Hansson T., Höpfl P., Knip M. Prevalence of celiac disease among children in Finland . N. Engl. J. Med. – 2003. Vol. 348. – P. 2517-2524.

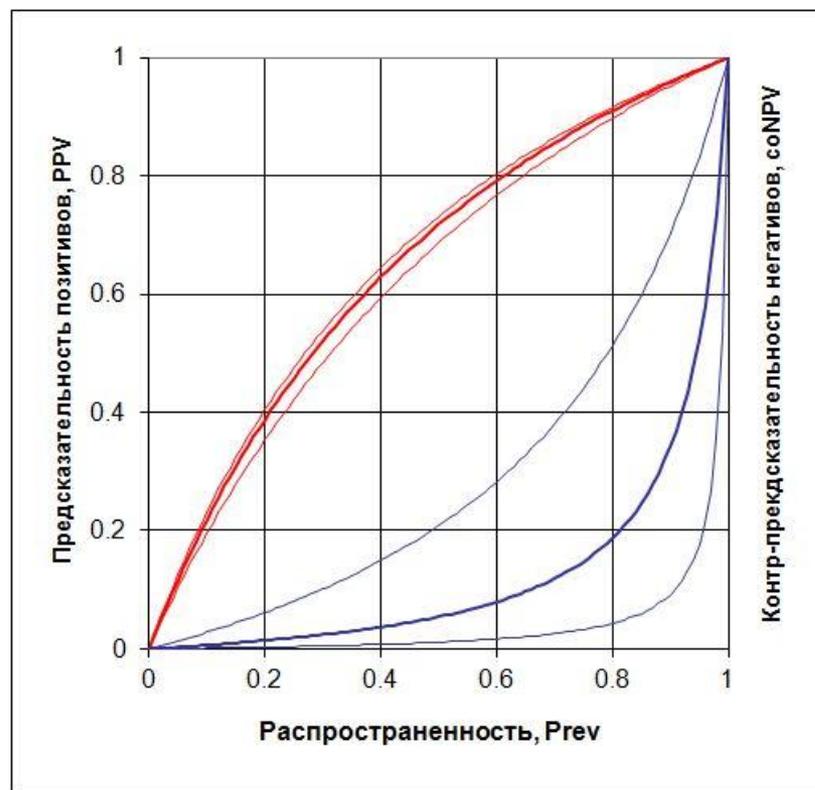
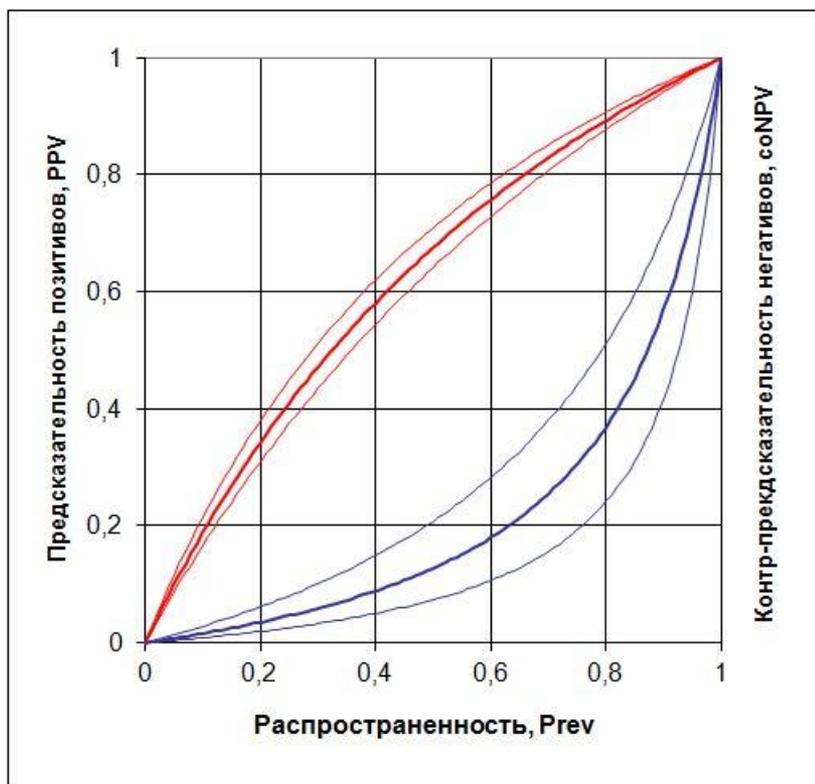
Гаплогруппы	Численности и частоты гаплогрупп в группах		
	Больные целиакией II-IV степени	Контрольная группа	Всего
DQ2 + DQ8	191	173	364
Остальные	17	219	236
Всего	208	392	600

Гаплогруппы	Тест на аутоантитела		
	Положительный	Отрицательный	Всего
DQ2+DQ8	54	1357	1465
Остальные	2	2214	2216
Всего	56	3571	3627



Вохмянина Н.В. Опыт *HLA*-генотипирования больных целиакией. Диагностическая значимость *HLA-DQ2* и *HLA-DQ8*. // Медицинская генетика, 2010. - Том 9, N 7. - С. 33-41.

Mäki M., Mustalahti K., Kokkonen J., Kulmala P., Haapalahti M., Karttunen T., Ilonen J., Laurila K., Dahlbom I., Hansson T., Höpfl P., Knip M. Prevalence of celiac disease among children in Finland . N. Engl. J. Med. – 2003. Vol. 348. – P. 2517-2524.



Галерея графиков предсказательностей

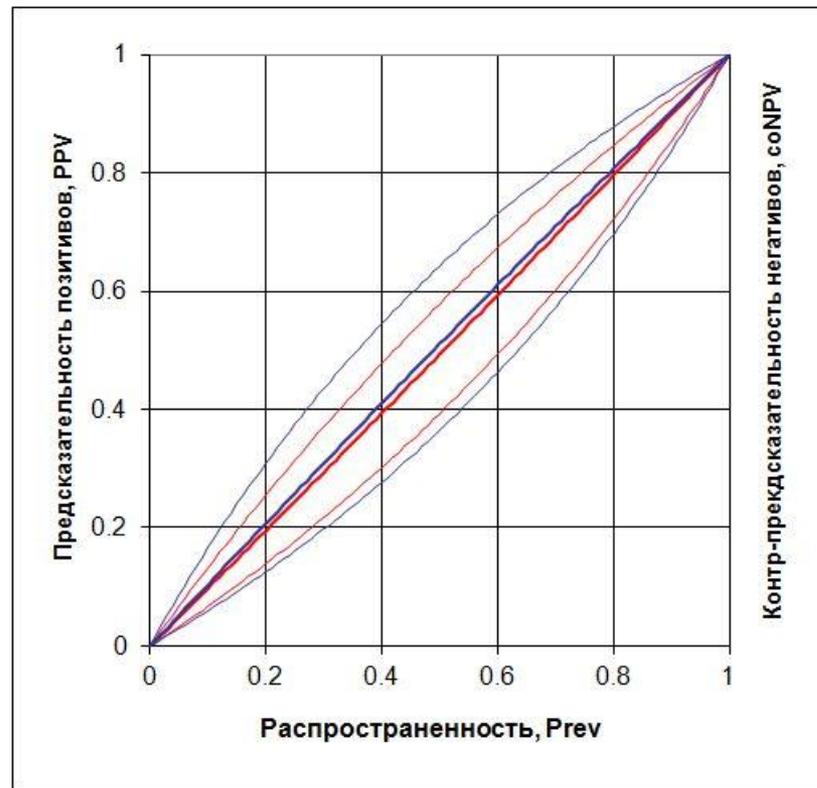
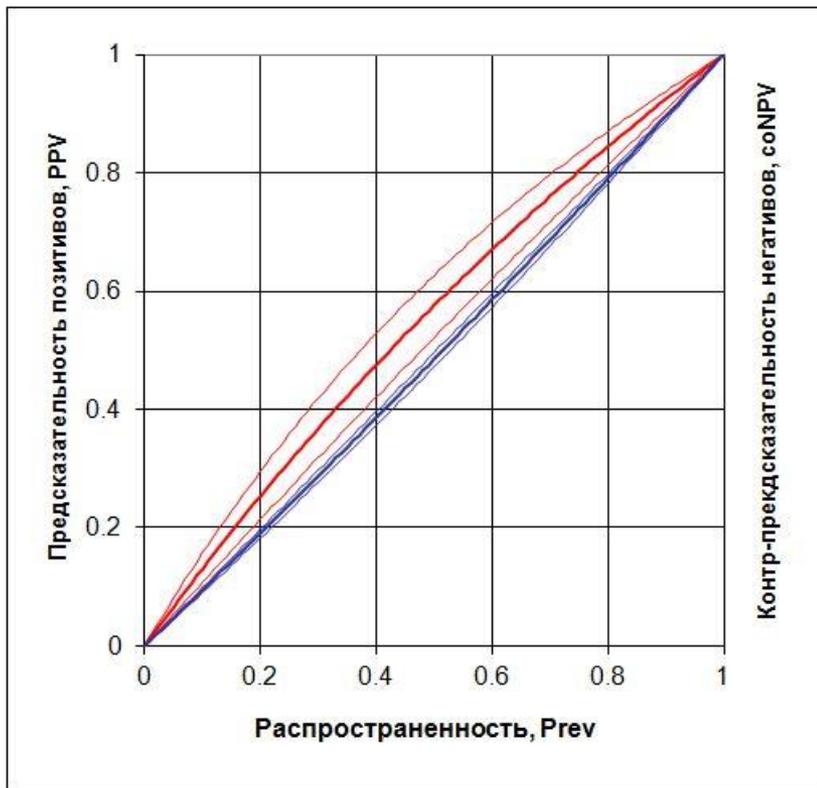


Lotufo P.A., Chae C.U., Ajani U.A., Hennekens C.H., Manson J.A.E. Male pattern baldness and coronary heart disease: The Physician's Health Study // *Arch. Intern. Medicine*, 2000. – Vol. 160. – P. 165-171.

Simon A., Worthen D. M., Mitas J. A. 1979. An evaluation of iridology // *JAMA*, V. 242, N 1, P. 1385-1389. 13.

Алопеция и ИБС

Иридодиагностика и почечная недостаточность





- Известно, что евнухи, когда они становятся евнухами в возрасте до 25 лет, не лысеют.
- Вряд ли найдется врач, который на основании этих данных будет рекомендовать молодым людям обзаводиться детьми до 25 лет, а потом становится евнухами для того, чтобы на 2% снизить риск развития у них ИБС.
- Однако это очень похоже на рекомендации медицинских генетиков, подавляющее большинство которых слишком часто бывают основаны на столь же клинически ничтожных значениях распознавательной и предсказательной способностей генетических маркеров.
- Редко значения отношения правдоподобий для позитивов в этих работах превышают значение $LR[+] > 2$.

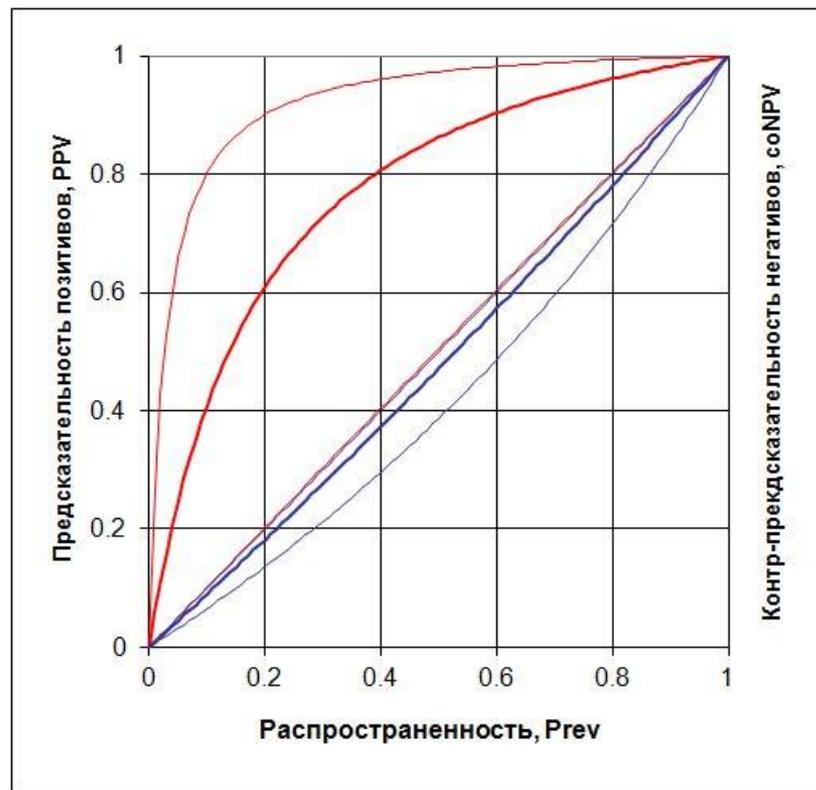
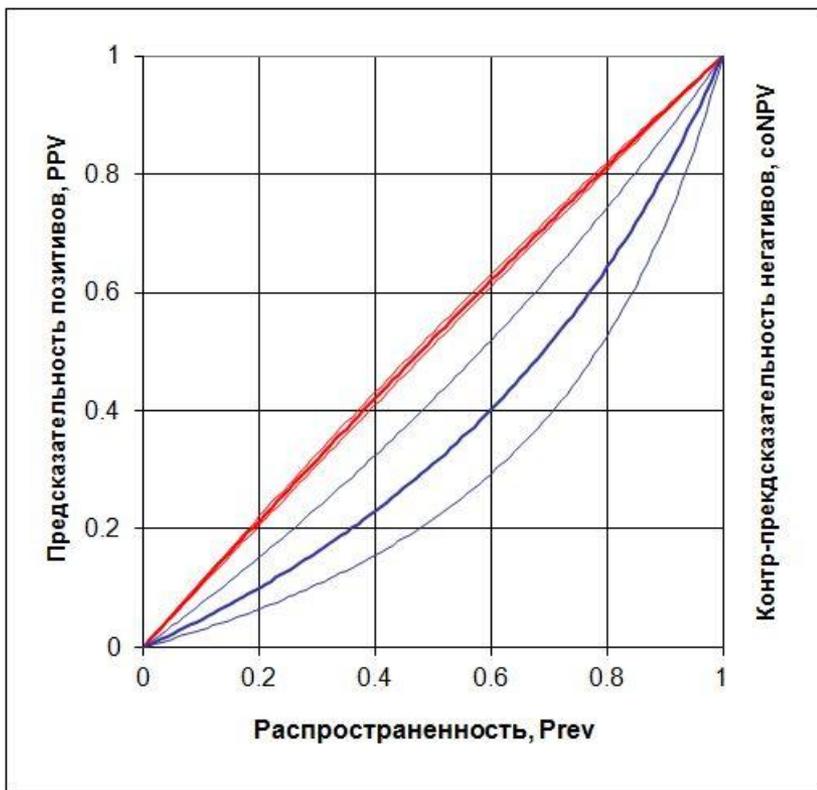


Druzhevskaya A.M, Ahmetov I.I., Astratenkova I.V., Rogozkin V.A. 2008. Association of the *ACTN3* R577X polymorphism with power athlete status in Russians // *Eur. J. Appl. Physiol.*, 2008. – Vol. 103. – P. 631–634.

Кундас Л.А., Жур К.В., Бышнеv Н.И. и др. Анализ молекулярно-генетических маркеров, ответственных за устойчивость к физическим нагрузкам, у представителей академической гребли // *Молекулярная и прикладная генетика: сб. науч. тр. Институт генетики и цитологии НАН Беларуси; (гл. ред. А.В. Кильчевский). 2013. - Минск: ГНУ «Институт генетики и цитологии НАН Беларуси», Т. 14. – С. 101-105.*

Ген *ACTN3* и элитные атлеты

Ген *PPARG* и элитные гребцы



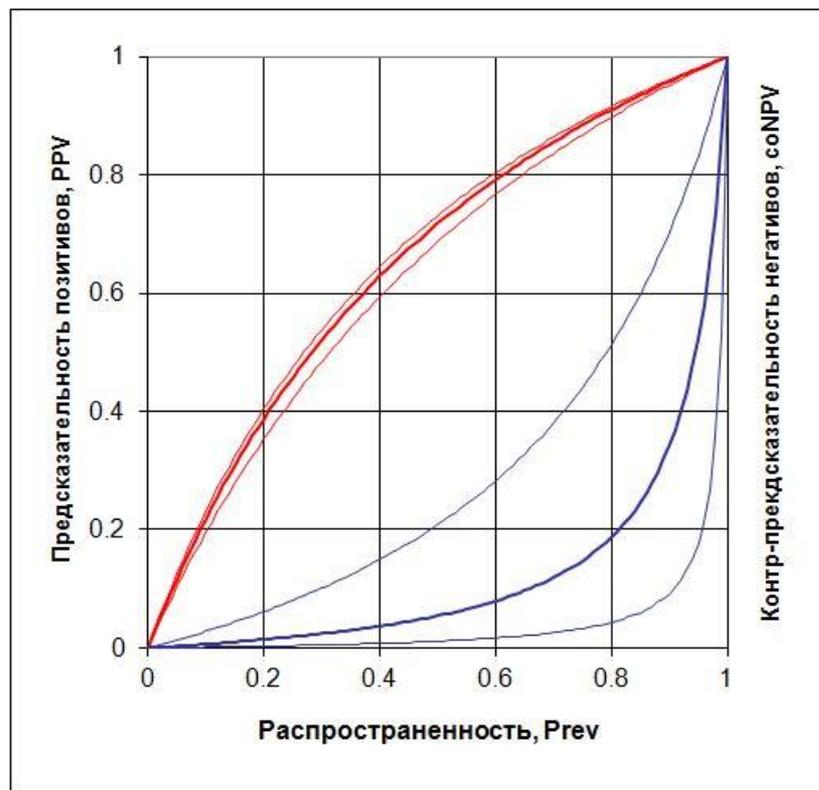
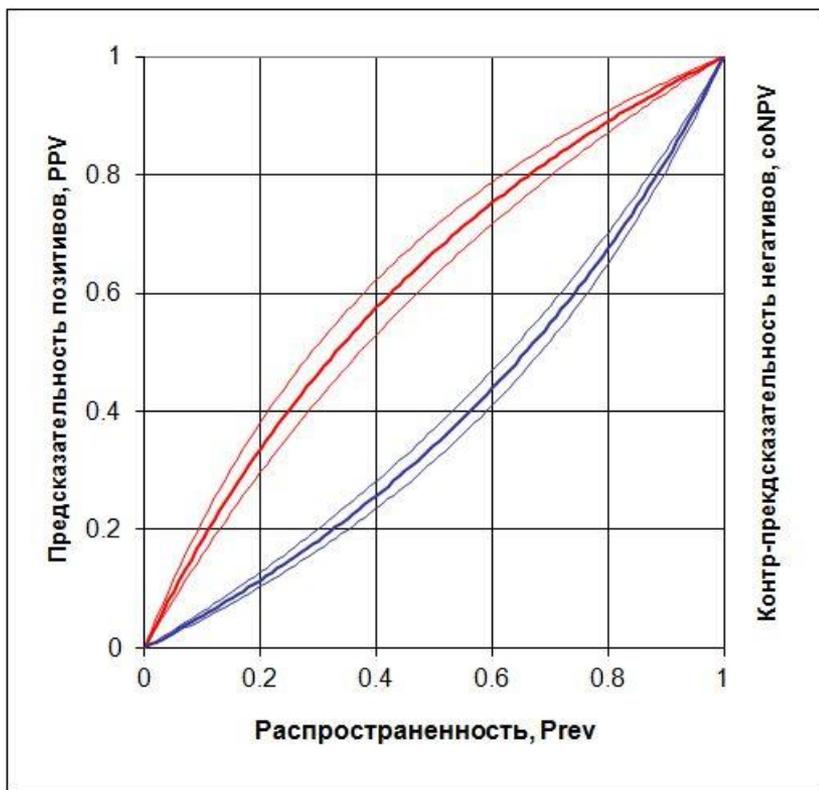


Mayeux R., Saunders A.M., Shea S., et al. Utility of the apolipoprotein E genotype in the diagnosis of Alzheimer's disease // *N. Engl. J. Med.*, 1998. – Vol. 338. – P. 506-511.

Mäki M., Mustalahti K., Kokkonen J., et al. 2003. Prevalence of celiac disease among children in Finland // *N. Engl. J. Med.*, 2003. – Vol. 348. – P. 2517-2524.

Ген АРОЕ и болезнь Альцгеймера

Галлотипы HLA и целиакия





Banks E., Reeves G., Beral V., Bull D., Crossley B., Simmonds M., Hilton E., Bailey S., Barrett N., Briers P., English R., Jackson A., Kutt E., Lavelle J., Rockall L., Wallis M.G., Wilson M., Patnick J. Influence of personal characteristics of individual women on sensitivity and specificity of mammography in the Million Women Study: cohort study // *BMJ*, 2004. - Vol. 329. – No. 7464. – P. 477-479.

Delaney K.P., Branson B.M., Uniyal A., Kerndt P.R., Keenan P.A., Jafa K., Gardner A.D., Jamieson D.J., Bulterys M. Performance of an oral fluid rapid HIV-1/2 test: experience from four CDC studies // *AIDS*, 2006. – Vol. 20. – P. 1655–1660.

Маммография и рак молочной железы

Экспресс-тест на ВИЧ, OraQuick

