

apps.uworld.com

Item 4 of 23 Question Id: 1232

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

	positive	negative		
Test positive	a True positive	b False positive	a+b	
Test negative	c False negative	d True negative	c+d	
	a+c	b+d		

→

	MI	No MI	
Test A Positive	200	50	250
Test A negative	120	80	200
	320	130	

MI = myocardial infarction.

©UWorld

Unlike specificity and sensitivity, PPV varies with disease prevalence. If disease prevalence increases, PPV increases; similarly, PPV decreases with decreasing prevalence. In this question, the patient's pre-test probability (which takes into account clinical judgment regarding how likely it is that he has an MI) is assumed to be equivalent to the prevalence of MI in the study, making the results directly translatable. Therefore, $PPV = 200 / (200+50) = 200/250 = 0.8$ or 80%. In other instances, clinicians may assume that pre-test prevalence equals disease prevalence in the population.

Educational objective:

Positive predictive value represents the probability of truly having a disease given a positive test result. It increases with increasing disease prevalence and decreases with decreasing disease prevalence.

apps.uworld.com

Item 5 of 23 Question Id: 19619

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A group of researchers wants to identify factors related to hospital-acquired bacteremia caused by methicillin-resistant *Staphylococcus aureus* (MRSA). A total of 45 patients were enrolled in the study after MRSA was isolated from a blood sample taken from them on the third or subsequent day after admission. In addition, 90 patients admitted to the hospital over the same period with a length of stay >2 days who did not have bacteremia were randomly selected. The frequency of factors such as insertion of a central line or urinary catheter and surgical site infection were then compared between the 2 groups. Which of the following is the most appropriate null hypothesis for this study?

- A. Hazard ratio is equal to 1 (4%)
- B. Hazard ratio is not equal to 1 (2%)
- C. Odds ratio is equal to 1 (52%)
- D. Odds ratio is not equal to 1 (9%)
- E. Relative risk is equal to 1 (23%)
- F. Relative risk is not equal to 1 (7%)

Omitted
Correct answer
C

52%
Answered correctly

02 secs
Time Spent

2023
Version

Explanation

A statistical hypothesis is an initial assumption (that may or may not be true) regarding population parameters

apps.uworld.com

Item 5 of 23

Question Id: 19619

Mark

Previous Next

Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A statistical hypothesis is an initial assumption (that may or may not be true) regarding population parameters (eg, mean values in two different groups) or the relationship between variables within a population (eg, association between an exposure and a disease). The **null hypothesis (H_0)** is a statement of no difference or **no association**. The alternative hypothesis (H_a) is a statement of difference or association.

In this case, patients are **initially identified** based on their **disease status** (ie, cases: patients with methicillin-resistant *Staphylococcus aureus* [MRSA] bacteremia, controls: patients without MRSA bacteremia). The **frequency of factors** (ie, exposure to insertion of a central line or urinary catheter, surgical site infection) is **subsequently compared** between cases and controls. Therefore, this is a **case-control study**; the most appropriate measure of association between an exposure and a disease in case-control studies is the **odds ratio (OR)**, which represents the odds of disease among exposed patients relative to nonexposed patients.

The OR may be interpreted as follows:

- **OR = 1.0 (null value)** indicates that the odds of exposure among cases are the same as the odds of exposure among controls; therefore, **exposure is not associated with the disease**.
- $OR \neq 1.0$ indicates that the odds of exposure are lower (ie, $OR < 1$) or higher (ie, $OR > 1$) among cases than among controls; therefore, exposure is associated with the disease.

Because H_0 is a statement of no association, the most appropriate H_0 for the study is that the $OR = 1$ (**Choice D**).

(Choices A, B, E, and F) Measures of association based on risk/incidence (eg, relative risk, hazard ratio) are calculated in cohort studies and experimental designs, in which participants are initially selected based on exposure status (ie, exposed or nonexposed to a risk factor or treatment) and then are followed over time to assess development of disease. However, in case-control studies, cases (by definition) *already have the*

apps.uworld.com

Item 5 of 23
Question Id: 19619

Mark

Previous Next

Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

appropriate measure of association between an exposure and a disease in case-control studies is the **Odds Ratio (OR)**, which represents the odds of disease among exposed patients relative to nonexposed patients.

The OR may be interpreted as follows:

- **OR = 1.0 (null value)** indicates that the odds of exposure among cases are the same as the odds of exposure among controls; therefore, **exposure is not associated with the disease**.
- $OR \neq 1.0$ indicates that the odds of exposure are lower (ie, $OR < 1$) or higher (ie, $OR > 1$) among cases than among controls; therefore, exposure is associated with the disease.

Because H_0 is a statement of no association, the most appropriate H_0 for the study is that the $OR = 1$ (**Choice D**).

(Choices A, B, E, and F) Measures of association based on risk/incidence (eg, relative risk, hazard ratio) are calculated in cohort studies and experimental designs, in which participants are initially selected based on exposure status (ie, exposed or nonexposed to a risk factor or treatment) and then are followed over time to assess development of disease. However, in case-control studies, cases (by definition) *already have the disease*. Therefore, case-control studies cannot calculate and compare the risk/incidence of disease between the case and control groups. Instead, case-control studies must compare the odds of exposure (to a risk modifier) between cases and controls.

Educational objective:

The odds ratio (OR) is a measure of association used in case-control studies. It quantifies the relationship between an exposure and a disease; its null value (ie, null hypothesis value) is always 1 (ie, $OR = 1$).

apps.uworld.com

Item 6 of 23 Question Id: 1283 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

Plasma homocysteine levels are measured in patients with acute coronary syndrome who are treated at a large community hospital. The mean plasma homocysteine level in this group is determined to be $11.1 \mu\text{mol/L}$ with a standard deviation of $1.2 \mu\text{mol/L}$. In a separate group of patients hospitalized on the general ward in the same hospital, the mean plasma level is $9.5 \mu\text{mol/L}$ and the standard deviation is $1.3 \mu\text{mol/L}$. Which of the following statistical methods should be used to compare the mean homocysteine levels of these 2 groups of patients?

- A. Two-sample t test (76%)
- B. Linear regression (1%)
- C. Correlation coefficient (4%)
- D. Chi-square test (16%)
- E. Meta-analysis (1%)

Omitted
Correct answer
A

76%
Answered correctly

01 sec
Time Spent

2023
Version

Explanation

Two-sample t test

Null hypothesis

Both samples drawn from

Alternate hypothesis

Both samples drawn from

apps.uworld.com

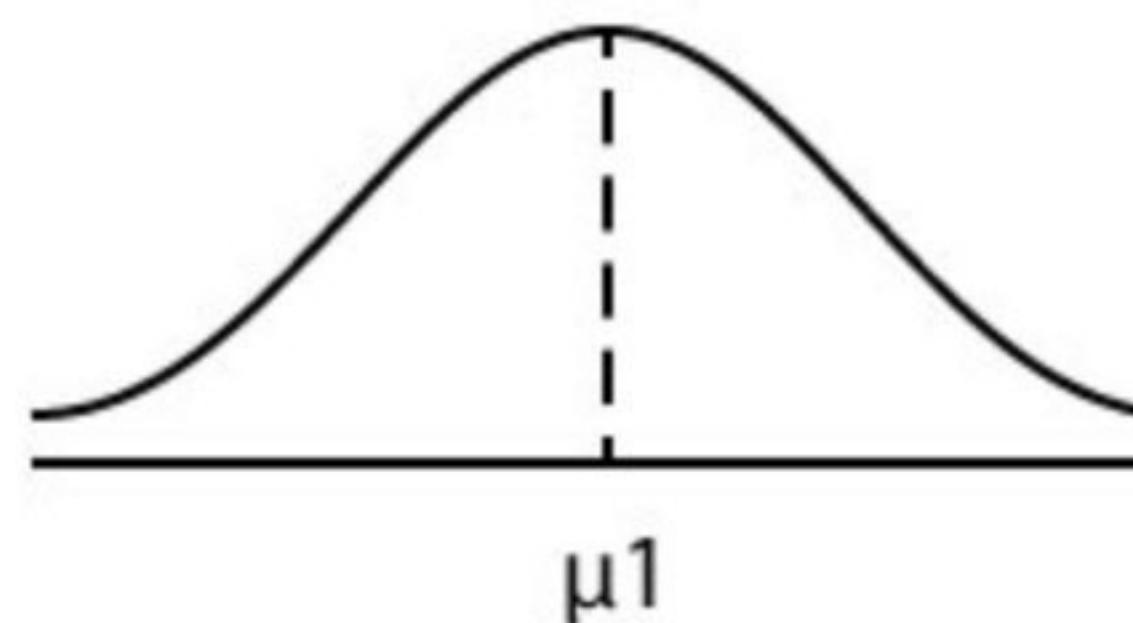
Item 6 of 23 AA Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

Mark

Two-sample t test

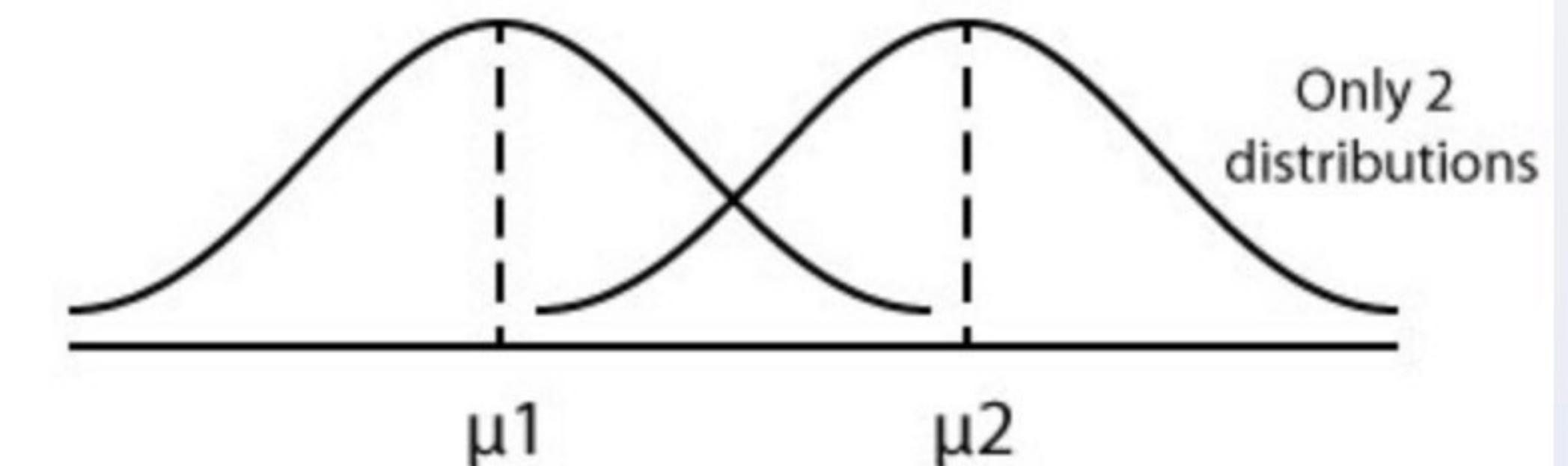
Null hypothesis

Both samples drawn from
the same population
 $(\mu_1 = \mu_2)$

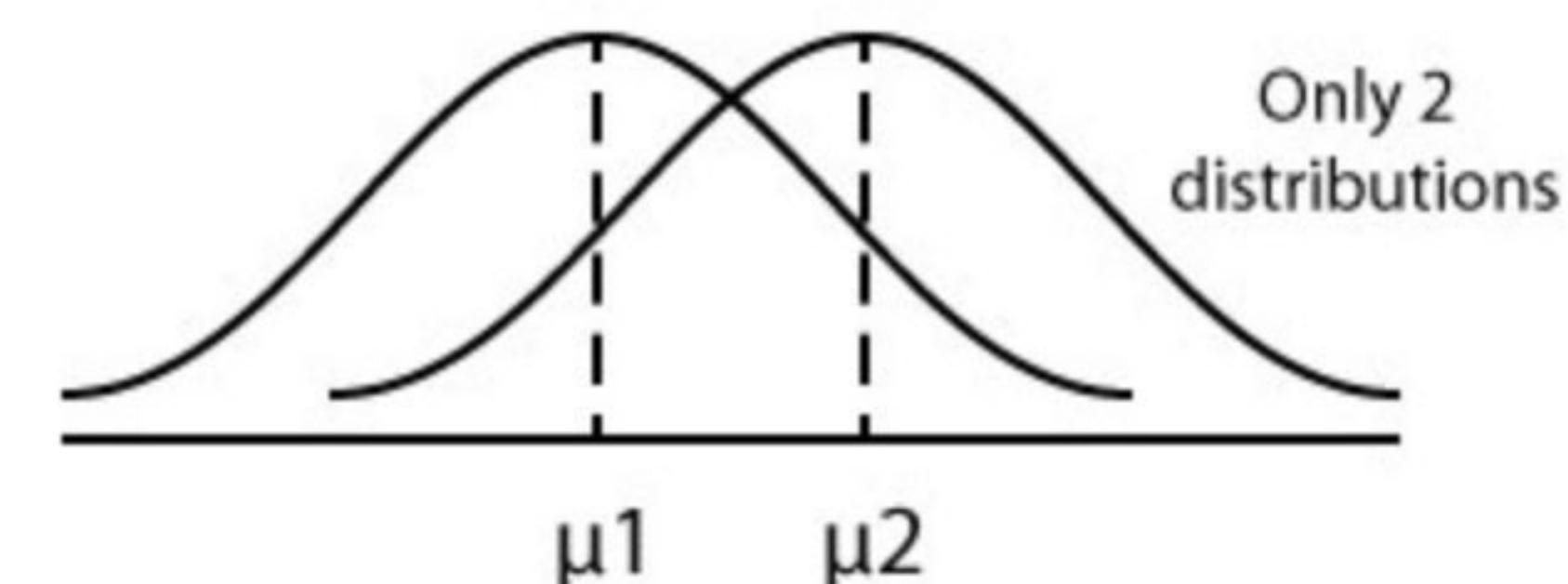


Alternate hypothesis

Both samples drawn from
different populations
 $(\mu_1 \neq \mu_2)$



Small apparent difference between the means of the samples
null hypothesis not rejected



Large apparent differences between the means of the samples
null hypothesis rejected

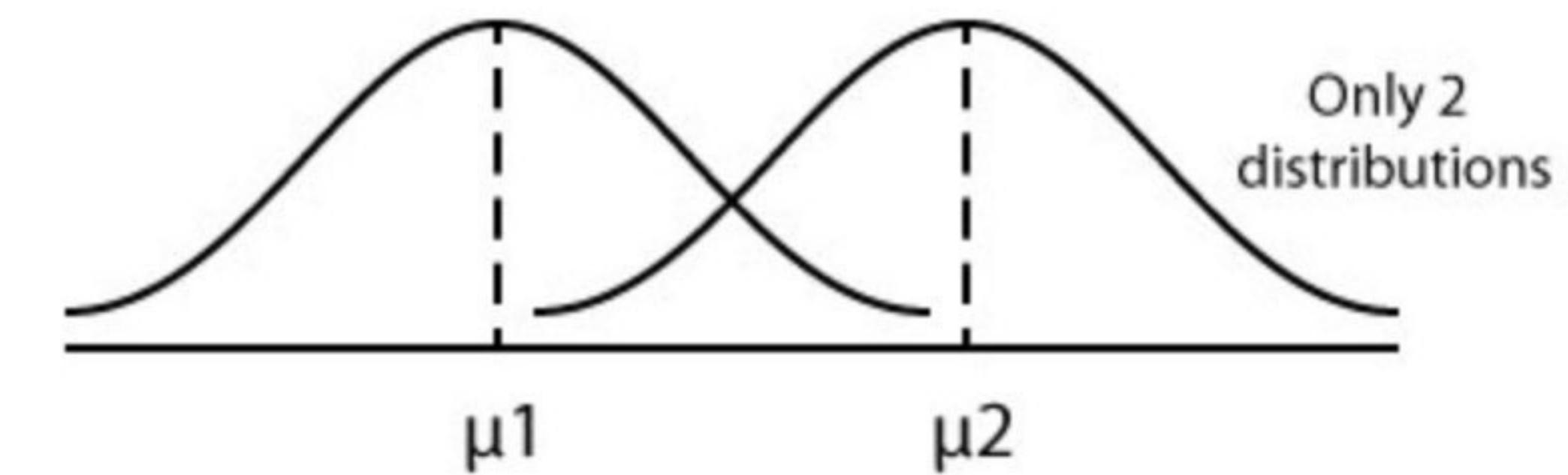


apps.uworld.com

Item 6 of 23 Question Id: 1283

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

Large apparent differences between the means of the samples
null hypothesis rejected



©UWorld

The two-sample t test is commonly employed to determine if the means of 2 populations are equal. Several statistical approaches can be used, but the basic numerical requirements needed to perform this test are the 2 mean values, the sample variances (eg, standard deviations), and the sample sizes. The t statistic is then calculated, from which the p value can be determined. If $p < 0.05$, the null hypothesis (which assumes that there is no difference between 2 groups) is rejected and the 2 means are assumed to be statistically different.

(Choice B) Linear regression is used to model the linear relationship between a dependent variable and an independent variable. For example, linear regression could be used to determine the relationship (described in terms of a trend line) between the number of cigarettes smoked per day and the number of yearly hospitalizations in COPD patients.

(Choice C) The correlation coefficient is a measure of the strength and direction of a linear relationship between 2 variables. For example, a study may report a correlation coefficient describing the association between estrogen levels and breast cancer risk in postmenopausal women. It is different from linear regression in that a single number is reported describing the strength and magnitude of the association.

(Choice D) The chi-square test is most appropriate for use with categorical data. It can be used to evaluate

apps.uworld.com

Item 6 of 23

Question Id: 1283

Mark

Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

calculated, from which the p value can be determined. If $p < 0.05$, the null hypothesis (which assumes that there is no difference between 2 groups) is rejected and the 2 means are assumed to be statistically different.

(Choice B) Linear regression is used to model the linear relationship between a dependent variable and an independent variable. For example, linear regression could be used to determine the relationship (described in terms of a trend line) between the number of cigarettes smoked per day and the number of yearly hospitalizations in COPD patients.

(Choice C) The correlation coefficient is a measure of the strength and direction of a linear relationship between 2 variables. For example, a study may report a correlation coefficient describing the association between estrogen levels and breast cancer risk in postmenopausal women. It is different from linear regression in that a single number is reported describing the strength and magnitude of the association.

(Choice D) The chi-square test is most appropriate for use with categorical data. It can be used to evaluate whether the expected frequency of an occurrence is consistent with the observed frequency of that occurrence ("goodness of fit"). For instance, a study evaluating Mendelian inheritance of red and green seed colors would use a chi-square test to compare the observed and expected proportions of each seed type.

(Choice E) Meta-analysis is an epidemiologic method of analyzing pooled data from several studies, thereby increasing the statistical power beyond that of the individual studies.

Educational objective:

The two-sample t test is a statistical method commonly employed to compare the means of 2 groups of subjects.

Biostatistics

Subject

Biostatistics & Epidemiology

System

Statistical tests

Topic

Copyright © UWorld. All rights reserved.

apps.uworld.com

Item 7 of 23 Question Id: 1284

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A research laboratory develops a new serologic test for detecting prostate cancer. The new assay is compared to transrectal ultrasound-guided prostate biopsy to determine sensitivity and specificity parameters. It is found that the test result is negative in 95% of patients who do not have the disease. If the new assay is used on 8 blood samples taken from patients without prostate cancer, what is the probability of all 8 test results coming back negative?

- A. 0.05×8 (3%)
- B. 0.95×8 (33%)
- C. 0.05^8 (3%)
- D. 0.95^8 (35%)
- E. $1 - 0.05^8$ (10%)
- F. $1 - 0.95^8$ (12%)

Omitted
Correct answer
D

35%
Answered correctly

07 secs
Time Spent

2023
Version

Explanation

Each of the 8 blood sample results is an **independent event** (ie, one patient's result has no impact on another's) with a 0.95 (95%) probability of correctly testing negative and a 0.05 (5%) probability of incorrectly testing positive. To calculate the chance of all 8 tests testing negative, **multiply the probability** of each test returning a

apps.uworld.com

Item 7 of 23
Question Id: 1284

Mark

Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

Each of the 8 blood sample results is an **independent event** (ie, one patient's result has no impact on another's) with a 0.95 (95%) probability of correctly testing negative and a 0.05 (5%) probability of incorrectly testing positive. To calculate the chance of all 8 tests testing negative, **multiply the probability** of each test returning a negative result:

$$0.95 \times 0.95 \times 0.95 \times 0.95 \times 0.95 \times 0.95 = 0.95^8$$

(Choices A and B) The value 0.05×8 is equivalent to adding the probability of a sample testing positive 8 times, and 0.95×8 is equivalent to adding the probability of a sample testing negative 8 times. However, the combined probability for a series of independent events is calculated using multiplication, not addition. For instance, when flipping a coin twice the probability of getting 2 heads in a row is not $0.5 + 0.5 = 0.5 \times 2 = 1$ (100%) but rather $0.5 \times 0.5 = 0.25$ (25%).

(Choice C) The probability of all 8 samples incorrectly testing positive is given by 0.05^8 .

(Choices E and F) Overall, 2 outcomes are possible: either all 8 samples will return the same test result or at least 1 sample will test differently. Because the total probability must add up to 1 (100%) and the probability that all 8 samples will return negative is given by 0.95^8 , the probability of at least 1 sample testing positive is $1 - 0.95^8$. Conversely, the probability that all 8 samples will return positive is 0.05^8 while the probability of at least 1 sample testing negative is $1 - 0.05^8$.

Educational objective:

If events are independent, the probability that all events will turn out the same is the product of the separate probabilities for each event. The probability of at least 1 event turning out differently is given as $1 - P(\text{all events being the same})$.

apps.uworld.com

Item 8 of 23 Question Id: 1271

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

Researchers are developing an enzyme-linked immunosorbent assay test for diagnosing rheumatoid arthritis. The test is designed to detect the presence of serum antibodies against citrullinated proteins. Two test populations with a differing prevalence of rheumatoid arthritis are selected. The researchers plan to assess the test's performance in the 2 populations by comparing a number of diagnostic test parameters. Which of the following performance measures is most likely to be different between the 2 test populations?

- A. Negative likelihood ratio (1%)
- B. Positive likelihood ratio (8%)
- C. Positive predictive value (80%)
- D. Sensitivity (6%)
- E. Specificity (3%)

Omitted
Correct answer
C

80%
Answered correctly

02 secs
Time Spent

2023
Version

Explanation

Sensitivity
 $TP / (TP + FN)$

Diseased population

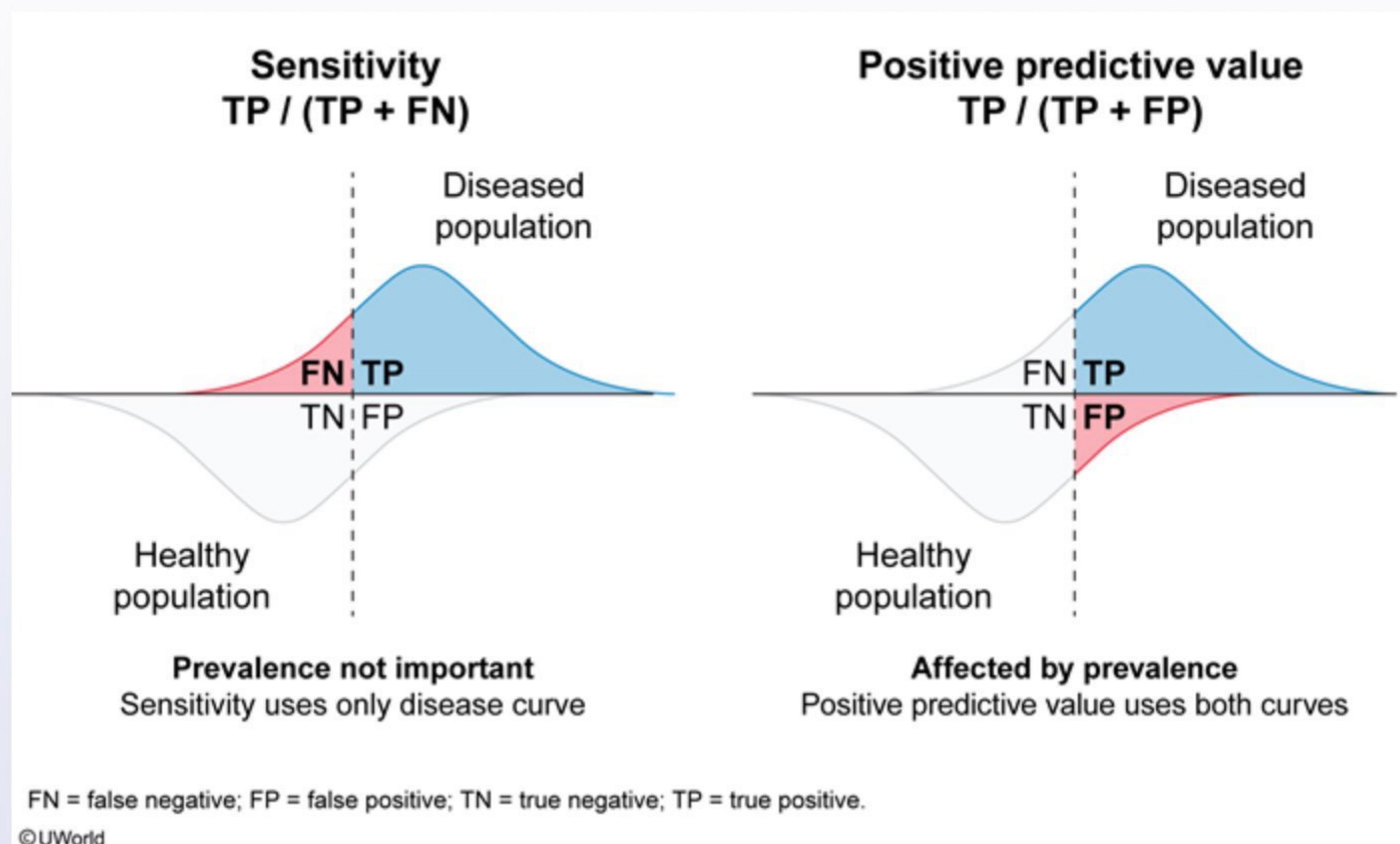
Positive predictive value
 $TP / (TP + FP)$

Diseased population

apps.uworld.com

Item 8 of 23 Question Id: 1271 Explanations

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings



Predictive values are performance measures of diagnostic tests that are **dependent on the prevalence** of disease in a population of interest. The positive predictive value (**PPV**) is the probability that someone who tests positive actually has the disease. It is calculated by dividing the number of true-positive results by the total number of positive results (ie, $\text{TP} / [\text{TP} + \text{FP}]$). The number of true-positive results depends on the sensitivity, and the number of false-positive results depends on the specificity; the relative proportion of each is determined by the prevalence of disease in the population.

Populations with a lower disease prevalence have fewer true-positive results and higher numbers of false

Item 8 of 23
Question Id: 1271

Mark

Previous Next

Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

by the prevalence of disease in the population.

Populations with a lower disease prevalence have fewer true-positive results and higher numbers of false positives, so the PPV decreases. As disease [prevalence increases](#), the number of true positives also increases, while the number of false positives decreases, resulting in a **higher PPV**. Similarly, the negative predictive value increases as the disease prevalence decreases.

(Choices A and B) Positive and negative likelihood ratios indicate how a particular positive or negative test result influences the pretest probability of having a disease. Likelihood ratios >1 indicate that the test result is associated with the presence of the disease; likelihood ratios <1 mean that the test result is associated with the absence of the disease. Because positive and negative likelihood ratios are based on sensitivity and specificity, they are not affected by disease prevalence.

(Choices D and E) Sensitivity and specificity are not affected by disease prevalence. This is because sensitivity is calculated using true positives and false negatives (only people with the disease), and specificity is calculated using true negatives and false positives (only people without the disease). In this case, the same test (with the same sensitivity and specificity characteristics) has been used on 2 different populations with a differing prevalence.

Educational objective:

Various parameters are used to evaluate the accuracy and usefulness of diagnostic tests. Positive and negative predictive values are influenced by disease prevalence in the target population; sensitivity, specificity, and likelihood ratios are not prevalence-dependent.

apps.uworld.com

Item 9 of 23 Question Id: 1278

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A 34-year-old man is admitted to the hospital with acute chest pain. An ECG is obtained in the emergency department and shows ST segment elevation in leads II, III, and aVF. A sample of blood is taken from the patient, and a new test is used to measure plasma homocysteine levels. The test is repeated 3 times with his blood sample, and the results are 11.8 µmol/L, 9.2 µmol/L, and 13.7 µmol/L (laboratory reference range: 4-14). Which of the following parameters is most likely to be low based on the results of the new test?

- A. Accuracy (12%)
- B. Precision (79%)
- C. Sensitivity (2%)
- D. Specificity (2%)
- E. Validity (3%)

Omitted

Correct answer

B



79%

Answered correctly



01 sec

Time Spent



2023

Version

Explanation

Precision & accuracy



apps.uworld.com

Item 9 of 23

Question Id: 1278

Mark

Previous Next

Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

Precision & accuracy

The image displays four target diagrams, each consisting of three concentric red rings. The bullseye is white.

- Top Left:** Several black dots are scattered across all three rings. This represents low precision (widely scattered) and low accuracy (not centered).
- Top Right:** All black dots are clustered tightly in the central white bullseye. This represents low precision (widely scattered) and high accuracy (centered).
- Bottom Left:** All black dots are clustered tightly in the outermost ring, far from the center. This represents high precision (tightly clustered) and low accuracy (not centered).
- Bottom Right:** All black dots are clustered tightly in the central white bullseye. This represents high precision (tightly clustered) and high accuracy (centered).

Low precision
Low accuracy

Low precision
High accuracy

apps.uworld.com

Item 9 of 23 Question Id: 1278

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

©UWorld

A measurement tool should be precise and accurate to be useful in the clinical setting. Precision and accuracy are defined as follows:

- **Precision** (reliability) is the ability of a test to reproduce identical or similar results with repeated measurements (bottom half of figure).
- **Accuracy** (validity) is the ability of a test to measure what it is supposed to measure. For a new test to be accurate, its results should be equivalent to the results obtained with a "gold standard" (eg, best conventional test available) on the same individual (right half of figure).

A precise test gives similar or very close results on repeat measurements. In this example, repeat measurements of the same sample yielded different results; therefore, the new test has low precision.

apps.uworld.com

Item 9 of 23
Question Id: 1278

Mark

Previous Next

Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A measurement tool should be precise and accurate to be useful in the clinical setting. Precision and accuracy are defined as follows:

- **Precision** (reliability) is the ability of a test to reproduce identical or similar results with repeated measurements (bottom half of figure).
- **Accuracy** (validity) is the ability of a test to measure what it is supposed to measure. For a new test to be accurate, its results should be equivalent to the results obtained with a "gold standard" (eg, best conventional test available) on the same individual (right half of figure).

A precise test gives similar or very close results on repeat measurements. In this example, repeat measurements of the same sample yielded different results; therefore, the new test has low precision.

(Choices A and E) To evaluate the accuracy of a new test, its results should be compared to those obtained with the gold standard test on the same individual. A test can be highly precise (ie, gives very similar results on repeat measurements) but inaccurate (ie, the measurements are all incorrect compared to the gold standard). A test that is highly imprecise is also highly inaccurate. In this case, as there were no test results obtained using the "gold standard", the accuracy/validity of the test cannot be determined.

(Choices C and D) Sensitivity is defined as a test's ability to identify the true presence of disease, whereas specificity is defined as a test's ability to identify the true absence of disease. Poor reliability can limit the sensitivity or specificity of a test if the results from healthy and diseased populations are close in value.

Educational objective:

A precise/reliable test is reproducible in that it gives similar results on repeat measurements. Reliability is maximal when random error is minimal.

apps.uworld.com

Item 10 of 23 Question Id: 1177

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

The results of a study investigating a new diagnostic test for acute myocardial infarction (MI) are given in the table below.

	MI	No MI
Test positive	75	20
Test negative	25	80

What is the sensitivity of the new diagnostic test?

- A. 25% (1%)
- B. 37.5% (0%)
- C. 50% (0%)
- D. 75% (88%)
- E. 79% (5%)
- F. 80% (2%)

Omitted
Correct answer
D

88%
Answered correctly

01 sec
Time Spent

2023
Version

Explanation

apps.uworld.com

Item 10 of 23 Question Id: 1177

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

	Positive condition	Negative condition	
Positive test result	TP	FP	$PPV = \frac{TP}{TP + FP}$
Negative test result	FN	TN	$NPV = \frac{TN}{TN + FN}$
Sensitivity = $\frac{TP}{TP + FN}$ Specificity = $\frac{TN}{TN + FP}$			
FN = false negative; FP = false positive; NPV = negative predictive value; PPV = positive predictive value; TN = true negative; TP = true positive.			

The data is presented in the standard 2×2 (contingency) table format. The **sensitivity** of a test refers to its ability to correctly identify individuals affected with a disease. It is the number of true positives divided by the total number of patients with the disease. Sensitivity should be high in screening tests in order to pick up all cases of a disease (decrease false negatives). Using the generic 2×2 (contingency) table:

Sensitivity = $TP / (TP + FN)$, where TP is true positives and FN is false negatives. In this case:

$$\text{Sensitivity} = 75 / (75 + 25) = 75/100 = 0.75 \text{ (or } 75\%)$$

(Choice A) The false negative rate is $(1 - \text{sensitivity}) = 1 - 0.75 = 0.25$ (or 25%). The false negative rate is not affected by disease prevalence.

(Choice B) The value 37.5% is obtained by dividing the TP (75) by the total number of patients in the study (75

apps.uworld.com

Item 10 of 23
Question Id: 1177

Mark

Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

ability to correctly identify individuals affected with a disease. It is the number of true positives divided by the total number of patients with the disease. Sensitivity should be high in screening tests in order to pick up all cases of a disease (decrease false negatives). Using the generic 2×2 (contingency) table:

Sensitivity = $TP / (TP + FN)$, where TP is true positives and FN is false negatives. In this case:

$$\text{Sensitivity} = 75 / (75 + 25) = 75/100 = 0.75 \text{ (or } 75\%)$$

(Choice A) The false negative rate is $(1 - \text{sensitivity}) = 1 - 0.75 = 0.25$ (or 25%). The false negative rate is not affected by disease prevalence.

(Choice B) The value 37.5% is obtained by dividing the TP (75) by the total number of patients in the study ($75 + 25 + 20 + 80 = 200$).

(Choice C) The number of patients with myocardial infarction (MI) based on the gold standard is 100 (= 75 + 25). The total number of patients in the study is 200 (= 75 + 25 + 20 + 80). Therefore, the prevalence of MI in this sample is $100/200 = 0.5$ (or 50%).

(Choice E) The positive predictive value (PPV) of the test (the probability that a person has the disease given a positive test result) is given by: $\text{PPV} = TP / (TP + FP)$, where FP is false positives. In this case, $\text{PPV} = 75 / (75 + 20) = 75/95 = 0.79$ (or 79%).

(Choice F) The test specificity (its ability to correctly identify individuals without the disease) is $80 / (80 + 20) = 80/100 = 0.80$ (or 80%).

Educational objective:

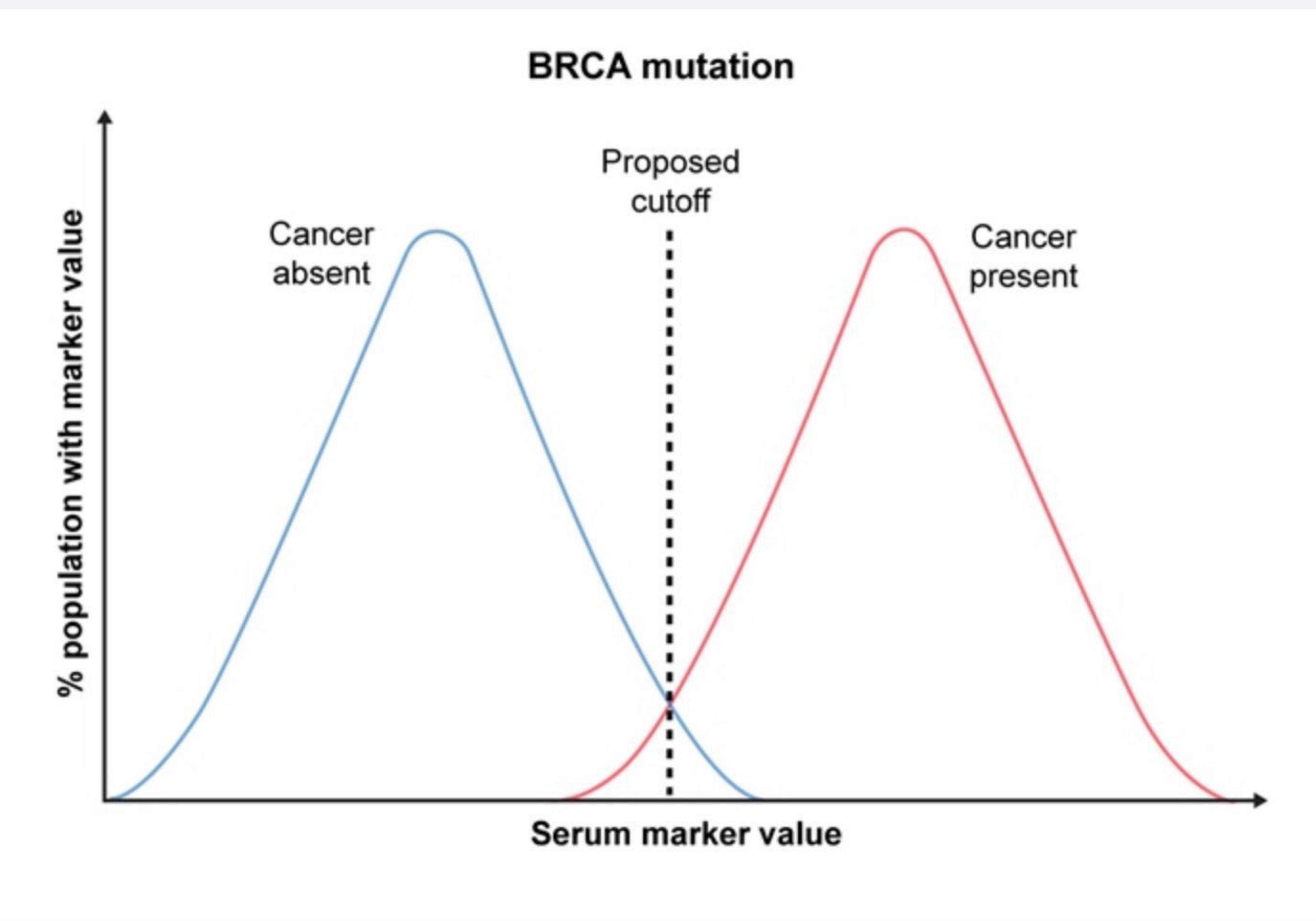
Sensitivity = true positives / (true positives + false negatives). Screening tests should have high sensitivity.

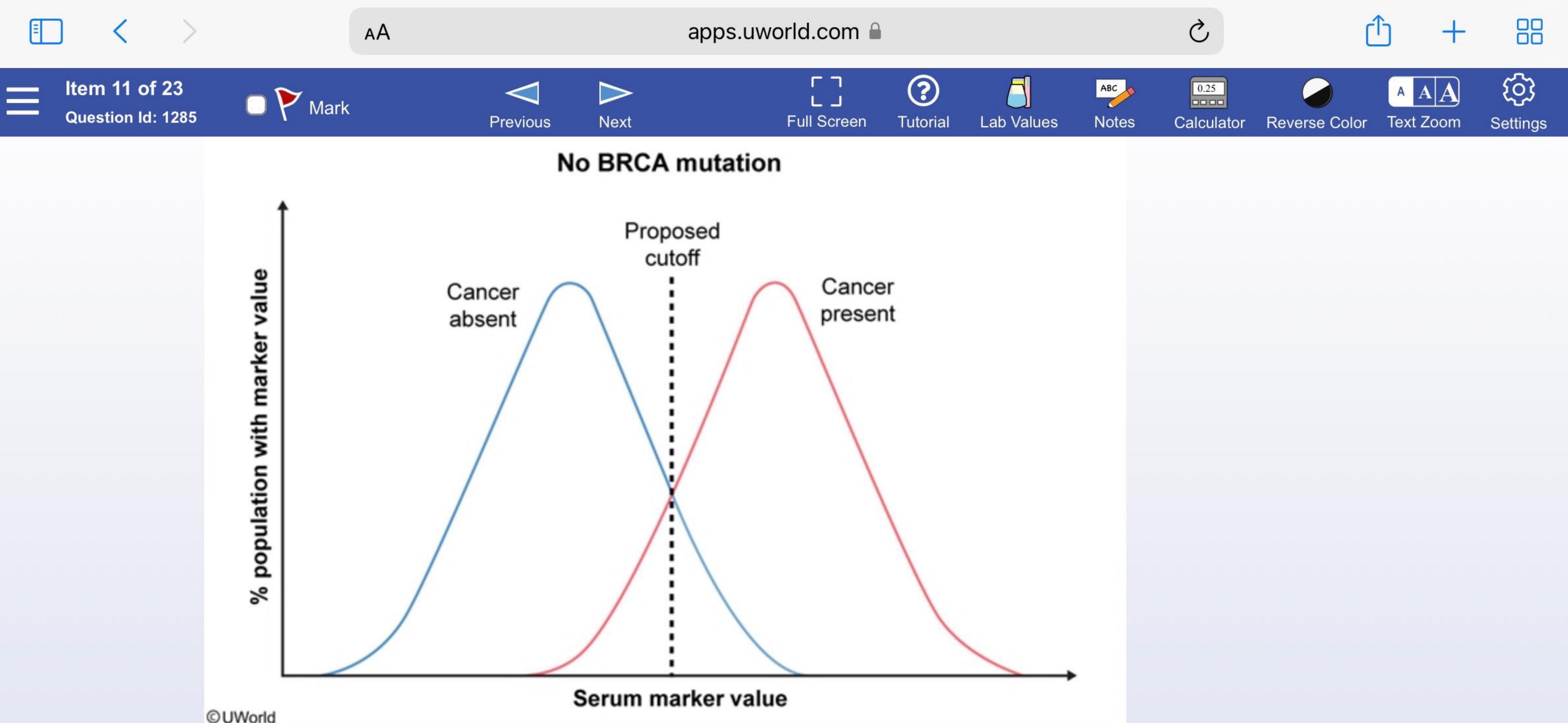
apps.uworld.com

Item 11 of 23 AA Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

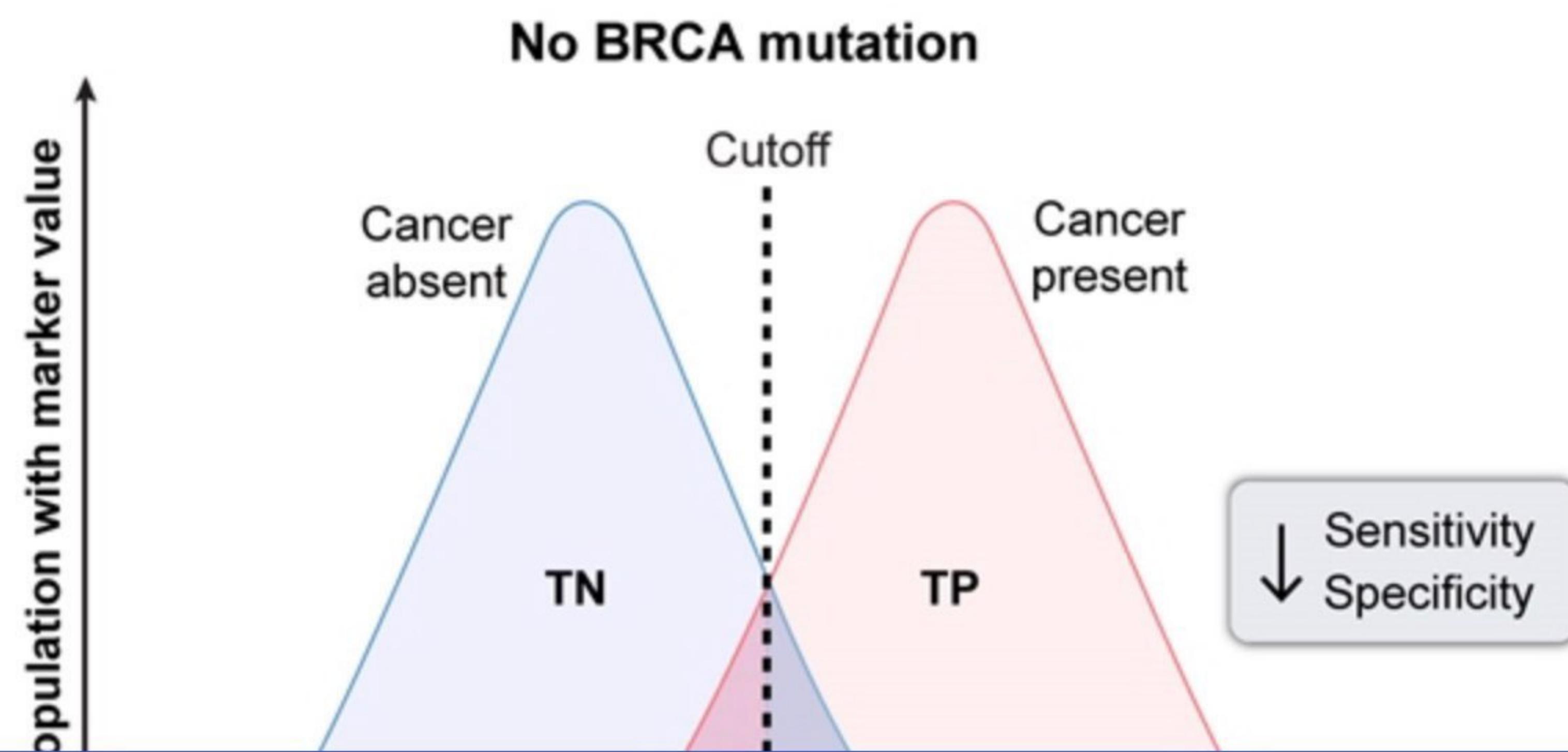
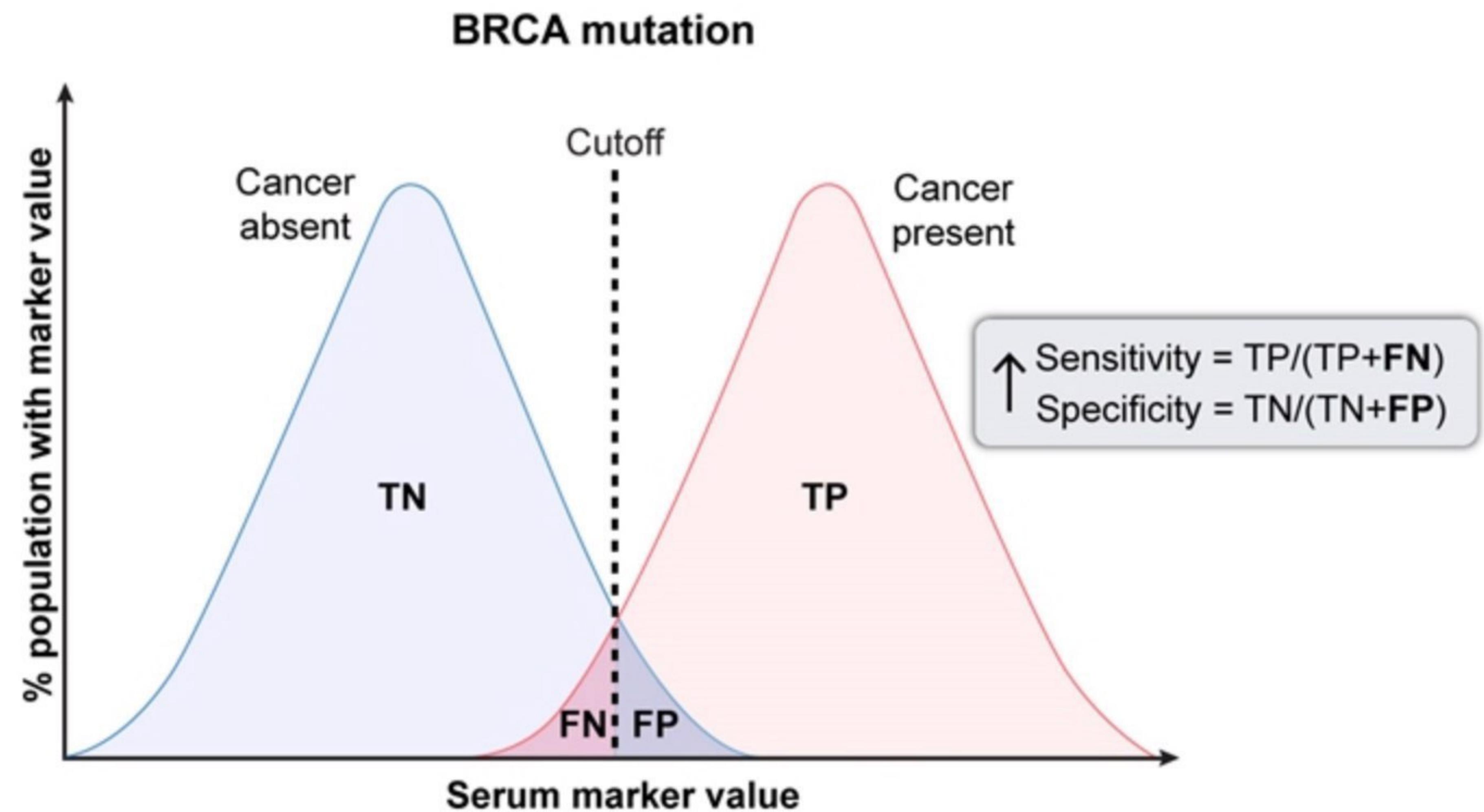
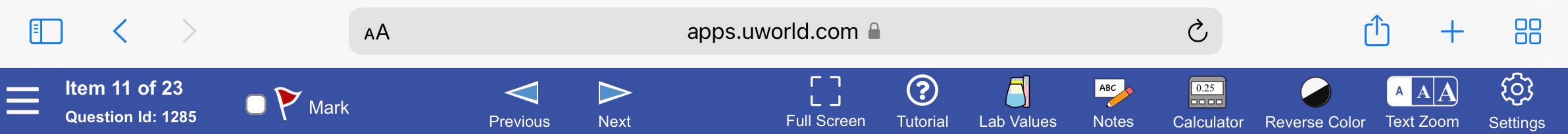
Mark Previous Next

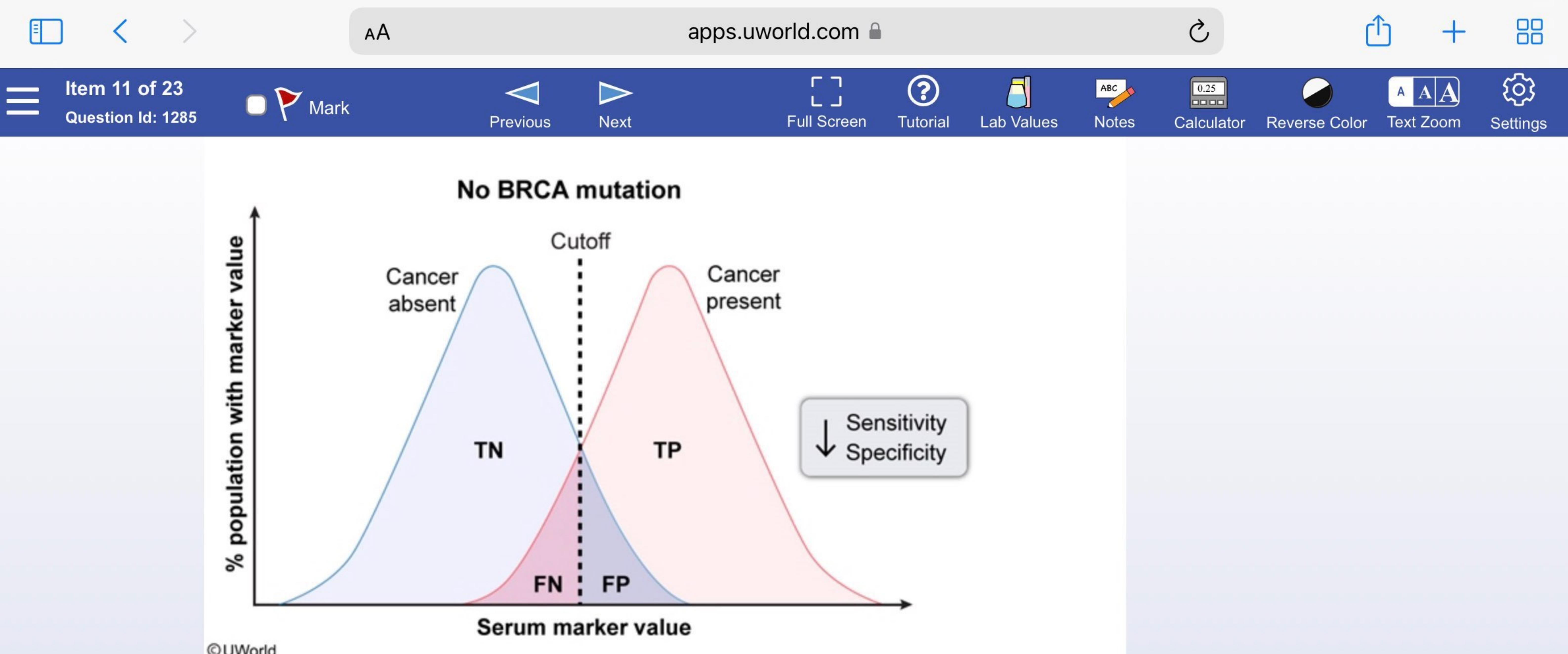
A new tumor marker is being investigated for its usefulness in diagnosing ovarian cancer. A sample of 400 women is stratified into 2 groups based on the presence or absence of *BRCA* mutations. Serum levels of the new marker are obtained in the 2 groups; in addition, both groups of women undergo conventional screening to determine their disease status. The curves on the top represent the distribution of the new serum marker in women with *BRCA* mutations, and the curves on the bottom represent the distribution of the new serum marker in women without the *BRCA* mutations.





- A. Higher sensitivity and higher specificity (64%)
- B. Higher sensitivity and lower specificity (8%)
- C. Higher sensitivity and same specificity (1%)
- D. Lower sensitivity and higher specificity (7%)





The cutoff value of a quantitative diagnostic test determines whether a given result is interpreted as positive or negative. When there is **overlap between** the serum values of the **healthy and diseased populations**, a cutoff value that correctly categorizes all individuals in both populations cannot be chosen. This **limits the sensitivity and specificity** of the test due to the presence of false positive (FP) and/or false negative (FN) individuals.

Sensitivity represents the ability of a test to correctly identify those with a given disease. It is calculated as the number of patients correctly testing positive (TP) divided by the total number of patients with the disease ($TP / [TP + FN]$). Specificity represents the ability of a test to correctly identify those without a given disease. It is calculated as the number of patients correctly testing negative (TN) divided by the total number of patients without the disease ($TN / [TN + FP]$).

apps.uworld.com

Item 11 of 23 Question Id: 1285 ©UWorld

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

The cutoff value of a quantitative diagnostic test determines whether a given result is interpreted as positive or negative. When there is **overlap between** the serum values of the **healthy and diseased populations**, a cutoff value that correctly categorizes all individuals in both populations cannot be chosen. This **limits the sensitivity and specificity** of the test due to the presence of false positive (FP) and/or false negative (FN) individuals.

Sensitivity represents the ability of a test to correctly identify those with a given disease. It is calculated as the number of patients correctly testing positive (TP) divided by the total number of patients with the disease ($TP / [TP + FN]$). Specificity represents the ability of a test to correctly identify those without a given disease. It is calculated as the number of patients correctly testing negative (TN) divided by the total number of patients without the disease ($TN / [TN + FP]$).

In this case, values of the serum marker in women with *BRCA* mutations show **decreased overlap** between the healthy and diseased curves of women with *BRCA* mutations (curves on the top) compared with women without *BRCA* mutations (curves on the bottom). The corresponding decrease in the number of FPs and FNs means the new serum marker has **higher sensitivity and specificity** (ie, better performance) in women with *BRCA* mutations.

Educational objective:

The degree of overlap between the healthy and the diseased population curves limits the maximum combined sensitivity and specificity of a quantitative diagnostic test. The degree to which sensitivity or specificity is affected depends on the chosen cutoff value.

Biostatistics
Subject

Biostatistics & Epidemiology
System

Sensitivity and specificity
Topic

Copyright © UWorld. All rights reserved.

apps.uworld.com

Item 12 of 23 Question Id: 19262

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A research group conducted a placebo-controlled clinical trial to assess whether a new drug to treat acute migraine with or without aura in adults is more effective than standard therapy. A total of 3,500 patients with acute migraine were enrolled in the study and randomly assigned to either the new drug or standard treatment. During the data analysis phase, the researchers decide to set alpha at 0.01 rather than 0.05. Which of the following is the most likely result of this change?

- A. Any significant findings will be reported with greater confidence (64%)
- B. There will be a higher probability of a type I error (10%)
- C. There will be a higher probability of finding statistically significant results (10%)
- D. There will be a lower probability of a type II error (5%)
- E. The study will have more statistical power (8%)

Omitted
Correct answer
A

64%
Answered correctly

03 secs
Time Spent

2023
Version

Explanation

Type I (α) and type II (β) errors

True status	
There is a true	There is NO true
reject H ₀	commit Type I error
fail to reject H ₀	commit Type II error

apps.uworld.com

Item 12 of 23 Question Id: 19262

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

Type I (α) and type II (β) errors

		True status	
		There is a true difference (ie, H_0 is false)	There is NO true difference (ie, H_0 is true)
Study result	Difference calculated as statistically significant (ie, reject H_0)	Correctly conclude there is a difference	Type I (α) error (Falsely conclude there is a difference)
	Difference calculated as NOT statistically significant (ie, fail to reject H_0)	Type II (β) error (Falsely conclude there is NO difference)	Correctly conclude there is NO difference

H_0 = null hypothesis of no difference.

©UWorld

Alpha (α) is the significance level used to establish the **statistical significance** of study results. It is usually predetermined by investigators as the threshold to reject a null hypothesis (H_0). The results of a study are considered statistically significant when the p -value (ie, probability of obtaining results as extreme as the observed results when H_0 is true) is less than the significance level (ie, p -value $< \alpha$). Alpha is also the probability of making a **type I error**, which is rejecting a true H_0 (ie, a false positive: finding a statistically significant difference when one does not truly exist). The **complement** of alpha/significance level is the **confidence level**.

The screenshot shows a mobile application interface for a study question. At the top, there is a header with the URL "apps.uworld.com". Below the header, a blue navigation bar contains various icons and text elements. From left to right, the icons are: a square with a vertical line, a left arrow, a right arrow, a double arrow, a magnifying glass, a double 'AA' symbol, a lock icon, a 'Mark' icon with a red flag, a 'Previous' arrow, a 'Next' arrow, a 'Full Screen' icon, a 'Tutorial' icon with a question mark, a 'Lab Values' icon with a test tube, a 'Notes' icon with a pencil, a 'Calculator' icon with '0.25', a 'Reverse Color' icon with a black circle, a 'Text Zoom' icon with three 'A's of increasing size, and a 'Settings' gear icon.

Alpha (α) is the significance level used to establish the **statistical significance** of study results. It is usually predetermined by investigators as the threshold to reject a null hypothesis (H_0). The results of a study are considered statistically significant when the *p*-value (ie, probability of obtaining results as extreme as the observed results when H_0 is true) is less than the significance level (ie, ***p*-value < α**). Alpha is also the probability of making a **type I error**, which is rejecting a true H_0 (ie, a false positive: finding a statistically significant difference when one does not truly exist). The **complement** of alpha/significance level is the **confidence level** ($1 - \alpha$), which represents the probability of not rejecting a true H_0 (ie, a true negative, or not finding a statistically significant difference when one does not truly exist).

The lower the significance level, the lower the probability of a type I error. However, this means that the threshold to attain statistical significance is more stringent. Therefore, **reducing the significance level** (eg, from $\alpha = 5\%$ to $\alpha = 1\%$) **decreases** the probability of a **type I error** along with the probability of **finding statistical significance** (eg, *p*-value < 0.05 to *p*-value < 0.01) (**Choices B and C**). A decreased significance level also means an **increased confidence level** ($1 - \alpha$); therefore, any statistically significant findings can be reported with a greater level of confidence.

(Choices D and E) A type II error occurs when a false H_0 is not rejected (ie, a false negative: not finding a statistically significant difference when one truly exists). The probability of a type II error (known as beta [β]) is the complement of power ($1 - \beta$). Statistical power is the probability of rejecting a false H_0 . That is, the power to find a statistically significant difference when one truly exists (ie, true positive). Reducing the significance level from 0.05 to 0.01 makes the threshold to attain statistical significance more stringent. Consequently, the study will have less statistical power to detect a statistically significant difference when one truly exists and will have a higher probability of a type II error (the complement of power).

Educational objective:

The screenshot shows a mobile application interface for a statistics question. At the top, there's a header with the URL "apps.uworld.com". Below the header is a toolbar with various icons: a square icon, a left arrow, a right arrow, a double arrow, a magnifying glass, a question mark, a calculator, a reverse color button, a text zoom button, and a settings gear icon. On the far left, there's a vertical menu icon and the text "Item 12 of 23" and "Question Id: 19262". To the right of the toolbar are buttons for "Mark", "Previous", "Next", "Full Screen", "Tutorial", "Lab Values", "Notes", "Calculator", "Reverse Color", "Text Zoom", and "Settings".

difference when one does not truly exist). The complement of alpha/significance level is the confidence level ($1 - \alpha$), which represents the probability of not rejecting a true H_0 (ie, a true negative, or not finding a statistically significant difference when one does not truly exist).

The lower the significance level, the lower the probability of a type I error. However, this means that the threshold to attain statistical significance is more stringent. Therefore, **reducing the significance level** (eg, from $\alpha = 5\%$ to $\alpha = 1\%$) **decreases** the probability of a **type I error** along with the probability of **finding statistical significance** (eg, p -value < 0.05 to p -value < 0.01) (**Choices B and C**). A decreased significance level also means an **increased confidence level** ($1 - \alpha$); therefore, any statistically significant findings can be reported with a greater level of confidence.

(Choices D and E) A type II error occurs when a false H_0 is not rejected (ie, a false negative: not finding a statistically significant difference when one truly exists). The probability of a type II error (known as beta [β]) is the complement of power ($1 - \beta$). Statistical power is the probability of rejecting a false H_0 . That is, the power to find a statistically significant difference when one truly exists (ie, true positive). Reducing the significance level from 0.05 to 0.01 makes the threshold to attain statistical significance more stringent. Consequently, the study will have less statistical power to detect a statistically significant difference when one truly exists and will have a higher probability of a type II error (the complement of power).

Educational objective:

Reducing the significance level alpha (α) in a study allows researchers to report any significant findings with greater confidence.

Biostatistics
Subject

Biostatistics & Epidemiology
System

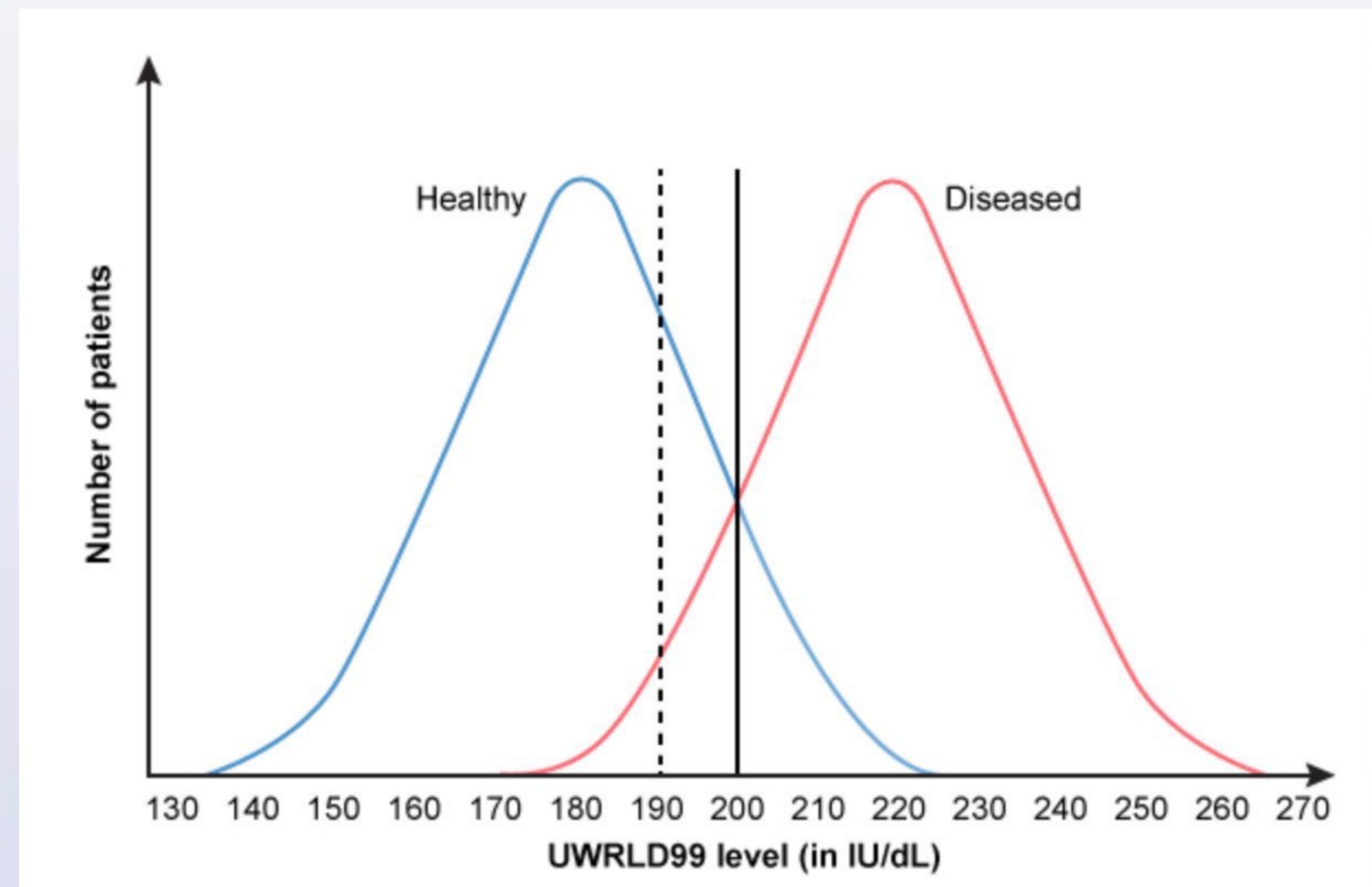
Hypothesis testing
Topic

apps.uworld.com

Item 13 of 23 Question Id: 1191

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A novel serum biomarker, UWRLD99, is being evaluated for early detection of pancreatic cancer. The concentration of UWRLD99 (measured in international units [IU] per dL) has been found to be elevated in patients with pancreatic adenocarcinoma. Test results in 200 volunteers ("healthy") and 190 patients with biopsy-proven pancreatic cancer ("diseased") are given in the figure below. The investigators determined the sensitivity and specificity of UWRLD99 for pancreatic cancer using a cutoff of 200 IU/dL (indicated by the solid vertical line).



If they had instead used a lower cutoff as indicated by the dashed vertical line, which of the following would most likely be seen?

apps.uworld.com

Item 13 of 23 Question Id: 1191

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

If they had instead used a lower cutoff as indicated by the dashed vertical line, which of the following would most likely be seen?

- A. Higher number of false negatives (8%)
- B. Higher positive predictive value (9%)
- C. Higher sensitivity (66%)
- D. Lower number of false positives (5%)
- E. Lower number of true positives (9%)

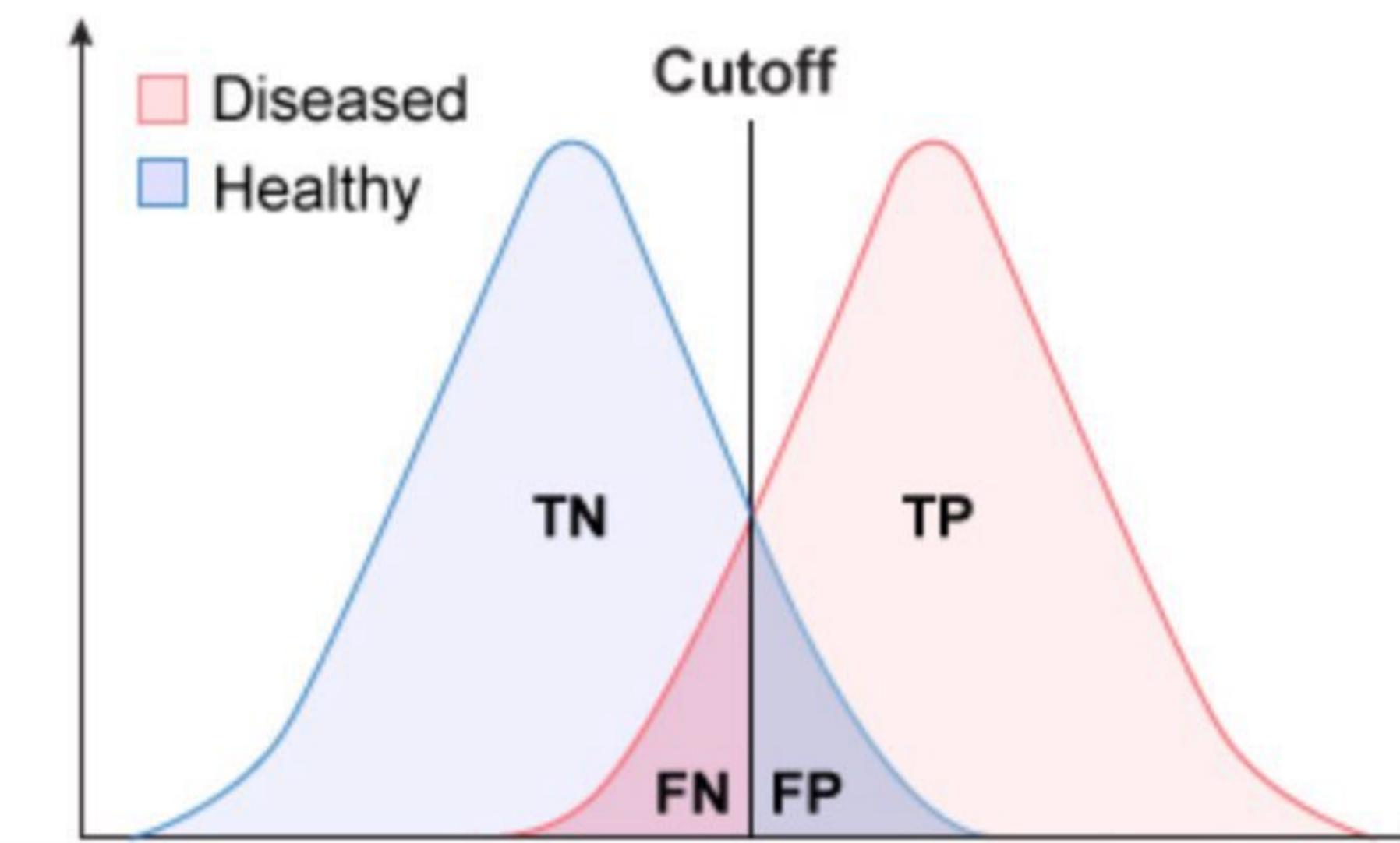
Omitted
Correct answer
C

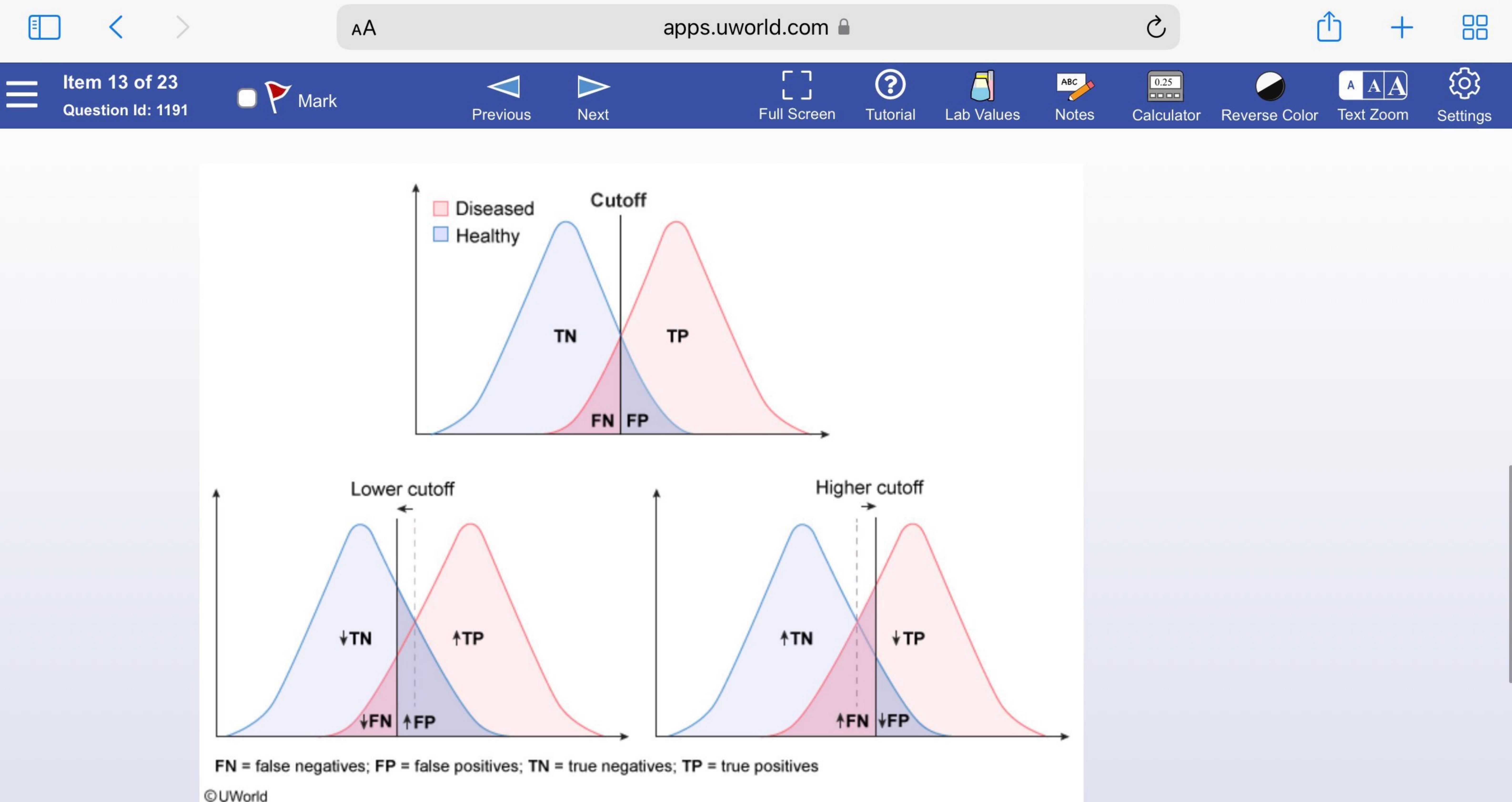
66%
Answered correctly

08 secs
Time Spent

2023
Version

Explanation





Important parameters of diagnostic tests include the following:

- True positives (TP) represent diseased individuals with positive test results.
- True negatives (TN) represent healthy individuals with negative test results.
- False positives (FP) represent healthy individuals with positive test results.
- False negatives (FN) represent diseased individuals with negative test results.

apps.uworld.com

Item 13 of 23 Question Id: 1191

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

Important parameters of diagnostic tests include the following:

- True positives (TP) represent diseased individuals with positive test results.
- True negatives (TN) represent healthy individuals with negative test results.
- False positives (FP) represent healthy individuals with positive test results.
- False negatives (FN) represent diseased individuals with negative test results.
- **Sensitivity** ($TP / [TP + FN]$) represents a test's ability to correctly identify diseased individuals from among all individuals. A test with high sensitivity has a low FN rate (important for screening purposes). With a highly sensitive test, most diseased patients will have a positive test result (and a **negative** test result would help rule **out** the disease [**SnNOut**]).
- **Specificity** ($TN / [TN + FP]$) represents the ability of a test to exclude those without the disease. A very specific test has a low FP rate (important for confirmatory tests). With a highly specific test, most healthy patients will have a negative test result (and a **positive** test result would help rule **in** the disease [**SpPIn**])).

The **cutoff value** of a quantitative diagnostic test determines whether a given result is interpreted as **positive or negative**. Depending on the disease or condition being tested for, sensitivity or specificity may be preferred and the cutoff value adjusted accordingly. A cutoff value just outside the overlapping region of the curves can maximize the sensitivity or specificity at 100% by correctly classifying all diseased or healthy individuals, respectively.

In the above example, changing the cutoff point to a **lower value** (shift to the left) would cause more patients with the disease to test positive ($\uparrow TP, \downarrow FN$), **increasing the sensitivity** of the test. However, as a consequence, more patients without the disease would also test positive ($\downarrow TN, \uparrow FP$), resulting in **decreased specificity**.

(Choices A, D, and E) Raising the cutoff value (shift to the right) would cause fewer individuals with the disease to have a positive test result ($\downarrow TP, \uparrow FN$), so sensitivity would decrease. Fewer individuals without the disease

apps.uworld.com

Item 13 of 23 Question Id: 1191

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

The **cutoff value** of a quantitative diagnostic test determines whether a given result is interpreted as **positive or negative**. Depending on the disease or condition being tested for, sensitivity or specificity may be preferred and the cutoff value adjusted accordingly. A cutoff value just outside the overlapping region of the curves can maximize the sensitivity or specificity at 100% by correctly classifying all diseased or healthy individuals, respectively.

In the above example, changing the cutoff point to a **lower value** (shift to the left) would cause more patients with the disease to test positive ($\uparrow TP, \downarrow FN$), **increasing the sensitivity** of the test. However, as a consequence, more patients without the disease would also test positive ($\downarrow TN, \uparrow FP$), resulting in **decreased specificity**.

(Choices A, D, and E) Raising the cutoff value (shift to the right) would cause fewer individuals with the disease to have a positive test result ($\downarrow TP, \uparrow FN$), so sensitivity would decrease. Fewer individuals without the disease would also test positive ($\uparrow TN, \downarrow FP$), so specificity would increase.

(Choice B) Positive predictive value ($PPV = TP / [TP + FP]$) represents the probability that a patient with a positive test result actually has the disease. In this case, the lower cutoff value would increase both TP and FP; how this affects the PPV depends on **disease prevalence**, which is not provided. In a population with very low disease prevalence, the number of FP would be expected to increase proportionately more than the number of TP, which would lower the PPV.

Educational objective:

The cutoff value of a quantitative diagnostic test determines whether a given result is interpreted as positive or negative. Lowering the cutoff point typically causes more patients with the disease to test positive, decreasing the number of false negatives and increasing test sensitivity. Consequently, more patients without the disease will also test positive, resulting in an increased number of false positives and decreased specificity.

apps.uworld.com

Item 14 of 23 Question Id: 19314 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A study was conducted to assess the age at menarche among young female gymnasts. Based on commitment to the sport, gymnasts were divided into two groups: competitive gymnasts and recreational gymnasts.

Age at menarche

	Sample size (n)	Mean, y	Standard deviation, y
Competitive	16	13.4	1.3
Recreational	22	12.3	0.8

Assuming that age at menarche is normally distributed, which of the following is closest to the probability that a randomly chosen competitive gymnast will have onset of menarche at age ≥ 16 ?

- A. 0.997 (3%)
- B. 0.950 (7%)
- C. 0.680 (3%)
- D. 0.160 (4%)
- E. 0.025 (72%)
- F. 0.0015 (9%)

Omitted

Correct answer

E



72%

Answered correctly



03 secs

Time Spent



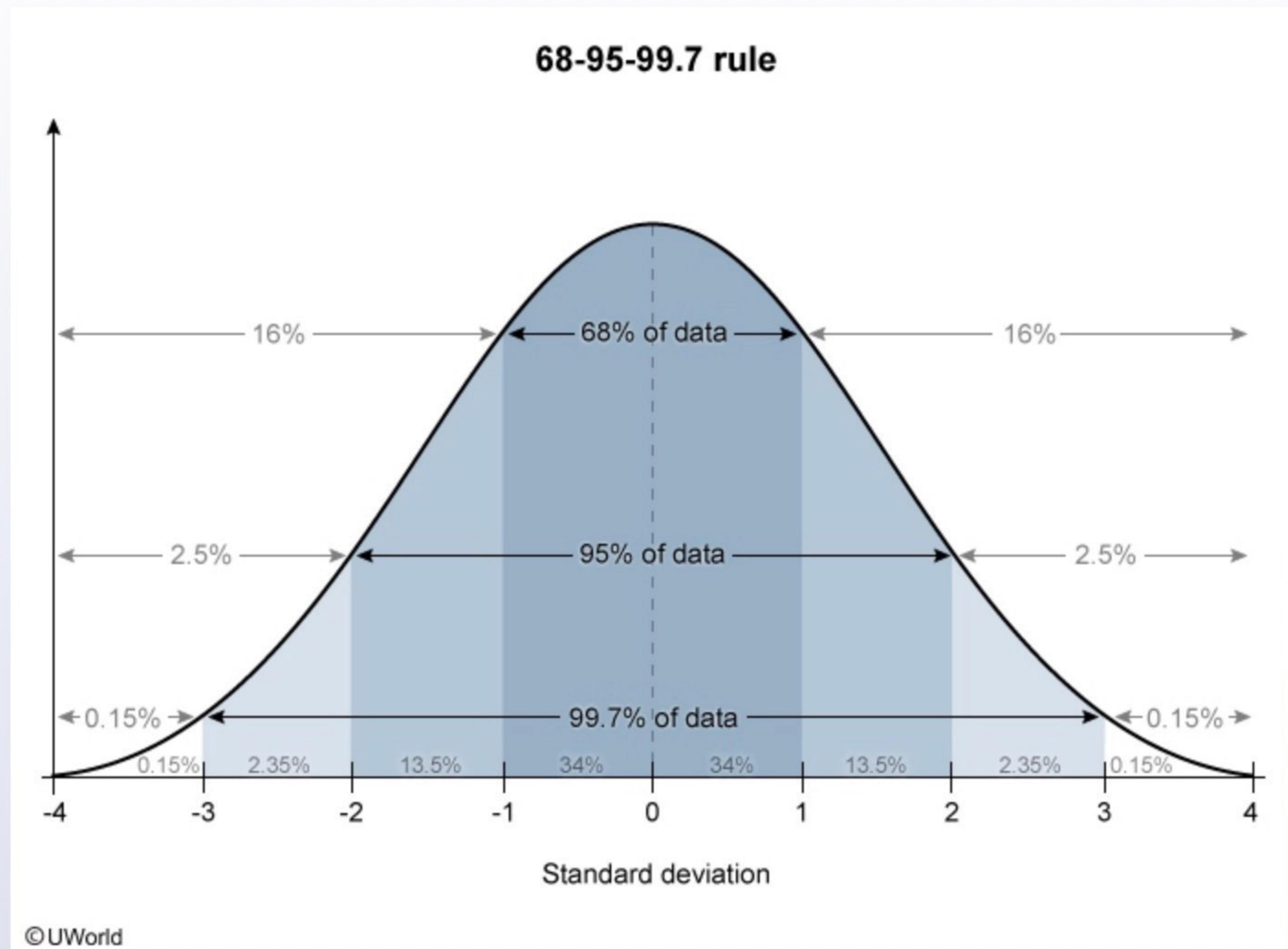
2023

Version

apps.uworld.com

Item 14 of 23 Question Id: 19314

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings



A normal (Gaussian) distribution is a symmetrical, bell-shaped distribution with a fixed percentage of observations lying within a certain distance of the mean. This distance is called the standard deviation (SD) and represents the degree of dispersion from the mean. The **68-95-99.7** rule for normal distributions states that 68% of all observations lie within 1 SD of the mean, 95% of all observations lie within 2 SDs of the mean, and 99.7% of all observations lie within 3 SDs of the mean.

apps.uworld.com

Item 14 of 23 Question Id: 19314

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A normal (Gaussian) distribution is a symmetrical, bell-shaped distribution with a fixed percentage of observations lying within a certain distance of the mean. This distance is called the standard deviation (SD) and represents the degree of dispersion from the mean. The **68-95-99.7** rule for normal distributions states that 68% of all observations lie within 1 SD of the mean, 95% of all observations lie within 2 SDs of the mean, and 99.7% of all observations lie within 3 SDs of the mean.

For competitive gymnasts in this sample, the **mean age at menarche is 13.4 years**, with a **SD of 1.3 years**.

Based on the 68-95-99.7 rule:

- 68% of observations lie within 1 SD: $13.4 \pm 1.3 = 12.1-14.7$.
- 95% of observations lie within **2 SDs**: $13.4 \pm 2.6 = 10.8-16.0$.
- 99.7% of observations lie within 3 SDs: $13.4 \pm 3.9 = 9.5-17.3$.

An onset of menarche at **age ≥ 16 years is 2 SDs** from the mean; therefore, **2.5%** of the observations must lie **above 16 years** (with 2.5% of observations below 10.8 years). The probability that a random competitive gymnast will have an onset of menarche at age ≥ 16 years is 0.025.

(Choices A, B, and C) The 68-95-99.7 rule states that 99.7% of observations lie within 3 SDs; therefore, 0.997 is the probability that a random competitive gymnast will have an onset of menarche between age 9.5 and 17.3 years. Similarly, 0.95 is the probability that a random competitive gymnast will have an onset of menarche between age 10.8 and 16.0 years (ie, within ± 2 SDs from the mean), and 0.68 is the probability that a random competitive gymnast will have an onset of menarche between age 12.1 and 14.7 years (ie, within ± 1 SD from the mean).

(Choices D and F) Based on the 68-95-99.7 rule, 32% (ie, $100\% - 68\%$) of observations lie outside 1 SD from the mean, with half (ie, $32/2 = 16\%$) above and half (16%) below 1 SD from the mean. Therefore, 0.160 is the

apps.uworld.com

Item 14 of 23 Question Id: 19314

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

- 95% of observations lie within 2 SDs: $13.4 \pm 2.6 = 10.8-16.0$.
- 99.7% of observations lie within 3 SDs: $13.4 \pm 3.9 = 9.5-17.3$.

An onset of menarche at age **≥16 years** is 2 SDs from the mean; therefore, **2.5%** of the observations must lie **above 16 years** (with 2.5% of observations below 10.8 years). The probability that a random competitive gymnast will have an onset of menarche at age ≥ 16 years is 0.025.

(Choices A, B, and C) The 68-95-99.7 rule states that 99.7% of observations lie within 3 SDs; therefore, 0.997 is the probability that a random competitive gymnast will have an onset of menarche between age 9.5 and 17.3 years. Similarly, 0.95 is the probability that a random competitive gymnast will have an onset of menarche between age 10.8 and 16.0 years (ie, within ± 2 SDs from the mean), and 0.68 is the probability that a random competitive gymnast will have an onset of menarche between age 12.1 and 14.7 years (ie, within ± 1 SD from the mean).

(Choices D and F) Based on the 68-95-99.7 rule, 32% (ie, $100\% - 68\%$) of observations lie outside 1 SD from the mean, with half (ie, $32/2 = 16\%$) above and half (16%) below 1 SD from the mean. Therefore, 0.160 is the probability that a random competitive gymnast will have an onset of menarche at either age ≤ 12.1 years (ie, ≥ 1 SD below the mean) or at age ≥ 14.7 years (ie, ≥ 1 SD above the mean). Similarly, 0.0015 is the probability that a random competitive gymnast will have an onset of menarche at either age ≤ 9.5 years (ie, ≥ 3 SDs below the mean) or at age ≥ 17.3 years (ie, ≥ 3 SDs above the mean).

Educational objective:

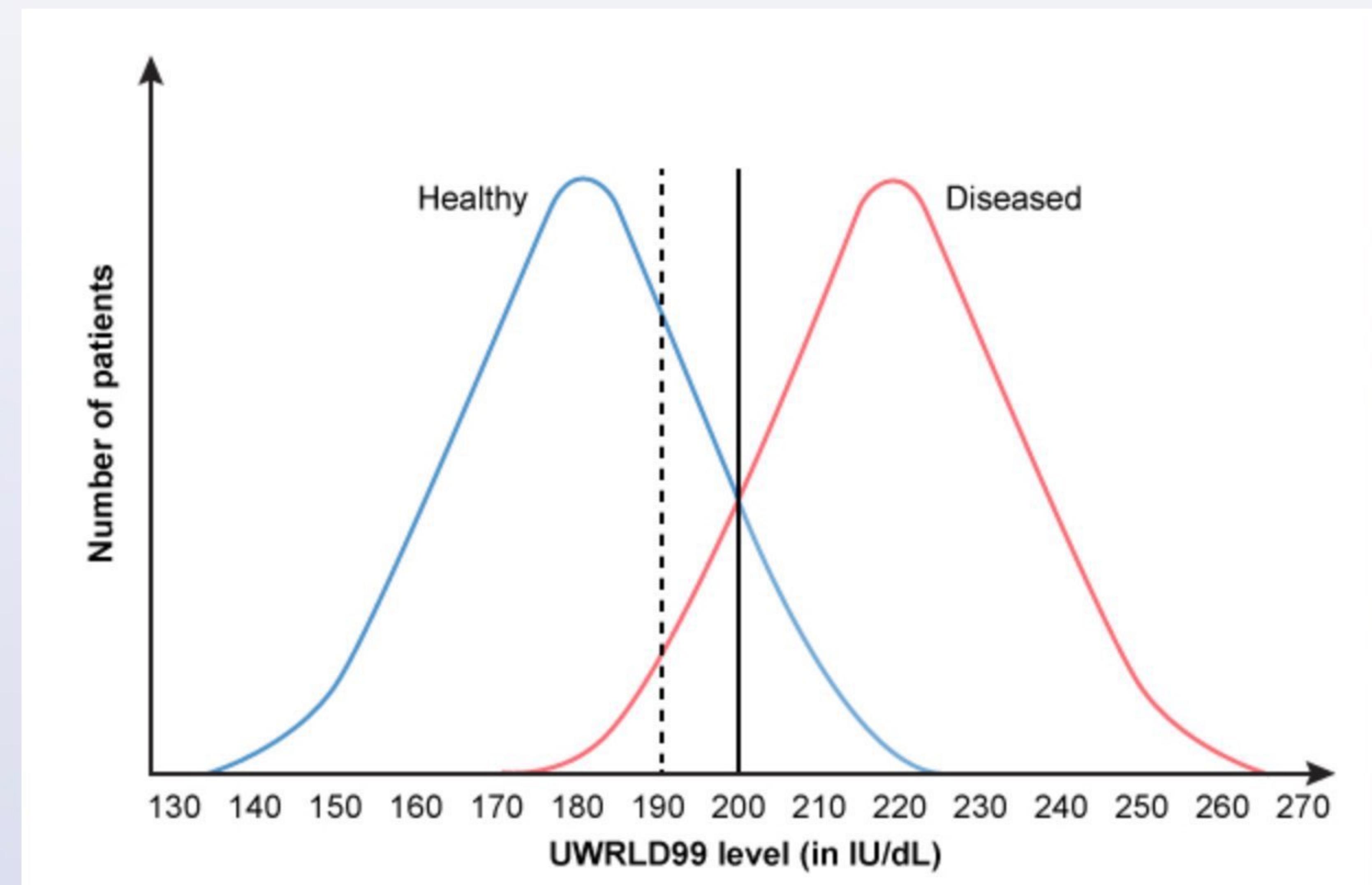
In a normal (bell-shaped) distribution, 68% of all values are within 1 standard deviation (SD) of the mean; 95% are within 2 SDs of the mean; and 99.7% are within 3 SDs of the mean.

apps.uworld.com

Item 13 of 23 Question Id: 1191

Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A novel serum biomarker, UWRLD99, is being evaluated for early detection of pancreatic cancer. The concentration of UWRLD99 (measured in international units [IU] per dL) has been found to be elevated in patients with pancreatic adenocarcinoma. Test results in 200 volunteers ("healthy") and 190 patients with biopsy-proven pancreatic cancer ("diseased") are given in the figure below. The investigators determined the sensitivity and specificity of UWRLD99 for pancreatic cancer using a cutoff of 200 IU/dL (indicated by the solid vertical line).



If they had instead used a lower cutoff as indicated by the dashed vertical line, which of the following would most likely be seen?