

apps.uworld.com

A **case-control study** is the most appropriate study design for evaluating the public health officials' claim. This is because the disease (acute myelogenous leukemia [AML]) is a rare condition occurring at a higher rate in this population, and a retrospective exposure (chemical waste exposure) needs to be evaluated. In case-control studies, 2 groups of subjects are created: **cases** (subjects **with** the disease of interest) and **controls** (subjects **without** the disease of interest). After the case and control groups are selected, exposure frequency to a specific variable (eg, chemical waste) within both groups is ascertained. If there is a statistically significant difference in exposure frequency between the 2 groups, it is likely that the variable in question is associated with disease development.

In this example, AML is the disease of interest; therefore, children with AML should be used as cases and children **without AML** should be used as **controls**. Cases and controls should be selected **regardless** of exposure status to the chemical waste (**Choices A and B**). Selecting subjects based on exposure status is inappropriate because comparing the frequency of exposure between the case and control groups is what determines whether the exposure is more prevalent among cases as compared to controls.

Ideally, exposure frequency among controls should be representative of that among the population of individuals "at risk" of becoming cases. In other words, for a given case-control study, controls are nondiseased individuals who could be considered cases if they had the disease. Often, controls and cases are matched based on independent variables (eg, age, sex) to decrease the effects of confounding.

(Choices D, E, and F) AML is the outcome of interest; therefore, children who have AML can be used only to form the cases group and cannot be used as controls. Cases should also be selected regardless of exposure status.

Educational objective:

Selection of control subjects in case-control studies is intended to provide an accurate estimation of exposure

difference in exposure frequency between the 2 groups, it is likely that the variable in question is associated with disease development.

In this example, AML is the disease of interest; therefore, children with AML should be used as cases and children **without AML** should be used as **controls**. Cases and controls should be selected **regardless** of exposure status to the chemical waste (**Choices A and B**). Selecting subjects based on exposure status is inappropriate because comparing the frequency of exposure between the case and control groups is what determines whether the exposure is more prevalent among cases as compared to controls.

Ideally, exposure frequency among controls should be representative of that among the population of individuals "at risk" of becoming cases. In other words, for a given case-control study, controls are nondiseased individuals who could be considered cases if they had the disease. Often, controls and cases are matched based on independent variables (eg, age, sex) to decrease the effects of confounding.

(Choices D, E, and F) AML is the outcome of interest; therefore, children who have AML can be used only to form the cases group and cannot be used as controls. Cases should also be selected regardless of exposure status.

Educational objective:

Selection of control subjects in case-control studies is intended to provide an accurate estimation of exposure frequency among the nondiseased general population. Cases and controls should be selected based on disease status, not exposure status.

Biostatistics
Subject

Biostatistics & Epidemiology
System

Study designs
Topic

Copyright © UWorld. All rights reserved.

apps.uworld.com

Item 23 of 40 Question Id: 1188 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A large prospective study evaluates the relationship between alcohol consumption and breast cancer. A total of 4000 middle-aged women are enrolled in the study through a random selection of residential addresses. Daily alcohol consumption and breast cancer incidence are assessed through the use of periodic questionnaires. On five-year follow-up, the investigators report the association between alcohol consumption and breast cancer as a relative risk of 1.32 (95% confidence interval = 0.90-1.85); 800 subjects were lost to follow-up by the end of the study, the majority of whom were moderate to heavy alcohol consumers. According to this information, which of the following biases is most likely to be present and may have affected the results?

- A. Lead-time bias (13%)
- B. Observer bias (6%)
- C. Random misclassification bias (11%)
- D. Recall bias (11%)
- E. Selection bias (56%)

Omitted
Correct answer
E

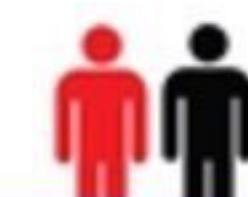
56%
Answered correctly

53 secs
Time Spent

2023
Version

Explanation

Attrition bias



apps.uworld.com

Item 23 of 40 Question Id: 1188 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

Attrition bias

The diagram illustrates Attrition bias through two groups: Unexposed and Exposed.

Unexposed Group: Starts with 7 black human icons. After an arrow, there are 5 black icons and 2 red icons. A bracket indicates "2 outcomes".

Exposed Group: Starts with 7 black human icons. A box encloses the first 4 icons. An arrow points to 3 red icons outside the box. A bracket indicates "3 outcomes". Another arrow points to 2 red icons inside the box. A bracket indicates "2 outcomes".

A legend on the right states: "Exposed & unexposed groups appear to have the same number of diseased cases, but the disease actually occurs more frequently in the exposed group".

©UWorld

Bias refers to a systematic error (due to nonrandom factors) that can result in an incorrect association between the exposure and outcome. In general, when bias is present, the results are consistently distorted in one

The screenshot shows a mobile application interface for a medical study. At the top, there's a header bar with icons for back, forward, and search. Below it is a toolbar with various functions: Item 23 of 40, Question Id: 1188, Mark (with a checkmark and a flag icon), Previous, Next, Full Screen, Tutorial, Lab Values, Notes, Calculator, Reverse Color, Text Zoom, and Settings. The main content area contains a detailed text about bias in studies, followed by three choice options (A, B, C) describing different types of bias.

Bias refers to a systematic error (due to nonrandom factors) that can result in an incorrect association between the exposure and outcome. In general, when bias is present, the **results are consistently distorted** in one direction. In prospective studies, if **loss to follow-up occurs disproportionately** between the exposed and unexposed groups, **attrition bias** can result if the lost subjects differ in their risk of developing the outcome compared to the remaining subjects. Attrition bias is a **form of selection bias** (a term that most often refers to systematic differences between groups in terms of treatment response or prognosis). Attrition bias does not occur when the losses happen randomly between the exposed and unexposed groups, as this simply leads to a smaller study population.

In this example, a substantial number of subjects ($800 / 4000 = 20\%$) were lost to follow-up, primarily those with higher levels of alcohol consumption. If these lost subjects had a higher incidence of breast cancer than those remaining in the study, then a selective loss of high-risk subjects would have occurred in the exposed group. As a result, the measure of the association between alcohol use and breast cancer would be underestimated, which might explain why no significant association was seen in this study (as the confidence interval of 0.90 to 1.85 crosses 1, which is the null value for relative risk).

(Choice A) Lead-time bias occurs when a screening test diagnoses a disease earlier than it would have appeared by natural history alone, so that the time from diagnosis until death appears prolonged even though there might actually be no improvement in survival.

(Choice B) Observer bias can occur in clinical trials when study participants or investigators are aware of individual treatment assignments. It can be prevented by performing a double-blind study in which neither participants nor investigators are aware of treatment assignments.

(Choice C) Misclassification bias occurs when either the exposure or the outcome is not identified correctly. Random (or nondifferential) misclassification affects all groups to the same extent. For example, if a pediatric

apps.uworld.com

a result, the measure of the association between alcohol use and breast cancer would be underestimated, which might explain why no significant association was seen in this study (as the confidence interval of 0.90 to 1.85 crosses 1, which is the null value for relative risk).

(Choice A) Lead-time bias occurs when a screening test diagnoses a disease earlier than it would have appeared by natural history alone, so that the time from diagnosis until death appears prolonged even though there might actually be no improvement in survival.

(Choice B) Observer bias can occur in clinical trials when study participants or investigators are aware of individual treatment assignments. It can be prevented by performing a double-blind study in which neither participants nor investigators are aware of treatment assignments.

(Choice C) Misclassification bias occurs when either the exposure or the outcome is not identified correctly. Random (or nondifferential) misclassification affects all groups to the same extent. For example, if a pediatric size sphygmomanometer cuff is used on all participants (treatment and control groups) as part of a study in adult patients, the blood pressure readings will be incorrect (due to the incorrect cuff size). However, the resulting misclassification will likely affect both groups to the same extent (because the same cuff was used).

(Choice D) Recall bias results from the inaccurate recall of past exposure status. It is a potential problem for retrospective studies such as case-control studies, particularly when questionnaires are used to inquire about distant past exposure. However, in this prospective study, exposure status is determined through periodic questionnaires assessing daily alcohol consumption (without inquiring about distant past exposure). In such a scenario, recall bias is less likely to affect results.

Educational objective:

In prospective studies, disproportionate loss to follow-up between the exposed and unexposed groups creates the potential for attrition bias, which is a form of selection bias. As a result, investigators generally try to achieve high patient follow-up rates in prospective studies.

apps.uworld.com

Item 27 of 40 Question Id: 1186 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A large cohort study is conducted to assess the association between smoking and squamous cell carcinoma of the esophagus among middle-aged Chinese men. During 10 years of follow-up, smokers have 5 times the risk of esophageal carcinoma compared to non-smokers (relative risk = 5.0, 95% confidence interval = 2.9-7.1). According to the study results, what percentage of squamous cell carcinoma of the esophagus in smokers can be attributed to smoking?

- A. 25% (12%)
- B. 50% (13%)
- C. 70% (10%)
- D. 80% (52%)
- E. 90% (11%)

Omitted
Correct answer
D

52%
Answered correctly

50 secs
Time Spent

2023
Version

Explanation

The **attributable risk percent in the exposed (ARP_{exposed})** is an important measure of the impact of a risk factor. ARP_{exposed} represents the excess risk in an exposed population that can be explained by exposure to a particular risk factor. It is calculated using the following formula:

$$\text{ARP}_{\text{exposed}} = 100 \times [(\text{risk in exposed} - \text{risk in unexposed}) / \text{risk in exposed}]$$

Item 27 of 40
Question Id: 1186

Mark

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

The **attributable risk percent in the exposed (ARP_{exposed})** is an important measure of the impact of a risk factor. ARP_{exposed} represents the excess risk in an exposed population that can be explained by exposure to a particular risk factor. It is calculated using the following formula:

$$\text{ARP}_{\text{exposed}} = 100 \times [(\text{risk in exposed} - \text{risk in unexposed})/\text{risk in exposed}]$$

This basic definition can be used to derive an equivalent formula involving relative risk (RR):

$$\text{ARP}_{\text{exposed}} = 100 \times [(\text{RR} - 1)/\text{RR}], \text{ where RR} = \text{risk in exposed/risk in unexposed}$$

Applying the formula to this example:

$$\text{ARP}_{\text{exposed}} = 100 \times [(\text{RR} - 1)/\text{RR}] = 100 \times [(5 - 1)/5] = 100 \times (4/5) = 100 \times 0.8 = 80\%$$

Therefore, according to the study results, 80% of esophageal squamous cell carcinoma cases in smokers were attributable to smoking.

The ARP_{exposed} is related to the attributable risk (AR), which is simply the difference between risk in the exposed and risk in the unexposed.

Educational objective:

The attributable risk percent (ARP) in the exposed represents the excess risk in the exposed population that can be attributed to the risk factor. It can be easily derived from the relative risk (RR) using the formula: $\text{ARP}_{\text{exposed}} = 100 \times [(\text{RR} - 1)/\text{RR}]$.

apps.uworld.com

Item 29 of 40 Question Id: 1231

Mark

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

Zeracizumab is an experimental angiogenesis inhibitor targeting vascular endothelial growth factor that is being tested for the treatment of advanced, chemotherapy-naive, non-squamous non-small cell lung cancer (NSCLC). Part of the drug company's evaluation process is to analyze the 1-year survival after treatment to determine the clinical efficacy of the experimental treatment. The results are given in the table below.

Zeracizumab-containing regimen	Standard chemotherapy
Alive at 1 year	40
Dead at 1 year	60

Which of the following best represents the number needed to harm for the zeracizumab-containing regimen?

- A. 2 (14%)
- B. 25 (16%)
- C. 40 (59%)
- D. 72 (6%)
- E. 94 (3%)

Omitted
Correct answer
C

59%
Answered correctly

10 secs
Time Spent

2023
Version

Explanation

Test Id: 302978395

Scanned with CamScanner

apps.uworld.com

Item 29 of 40 Question Id: 1231 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

Unfortunately, not all interventions or new treatments end up helping patients. The **number needed to harm (NNH)** represents the number of people who must be treated before 1 additional adverse event occurs. It is calculated in a manner similar to the number needed to treat (NNT) but using the **absolute risk increase** (attributable risk) instead of the absolute risk reduction:

$$\text{NNH} = 1 / \text{Absolute risk increase}$$

To determine the absolute risk increase, first **calculate the adverse event rate** (eg, death) in the experimental and control groups. In this case, the adverse event of interest is death at 1 year. There were $40 + 60 = 100$ people treated with the experimental treatment (ie, zirconium-containing regimen); of those, 60 were dead at 1 year. Similarly, there were $51 + 69 = 120$ people treated with standard chemotherapy; of those, 69 were dead at 1 year. Therefore:

$$\text{Adverse event rate in experimental group} = 60 / 100 = 0.60 \text{ (ie, } 60\%)$$

$$\text{Adverse event rate in control group} = 69 / 120 = 0.575 \text{ (ie, } 57.5\%)$$

The absolute risk increase can then be calculated by **subtracting** the adverse event rate in the control group from the adverse event rate in the treatment group:

$$\text{Absolute risk increase} = 0.60 - 0.575 = 0.025 \text{ (ie, } 2.5\%)$$

The absolute risk increase attributable to the treatment is **2.5%** (60% risk of being dead at 1 year in the treatment group compared to 57.5% in the control group). The NNH then is simply the **inverse** of the absolute risk increase:

$$\text{NNH} = 1 / 0.025 = 40$$

This result indicates that, on average, **40 patients need to be treated** with a zirconium-containing regimen

apps.uworld.com

Item 29 of 40 Question Id: 1231 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

people treated with the experimental treatment (ie, zeraclizumab-containing regimen), of those, 60 were dead at

1 year. Similarly, there were $51 + 69 = 120$ people treated with standard chemotherapy; of those, 69 were dead at 1 year. Therefore:

$$\text{Adverse event rate in experimental group} = 60 / 100 = 0.60 \text{ (ie, } 60\%)$$

$$\text{Adverse event rate in control group} = 69 / 120 = 0.575 \text{ (ie, } 57.5\%)$$

The absolute risk increase can then be calculated by **subtracting** the adverse event rate in the control group from the adverse event rate in the treatment group:

$$\text{Absolute risk increase} = 0.60 - 0.575 = 0.025 \text{ (ie, } 2.5\%)$$

The absolute risk increase attributable to the treatment is **2.5%** (60% risk of being dead at 1 year in the treatment group compared to 57.5% in the control group). The NNH then is simply the **inverse** of the absolute risk increase:

$$\text{NNH} = 1 / 0.025 = 40$$

This result indicates that, on average, **40 patients need to be treated** with a zeraclizumab-containing regimen for 1 additional person to experience an adverse event (ie, death in 1 year).

Note that the data is NOT presented in the standard format of a contingency (2×2) table, so care should be exercised in selecting the appropriate values and applying the formulas (see [alternate solution](#)).

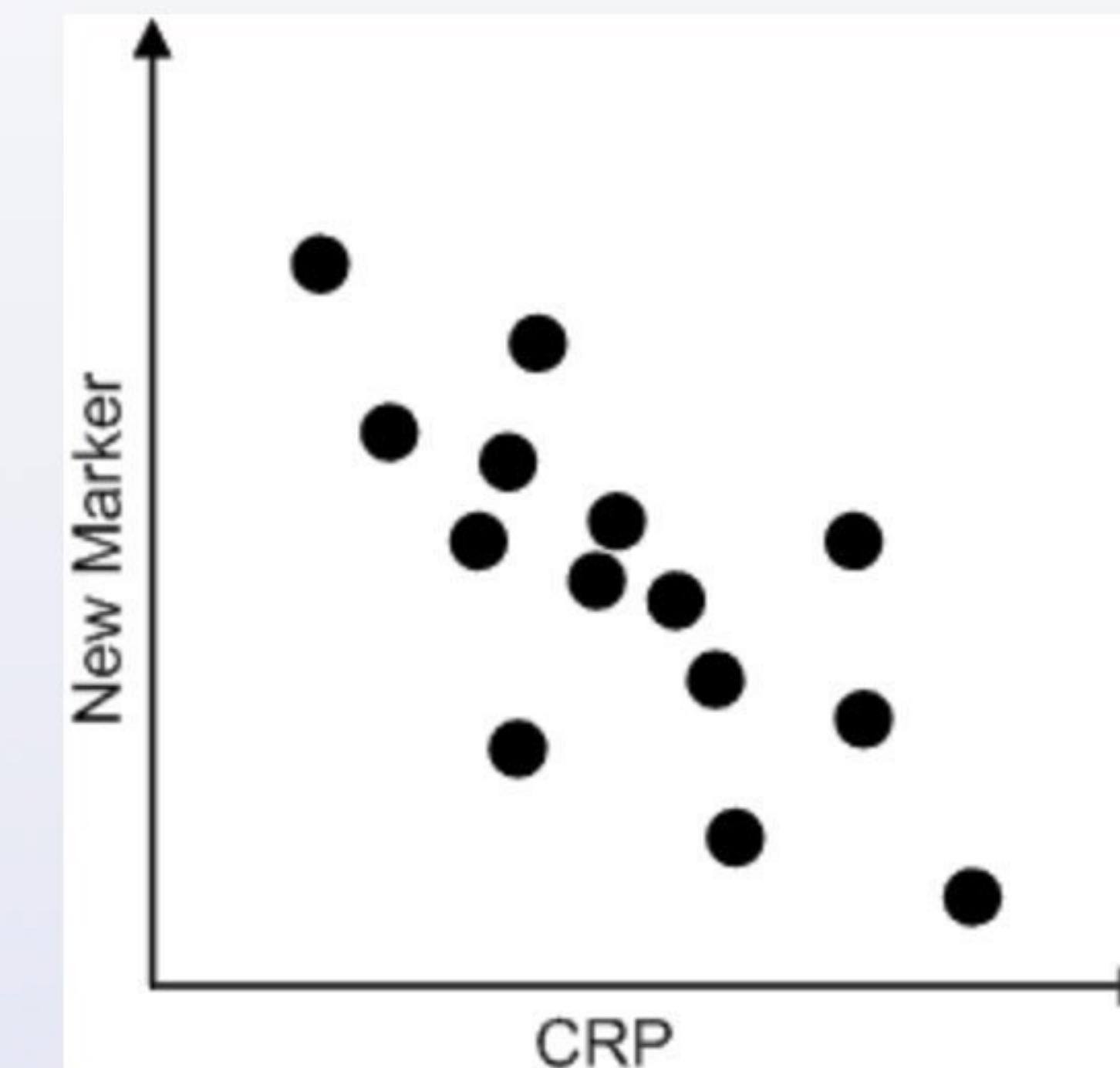
Educational objective:

The number needed to harm (NNH) represents the number of people who must be treated before 1 additional adverse event occurs. In order to calculate NNH, the absolute risk increase between the treatment and control groups must be known: $\text{NNH} = 1 / \text{Absolute risk increase}$

apps.uworld.com

Item 35 of 40 Question Id: 1210 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

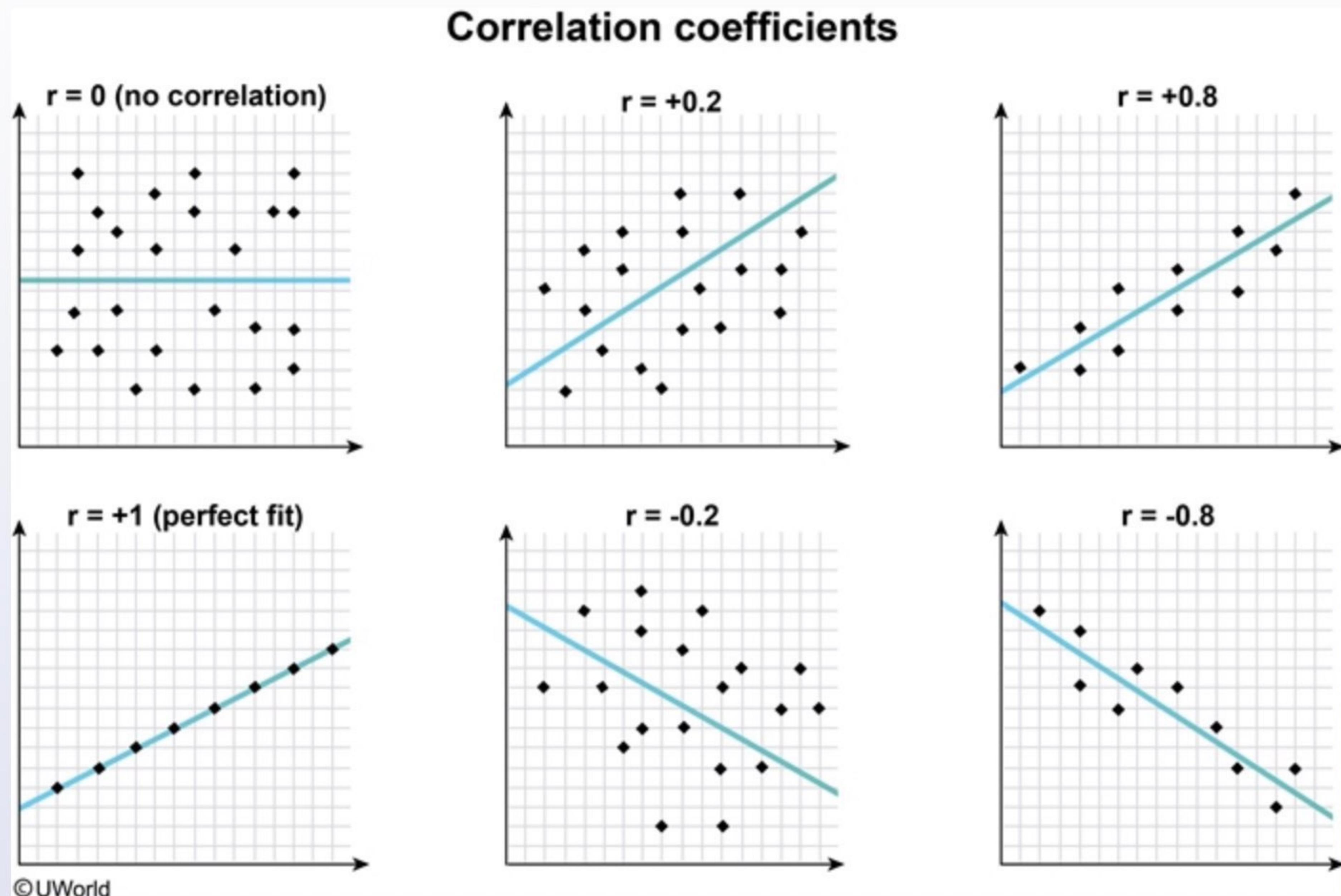
Inflammatory biological markers are clinically useful in a number of different ways, one being the assessment of disease activity in conditions such as systemic lupus erythematosus (SLE). A new inflammatory marker is being investigated in patients with active SLE flares. When the blood level of the new marker (in mg/L) is plotted against the C-reactive protein (CRP) level (also in mg/L), the following plot is obtained:



Based on the plot, the correlation coefficient between the 2 variables is closest to which of the following values?

- A. +0.8 (4%)
- B. +0.2 (5%)
- C. 0 (6%)
- D. -0.2 (10%)
- E. -0.8 (73%)

Item 35 of 40 Question Id: 1210 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings



Scatter plots are useful for crude data analysis. If a **linear** association is present between 2 variables, a **correlation coefficient** (r) mathematically describes how well a "line of best fit" (blue line in Figure) would correspond to the data points plotted. The value of r ranges from **-1 to +1** and describes 2 important characteristics of an association: the strength and the polarity. The closer the r value is to its margins $[-1, 1]$, the stronger the association.

An increase in C-reactive protein (CRP) level is associated with a decrease in the new marker's level; therefore, r

apps.uworld.com

Item 35 of 40 Question Id: 1210 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

Scatter plots are useful for crude data analysis. If a **linear** association is present between 2 variables, a **correlation coefficient** (r) mathematically describes how well a "line of best fit" (blue line in Figure) would correspond to the data points plotted. The value of r ranges from **-1 to +1** and describes 2 important characteristics of an association: the strength and the polarity. The closer the r value is to its margins $[-1, 1]$, the stronger the association.

An increase in C-reactive protein (CRP) level is associated with a decrease in the new marker's level; therefore, $r < 0$. Although the scatter plot does not demonstrate a perfect linear arrangement, it does show a reasonably strong linear association as there is minimal variation along the line of best fit (points are relatively close to the line). Therefore, the most appropriate answer among the options given is $r = -0.8$.

(Choices A and B) In this example, r cannot be >0 as an increase in CRP level is associated with a decrease in the new marker's level.

(Choice C) An $r = 0$ indicates that there is no association (ie, a random distribution).

(Choice D) An $r = -0.2$ would also indicate a negative association; however, there would be more variation along the line of best fit (as seen in the Figure).

The value of r is **not** the slope of the line of best fit. For instance, assuming a data set with a maximum positive association (ie, all data points forming a straight line with a positive slope), the line of best fit will always have $r = +1$, even if its actual slope is 0.2, 5.5, or any other positive value (as seen in the Figure).

Educational objective:

The correlation coefficient (r) ranges from **-1 to +1** and describes the strength and polarity of a linear association.

apps.uworld.com

Item 36 of 40 Question Id: 1205 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A study is conducted to assess the effect of beta-blocker therapy in patients with acute myocardial infarction (MI). The study found that 20 patients out of 90 who took a beta-blocker during the week prior to developing an MI had a major arrhythmia event during hospitalization. The study also determined that 30 patients out of 70 who did not take any beta-blocker developed a major arrhythmia. What was the odds ratio of developing major arrhythmia in patients who took beta-blockers compared to those who did not take beta-blockers?

- A. $(20 \times 30) / (70 \times 40)$ (5%)
- B. $(20 \times 40) / (70 \times 30)$ (53%)
- C. $(20 \times 70) / (30 \times 40)$ (6%)
- D. $(20 / 50) / (70 / 110)$ (2%)
- E. $(20 / 90) / (30 / 70)$ (32%)

Omitted
Correct answer
B

53%
Answered correctly

37 secs
Time Spent

2023
Version

Explanation

The odds of an event occurring are defined as the probability of that event happening divided by the probability of the event not happening. For instance, when [rolling dice](#), the probability of rolling a one is $1/6$, whereas the probability of rolling any other number is $5/6$. Therefore, the odds of rolling a one are $1/5$.

The **odds ratio (OR)** is a measure of association that compares the odds of an outcome occurring based on

apps.uworld.com

Item 36 of 40 Question Id: 1205 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

The odds of an event occurring are defined as the probability of that event happening divided by the probability of the event not happening. For instance, when [rolling dice](#), the probability of rolling a one is 1/6, whereas the probability of rolling any other number is 5/6. Therefore, the odds of rolling a one are 1/5.

The **odds ratio (OR)** is a measure of association that compares the odds of an outcome occurring based on exposure status. In this case, it represents the odds that the outcome (eg, major arrhythmia) occurred in the presence of the exposure (eg, beta blocker use) compared with the odds that the outcome occurred in the absence of that exposure. The OR can be calculated using a standard contingency table (with exposures in the rows and outcomes in the columns).

	Major arrhythmia	No arrhythmia	
Beta blocker (exposed group)	20 (a)	70 (b)	90
No beta blocker (unexposed group)	30 (c)	40 (d)	70
	50	110	

©UWorld

Odds of developing major arrhythmia in exposed group = 20/70

Odds of developing major arrhythmia in unexposed group = 30/40

(note that the total number of people in the exposed and unexposed groups is **not** used to calculate odds, unlike calculating probabilities)

apps.uworld.com

Item 36 of 40 Question Id: 1205 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

Odds of developing major arrhythmia in exposed group = 20/70

Odds of developing major arrhythmia in unexposed group = 30/40

(note that the total number of people in the exposed and unexposed groups is **not** used to calculate odds, unlike calculating probabilities)

Therefore, the OR of developing a major arrhythmia in patients who took beta blockers compared with those who did not is as follows:

$$(a/b) / (c/d) = (20/70) / (30/40) = (20 \times 40) / (70 \times 30) = 0.38$$

In this case, the OR calculated was the OR of disease (among exposed individuals relative to nonexposed individuals), given by $(a/b) / (c/d)$; in some cases, the OR of exposure (among diseased individuals relative to nondiseased individuals) is calculated, given by $(a/c) / (b/d)$. Both OR definitions are mathematically equivalent and reduce (by a [mathematical shortcut](#)) to: $OR = ad/bc$. Any of these formulas would give the same result.

Because the OR is <1 , the odds of developing a major arrhythmia is lower for patients who were taking beta blockers compared with those who were not.

(Choice E) Relative risk (RR) is the probability of an outcome occurring in the exposed group compared with the probability of the outcome occurring in the unexposed group. It is calculated using the following formula:

$$RR = [a/(a + b)] / [c/(c + d)] = (20/90) / (30/70)$$

RR often has more clinical utility than the OR since it gives the change in disease risk based on the presence or absence of risk factors. However, RR cannot be calculated in case-control studies as patients are preselected based on disease status (cases and controls) instead of being sorted into exposure groups and followed over time to assess disease outcomes.

apps.uworld.com

Item 36 of 40
Question Id: 1205

Mark

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

individuals), given by $(a/b) / (c/d)$; in some cases, the OR of exposure (among diseased individuals relative to nondiseased individuals) is calculated, given by $(a/c) / (b/d)$. Both OR definitions are mathematically equivalent and reduce (by a mathematical shortcut) to: $OR = ad/bc$. Any of these formulas would give the same result.

Because the OR is <1 , the odds of developing a major arrhythmia is lower for patients who were taking beta blockers compared with those who were not.

(Choice E) Relative risk (RR) is the probability of an outcome occurring in the exposed group compared with the probability of the outcome occurring in the unexposed group. It is calculated using the following formula:

$$RR = [a/(a + b)] / [c/(c + d)] = (20/90) / (30/70)$$

RR often has more clinical utility than the OR since it gives the change in disease risk based on the presence or absence of risk factors. However, RR cannot be calculated in case-control studies as patients are preselected based on disease status (cases and controls) instead of being sorted into exposure groups and followed over time to assess disease outcomes.

Educational objective:

The odds ratio (OR) is a measure of the strength of an association between an exposure and an outcome. Specifically, it represents the odds that an outcome occurred in the presence of a particular exposure compared with the odds that the outcome occurred in the absence of that exposure. In a standard contingency table, $OR = ad/bc$.

Biostatistics
Subject

Biostatistics & Epidemiology
System

Odds ratio
Topic



AA

apps.uworld.com



Item 37 of 40

Question Id: 1207

Previous NextFull ScreenTutorialLab ValuesNotesCalculatorReverse ColorA A ASettings

A study evaluated the role of angiotensin-converting enzyme (ACE) inhibitors in the prevention of coronary events in patients with diabetes. During 5 years of follow-up, 120 out of 400 diabetic patients who had taken an ACE inhibitor developed an acute coronary event. Over the same time, 100 out of 300 diabetic patients who had not taken an ACE inhibitor experienced coronary events. What was the relative risk of developing a coronary event in diabetic patients who were taking ACE inhibitors compared to diabetic patients who were not taking ACE inhibitors?

- A. $(120 \times 100) / (280 \times 200)$ (3%)
- B. $(120 \times 200) / (100 \times 280)$ (11%)
- C. $(120 \times 280) / (100 \times 200)$ (4%)
- D. $(120/220) / (280/480)$ (3%)
- E. $(120/400) / (100/300)$ (77%)

Omitted
Correct answer
E

77%
Answered correctly

34 secs
Time Spent

2023
Version

Explanation

It is important to know how to calculate some basic measures of association when given raw data. The first step is to organize the data into a 2×2 (contingency) table as shown below. The [standard format](#) includes the exposures in the rows (with "exposure present" listed above "exposure absent") and the outcomes in the

apps.uworld.com

Item 37 of 40 Question Id: 1207 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

It is important to know how to calculate some basic measures of association when given raw data. The first step is to organize the data into a 2×2 (contingency) table as shown below. The [standard format](#) includes the exposures in the rows (with "exposure present" listed above "exposure absent") and the outcomes in the columns (with "outcome present" listed before "outcome absent").

	Coronary event	No coronary event	
ACE inhibitor	120 (a)	280 (b)	400
No ACE inhibitor	100 (c)	200 (d)	300
	220	480	700

The relative risk (RR) represents the risk of an outcome (eg, coronary event) in the exposed (eg, angiotensin-converting enzyme [ACE] inhibitor) divided by the risk of that outcome in the unexposed (eg, no ACE inhibitor). For cohort studies, it is calculated using the following formula:

$$RR = [a/(a+b)] / [c/(c+d)] = (120/400)/(100/300) = 0.90$$

Understanding the values calculated (rather than memorizing several formulas) would prevent errors that could occur if the standard format for the 2×2 table is not followed. By definition, RR = risk among the exposed/risk among the unexposed. In this example, 400 diabetic patients were taking ACE inhibitors (exposed) and 120 of those had a coronary event. Therefore, the risk among the exposed is 120/400. Similarly, 300 diabetic patients were not taking ACE inhibitors (unexposed) and 100 of those had a coronary event. Therefore, the risk among the unexposed is 100/300. The RR is the ratio of those 2 values: $RR = (120/400)/(100/300)$.

apps.uworld.com

Item 37 of 40 Question Id: 1207 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

220 480 700

The relative risk (RR) represents the risk of an outcome (eg, coronary event) in the exposed (eg, angiotensin-converting enzyme [ACE] inhibitor) divided by the risk of that outcome in the unexposed (eg, no ACE inhibitor). For cohort studies, it is calculated using the following formula:

$$RR = [a/(a+b)] / [c/(c+d)] = (120/400)/(100/300) = 0.90$$

Understanding the values calculated (rather than memorizing several formulas) would prevent errors that could occur if the standard format for the 2×2 table is not followed. By definition, RR = risk among the exposed/risk among the unexposed. In this example, 400 diabetic patients were taking ACE inhibitors (exposed) and 120 of those had a coronary event. Therefore, the risk among the exposed is 120/400. Similarly, 300 diabetic patients were not taking ACE inhibitors (unexposed) and 100 of those had a coronary event. Therefore, the risk among the unexposed is 100/300. The RR is the ratio of those 2 values: $RR = (120/400)/(100/300)$.

(Choice B) The odds ratio (OR) is calculated using the following formula: $OR = ad/bc$. The OR is often calculated in case-control studies where RR cannot be obtained because such studies do not follow patients over time to determine the incidence of a new outcome.

Educational objective:

The relative risk (RR) represents the risk of an outcome in the exposed divided by the risk of that outcome in the unexposed. Applying the correct formula for RR calculations depends on the proper formatting of a 2×2 (contingency) table.

apps.uworld.com

Item 38 of 40 Question Id: 1204 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A study determines that the mean blood cholesterol level is 195 mg/dL in 200 non-diabetic hospitalized patients and 210 mg/dL in 180 diabetic hospitalized patients. The probability that the observed difference is due to chance alone is reported to be 5%. There is also a 20% probability of concluding that there is no difference in blood cholesterol level when there is one in reality. What is the power of the study?

- A. 0.05 (7%)
- B. 0.20 (6%)
- C. 0.50 (3%)
- D. 0.80 (72%)
- E. 0.95 (10%)

Omitted
Correct answer
D

72%
Answered correctly

39 secs
Time Spent

2023
Version

Explanation

The **power** of a study is the ability of a study to detect a difference between groups when such a difference truly exists. Power is related to **type II error (β)**, which is the probability of concluding there is no difference between groups when one truly exists. Mathematically, power is given by:

$$\text{Power} = 1 - \beta$$

apps.uworld.com

Item 38 of 40 Question Id: 1204 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

exists. Power is related to **type II error (β)**, which is the probability of concluding there is no difference between groups when one truly exists. Mathematically, power is given by:

$$\text{Power} = 1 - \beta$$

In this example, the power of the study is the probability of detecting a difference in blood cholesterol level between diabetics and non-diabetics if there is a real difference. The probability of concluding that there is no difference in blood cholesterol level when in reality there is one is given as 20%; this corresponds to the definition of β (ie, $\beta = 0.20$ in this example). Therefore:

$$\text{Power} = 1 - \beta = 1 - 0.20 = 0.80$$

(Choice A) Type I error (α) describes the probability of seeing a difference when there is no difference in reality. The value of α is generally compared to the probability that the observed difference is due to chance alone (a simplified explanation of the p-value). In this example, the probability that the observed difference between diabetic and nondiabetic patients is due to chance alone is given as 5% (0.05).

(Choices B and E) Type II error (β) is 0.20, as explained above. The value 0.95 corresponds to $(1 - \alpha)$, but power is given by $(1 - \beta)$.

Educational objective:

The power of a study indicates the probability of seeing a difference when there is one. The formula is Power = $1 - \beta$, where β is the type II error rate.

Biostatistics
Subject

Biostatistics & Epidemiology
System

Power and sample size
Topic

apps.uworld.com

Item 39 of 40 Question Id: 1272 Mark Previous Next Full Screen Tutorial Lab Values Notes Calculator Reverse Color Text Zoom Settings

A 62-year-old man with diabetes, hypertension, and hyperlipidemia comes to the emergency department of an academic medical center with chest pain, nausea, vomiting, and diaphoresis. An electrocardiogram demonstrates ST elevation in the anterior leads, and cardiac enzymes are markedly elevated. Investigators at the center are designing a randomized control trial to test the hypothesis that drug B will decrease the mortality associated with acute ST-elevation myocardial infarction compared to standard of care. To ensure that investigators will not miss a difference between drug B and standard of care (if a difference truly exists), which of the following would they want to maximize?

- A. α (6%)
- B. β (9%)
- C. Type I error (6%)
- D. Type II error (5%)
- E. $1 - \beta$ (72%)

Omitted
Correct answer
E

72%
Answered correctly

08 secs
Time Spent

2023
Version

Explanation

Statistical **power**, $(1 - \beta)$, represents a study's ability to detect a difference when one exists. It is the probability of rejecting the null hypothesis when it is truly false - ie, the probability of finding a true relationship. Power

Item 39 of 40
Question Id: 1272

Mark

Previous Next

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

Statistical **power**, $(1 - \beta)$, represents a study's ability to detect a difference when one exists. It is the probability of rejecting the null hypothesis when it is truly false - ie, the probability of finding a true relationship. Power depends on sample size and the difference in outcome between the groups being tested. In this study, the researchers want to detect a difference between drug B and standard of care if one exists; they want to maximize power.

β is the probability of committing a type II error (**Choices B and D**). Type II error occurs when researchers fail to reject the null hypothesis when it is truly false. This causes investigators to miss true relationships. An example of a type II error would be a study finding that aspirin does not affect platelet function when, in fact, it does. Therefore, if β is set at 0.2, the power will be $(1 - \beta) = 80\%$; there will be an 80% chance of rejecting the null hypothesis when it is truly false.

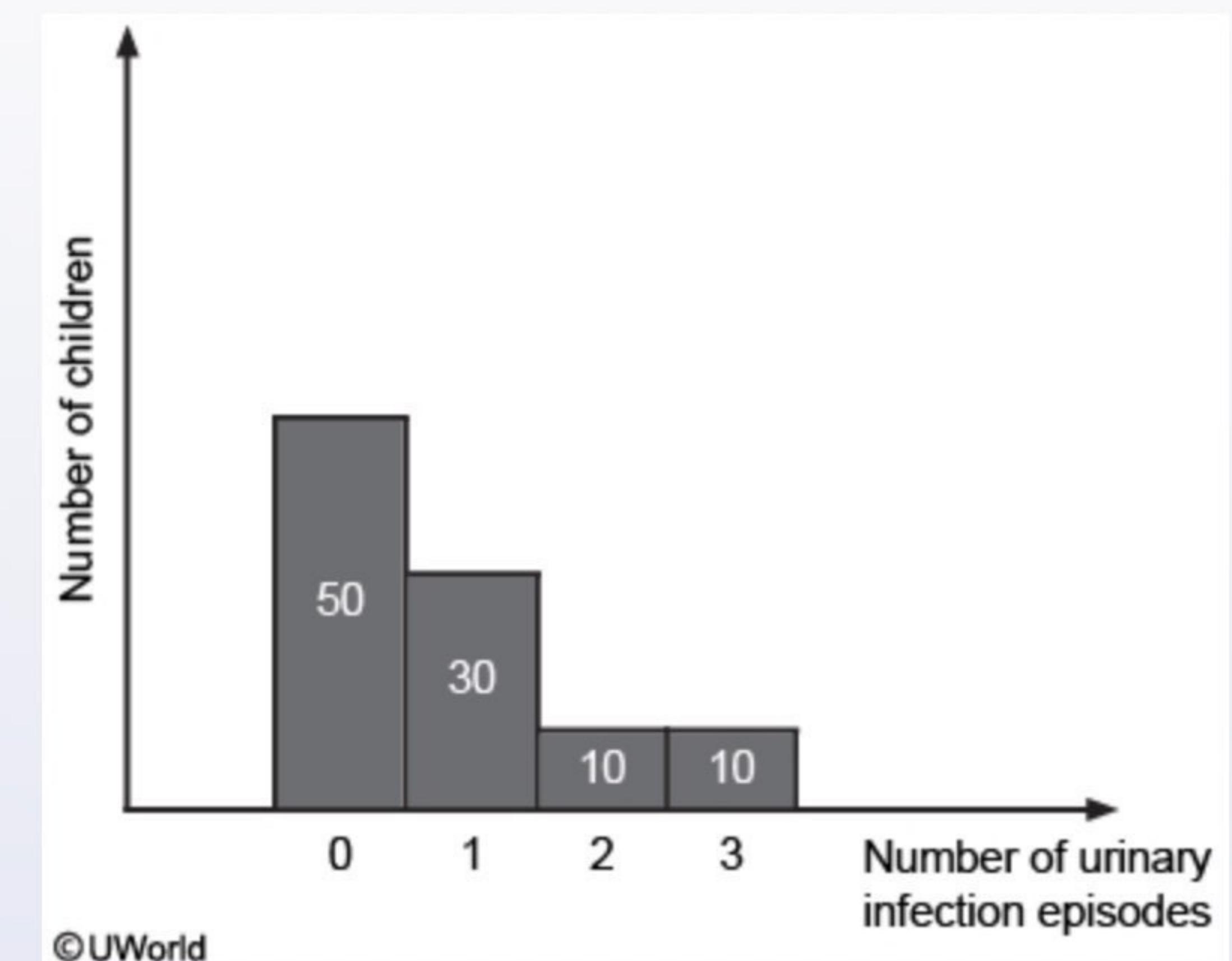
(Choices A and C) Type I error occurs when researchers reject the null hypothesis when the null hypothesis is really true. That is, a study finds a statistically significant difference between 2 groups when one does not truly exist. An example would be if a study concluded that candy improves heart failure mortality when, in fact, it does not. α is the maximum probability of making a type I error that a researcher is willing to accept. Generally, α is compared to the p-value, the probability of observing a given result (or more extreme) due to chance alone assuming the null hypothesis is true (eg, if there is no real difference between the groups). The value of α is typically set at 0.05, meaning that researchers are willing to accept up to a 5% chance of making a type I error. In such a scenario, if $p < 0.05$, the result is said to be statistically significant.

Educational objective:

Power $(1 - \beta)$ is the probability of rejecting a null hypothesis when it is truly false. It is typically set at 80% and depends on sample size and difference between outcomes.

Measures and distribution of data

A sample of children age 2-5 years is chosen at random from an outpatient clinic. The number of urinary tract infections (UTIs) over a 1-year period is given in the figure below.



What was the average number of UTI episodes over 1 year for a child in this sample?

- A. Between 0 and 1
 - B. 1
 - C. Between 1 and 2
 - D. 2
 - E. Between 2 and 3

[Proceed To Next Item](#)

Researchers conducted a randomized controlled trial to assess the effectiveness of a new drug to reduce severe cutaneous reactions in patients with rheumatoid arthritis who are treated with adalimumab. Of 150 adults with rheumatoid arthritis who are treated with adalimumab, 75 received the new drug in addition to adalimumab, and 75 received a placebo in addition to adalimumab. Results show that 6 patients in the new drug group developed severe cutaneous reactions, compared to 9 in the placebo group. Which of the following represents the relative risk reduction for severe cutaneous reactions among patients in the new drug group?

- A. 0.08
- B. 0.10
- C. 0.12
- D. 0.33
- E. 0.67

[Proceed To Next Item](#)



tion Id: 19308

98%

98%

A study assessed the association between a new vaccine and traveler's diarrhea (TD). Researchers selected a random sample of people who intended to travel to regions where they were at increased risk for TD and who had received the new vaccine and another independent random sample of people who intended to travel to the same regions and who had not received the new vaccine. These 2 samples of travelers were assessed for the occurrence of TD during the trip and for 7 days after returning home. Which of the following measures of association are the investigators most likely to report?

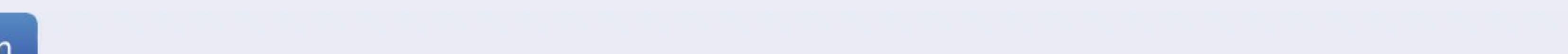
- A. Incidence
- B. Median survival time
- C. Odds ratio
- D. Prevalence
- E. Relative risk

[Proceed To Next Item](#)



A large study of serum folate levels in a sample of women age 16-45 reveals that this parameter is normally distributed with a mean of 5.0 ng/mL and a standard deviation of 0.5 ng/mL. According to the study results, 95% of serum folate observations in these patients will lie approximately between which of the following limits?

- A. 3.5 and 6.0 ng/mL
- B. 3.5 and 6.5 ng/mL
- C. 4.0 and 6.0 ng/mL
- D. 4.0 and 5.5 ng/mL
- E. 4.5 and 5.5 ng/mL


[Proceed To Next Item](#)

A new portable cholesterol-measuring device is being developed for use in medically underserved communities. During clinical trials of an early prototype, a patient's cholesterol level is found to be 200 mg/dL, 201 mg/dL, and 200 mg/dL on 3 separate measurements of the same blood sample. Using the gold standard measurement method, the same sample is found to have a cholesterol level of 260 mg/dL. Which of the following descriptions best characterizes the new cholesterol-measuring device?

- A. High accuracy; high precision
 - B. High accuracy; low precision
 - C. High sensitivity; low specificity
 - D. Low accuracy; high precision
 - E. Low accuracy; low precision
 - F. Low sensitivity; high specificity

[Proceed To Next Item](#)



A 65-year-old man with a history of congestive heart failure is hospitalized with chest pain and hypotension requiring admission to the cardiac care unit. An intra-arterial line is placed for direct blood pressure monitoring. Consecutive readings of his intra-arterial blood pressure are 75, 110, 80, 90, 75, and 110 mm Hg. Which of the following represents the median of these blood pressure readings?

- A. 80 mm Hg
- B. 85 mm Hg
- C. 90 mm Hg
- D. 100 mm Hg
- E. 110 mm Hg

Proceed To Next Item

An intern in the endocrinology department is asked to complete a chart review of cholesterol levels of all patients hospitalized with diabetes mellitus-related cardiovascular complications in the past 6 months. The intern's preliminary data analysis reveals that there were 400 such patients, and that serum cholesterol levels were normally distributed among the patients with a mean of 220 mg/dL and a standard deviation of 10 mg/dL. Based on these results, how many patients in this study would be expected to have serum cholesterol ≥ 240 mg/dL?

- A. 2
- B. 10
- C. 20
- D. 64
- E. 128

[Proceed To Next Item](#)

X

Blood pressure measurements are obtained from a sample of individuals with no known medical conditions. For systolic blood pressure (SBP), the mean measurements and associated standard deviations (SDs) are shown by age group for men and women:

	Men	Women
Age group	Mean SBP (mm Hg) \pm SD	Mean SBP (mm Hg) \pm SD
35-44	120 ± 20	124 ± 18
45-54	131 ± 21	137 ± 24
55-64	141 ± 19	140 ± 20

If hypertension is defined as SBP >140 mm Hg, approximately what percentage of men age 35-44 in this sample will be classified as having hypertension, assuming a normal (Gaussian) distribution?

- A. 16%
- B. 34%
- C. 50%
- D. 68%
- E. 84%
- F. 95%

[Proceed To Next Item](#)



tion Id: 108029



Previous

Next

Exit Mode

Tutorial

Lab Values

Notes

Calculator

Reverse Color

Text Zoom

Settings

A researcher selects a random sample of 100 men age 18-24 and determines their mean serum total cholesterol is 180 mg/dL with a standard deviation of 40 mg/dL. According to national statistics, cholesterol levels for the population of men age 18-24 follow a normal (gaussian) distribution. Based on this information, approximately 50% of the men in the sample will have which of the following serum total cholesterol levels?

- A. 100-260 mg/dL
- B. <140 mg/dL or >220 mg/dL
- C. 140-220 mg/dL
- D. <180 mg/dL
- E. >260 mg/dL

[Proceed To Next Item](#)

spend

End Block

Pulmonary capillary wedge pressure (PCWP) measurements can be used to estimate left atrial pressure; the normal range is between 6-12 mm Hg, and recorded values are whole numbers. A patient in the intensive care unit has 20 serial PCWP measurements taken over the course of 2 hours. Among these 20 observations, the maximal recorded value is 12 mm Hg and the minimal recorded value is 10 mm Hg. If the next measurement is 26 mm Hg, which of the following is most likely to remain unchanged?

- A. Mean
 - B. Mode
 - C. Range
 - D. Standard deviation
 - E. Variance

[Proceed To Next Item](#)



A medical student is conducting a chart review of patients admitted to the local emergency department with acute pancreatitis. He is interested in comparing blood glucose levels between mild and severe cases. Before attempting to compare the mean blood glucose levels between the 2 groups of patients, he decides to conduct a descriptive analysis of the variables in each group. Blood glucose levels in patients with severe acute pancreatitis were found to have a strong positively skewed distribution. Which of the following is most likely to be true regarding the data for the severe acute pancreatitis group?

- A. The mean is equal to the median
- B. The mean is equal to the mode
- C. The mean is greater than the median
- D. The median is greater than the mean
- E. The mode is greater than the mean

[Proceed To Next Item](#)

X

tion Id: 1209

Mark

Previous

Next

Full Screen

Tutorial

Lab Values

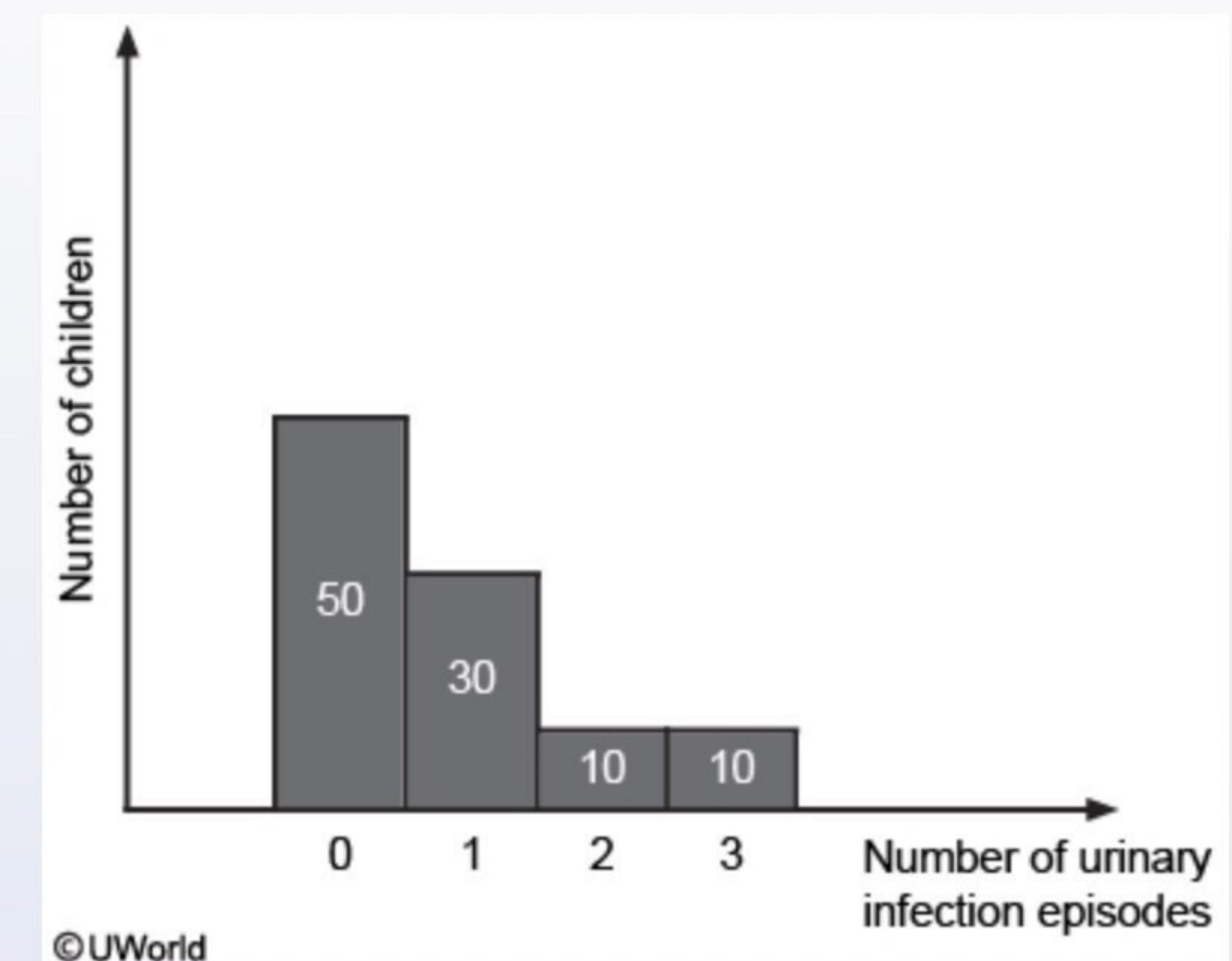
ABC
Notes0.25
Calculator

Reverse Color

Text Zoom

Settings

A sample of children age 2-5 years is chosen at random from an outpatient clinic. The number of urinary tract infections (UTIs) over a 1-year period is given in the figure below.



What was the average number of UTI episodes over 1 year for a child in this sample?

- A. Between 0 and 1 (75%)
- B. 1 (4%)
- C. Between 1 and 2 (16%)
- D. 2 (1%)
- E. Between 2 and 3 (2%)

Omitted
Correct answer

75%
Answered correctly

16 secs
Time Spent

2023
Version



- D. 2 (1%)
- E. Between 2 and 3 (2%)

Omitted

Correct answer
A75%
Answered correctly16 secs
Time Spent2023
Version

Explanation

This bar graph illustrates the frequency of urinary tract infection (UTI) episodes in the sample. According to the figure, 50 children had 0 UTIs, 30 children had 1 UTI, 10 had 2 UTIs, and 10 had 3 UTIs.

In general, the average (or mean) of a dataset is the sum of the values divided by the total number of values. In this example, the average number of UTI episodes per child is the sum (ie, total number) of UTIs divided by the total sample size (ie, total number of children). The total sample size is: $50 + 30 + 10 + 10 = 100$ children. The total number of UTIs is: $(0 \times 50) + (1 \times 30) + (2 \times 10) + (3 \times 10) = 0 + 30 + 20 + 30 = 80$ UTIs. The average is obtained by dividing the total number of UTIs (80) by the total sample size (100). Therefore, the average number of UTI episodes per year = $80/100 = 0.8$ UTIs for a child in this sample.

In conclusion, children age 2-5 years in this clinic experienced between 0 and 1 UTI over 1 year on average.

Educational objective:

The average (or mean) of a dataset of values is the sum of the values divided by the total number of values.

Biostatistics
Subject

Biostatistics & Epidemiology
System

Central tendency (mean, median, mode, outliers)
Topic

Copyright © UWorld. All rights reserved.



dback

End Block