SCIENCE OLYMPIAD

Exploring the World of Science

Disease Detectives

The starred questions can be used as tie breakers

Total Points: 212

Part 1: Lyme Disease

Lyme disease is a multisystem illness caused by Borrelia burgdorferi, a spirochete transmitted by certain species of Ixodes ticks. Approximately 30,000 confirmed and probable cases of Lyme disease were reported in the United States in 2012, primarily from high-incidence states in the Northeast (Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont) and upper Midwest (Minnesota and Wisconsin) (1,2).† Common manifestations include cutaneous, neurologic, and rheumatologic signs and symptoms.

Symptomatic infection of the heart is rare in recognized Lyme disease cases and usually resolves promptly with appropriate antibiotic therapy. Nonetheless, cardiac involvement occasionally can cause life-threatening cardiac conduction abnormalities. During November 2012–July 2013, 10 women and fifteen men (ranging in age from 26 to 38 years) from high-incidence Lyme disease states experienced sudden cardiac death and, on postmortem examination, were found to have evidence of Lyme carditis. 10 of the men and 3 of the women were not wearing any anti mosquito repellant. 5 of the men and 7 of the women were outside for less than 10 hours in tick season. The deaths were investigated by the Connecticut Department of Public Health, Massachusetts Department of Public Health, New Hampshire Department of Public Health, New York State Department of Health, and CDC. As a disease detective, you are trying to figure out the cause of these sudden deaths and this rare form of Lyme disease.

Adapted for this exercise from cdc.gov

1. (3 points) Name the 3 things epidemiologists use to characterize an outbreak.

Person
Place
Time

2. (2 points) Why were the people infected not elderly or children, who are usually more susceptible to disease?

This type of disease is related to outdoor activities and the elderly would not be doing this.

*3. (6 points) Name the three common types of studies you could use to analyze THIS TYPE OF OUTBREAK. (HINT: Pay attention to the relationship between time and disease status for this outbreak)

RETROSPECTIVE Cohort (MUST say retrospective)
Case Control
Cross Sectional

4. (3 points) What is the difference between a prospective and a retrospective study?

Prospective study participants are enrolled before health outcome occurs, retrospective starts with outcome and works back to potential causes.

5. (3 points) What are three advantages of a cohort study?

Temporal sequence is proper, study multiple outcomes in one study, easy calculations of risk, find incubation period of disease, good for rare exposure (any 3)

6. (3 points) What are three disadvantages of a case control study?

No temporality, bias in control selection, information bias, only a single outcome is studied (any 3).

7. (3 points) What are the three components of the Epidemiological Triad?

Agent
Host
Environment

8. (1 point) Name one of the rates that any of the above study types use

Attack rate, odds ratio, relative risk (any of these is fine)

You decide to conduct a study in which you match a group of people to the infected people and see if there is any specific exposure that may be the cause of this disease.

9. (3 points) What type of study design are you using?

Case control

10. (5 points) Name 3 ways you can reduce bias in this study design.

Match cases and controls based on factors such as age, race, sex. (Must have some sort of matching)
The other 2 ways can be anything logical, like making sure people are not picked from one area, they are not predisposed to disease, there is a large number of people, etc.

You recruit 10 women and 15 men who were not diseased to use as a comparison group.

11. (1 point). What is this group called?

Control group

12. (5 points) 4 of the healthy women and 3 of the healthy men were not wearing any anti mosquito repellent; draw a 2x2 table for this exposure.

	Case	**Control**
Exposed	13	7
Unexposed	12	18

*13. (4 points) Calculate the relative risk for this exposure. (Please show your work)

You should not calculate relative risk because this is a case control and this is not the correct measure of association.

*14. (4 points) Calculate the odds ratio for this exposure. (Please show your work)

Ad/bc; (13*18)/(12*7) = 2.79

15. (4 points) Is there an increased association between exposure and sickness? Explain why using the above calculation(s).

Yes, there is a 2.79 times more likely chance that those who were sick were exposed (as in, not wearing mosquito repellent). The odds ratio calculates for this.

15b. (3 points) In an outbreak or an epidemic, evidence is accumulated linking disease to a causative organism or substance. What is the name of the criteria used to prove that an organism causes a disease?

Bradford Hill Criteria

16. (5 points) List 4 types of data collected and used by epidemiologists

Demographic, Identification, Risk Factors, Clinical Disease Data

17. (5 points) 10 of the healthy women and 5 of the healthy men were outside for less than 10 hours during tick season; draw a 2x2 table for this exposure.

	Case	**Control**
Exposed	12	15
Unexposed	13	10

*18. (4 points) Calculate the relative risk for this exposure. (Please show your work)

You should not calculate relative risk because this is a case control and this is not the correct measure of association.

*19. (4 points) Calculate the odds ratio for this exposure. (Please show your work)

Ad/bc; (12*10)/(13*15) = 0.62

*20. (7 points) Please explain how to calculate the McNemar test for paired data and what this test is.

(give partial credit for this explanation if it is not detailed enough)
It is a paired version of the Chi-square test for matched pair data which is used to compare paired proportions. It can be used to analyze retrospective case control studies and it calculates a P value. It only uses the number of discordant pairs (the number of pairs for which the control was exposed to the risk but the case was not) and the number of pairs where case was exposed and control was not (R and S). Calculate Chi Square using

$$\chi^2 = \frac{(|R-S|-1)^2}{R+S}$$

21. (6 points) For this exposure, please calculate the log odds ratio and show your work.

=log(odds ratio) = log(.62) = -.207

22. (8 points) For this exposure, please calculate the bias corrected log odds ratio and show your work.

Give 2 points for writing formula

= l'(o) = LOG[{(N11+0.5) (N22+0.5)}/ {(N12+0.5) (N21+0.5)}]
Log{(12+.5)(10+.5)}/{(13+.5)(15+.5)}
=0.202

23. (4 points) Is there an increased association between exposure and sickness? Explain why using the above calculation(s).

No, there is a 0.62 times more likely chance that staying outside for less than 10 hours causes the disease. This means that it could be a protective

effect, because the more time that is spent outdoors the more likely getting the disease is.

24. (3 points) After you have evaluated your hypotheses, what steps should you carry out in order to further characterize this outbreak?

Refine hypothesis, carry out control and prevention and communicate findings

Part 2: Severe Stomach Illness

Hotel Sketc had its annual dinner party and served delicious food. After the party, many attendants reported feeling sick with diarrhea, vomiting, and gastrointestinal irritation. As disease detectives, you decide to investigate this outbreak of sudden illness.

	Cases				**Controls**			
Food	**Ate**	**Did Not Eat**	**Total**	**%Ate**	**Ate**	**Did Not Eat**	**Total**	**%Ate**
Baked Ham	29	17	46	**63.04**	17	12	**29**	58.62
Spinach	26	20	46	**56.52**	17	**12**	29	58.62
Mashed potatoes	23	**23**	46	50.00	**14**	**15**	29	48.28
Jell-O	**16**	**30**	46	34.78	**7**	**22**	29	24.14
Rolls	**21**	25	46	45.65	**16**	**13**	29	55.17
Cake	27	19	46	**58.70**	**13**	**16**	29	44.83
Ice cream	**43**	**3**	46	93.48	11	18	29	**37.93**
Milk	2	**44**	46	4.35	2	27	29	**6.90**
Water	**13**	**33**	46	28.26	11	18	29	**37.93**

25. (8 points) Fill in the missing parts of the table.

26. (2 points) Which food has the highest risk of illness associated with it?

Ice Cream

27. (2 points) Which food had the lowest risk of illness in the case group associated with it?

Milk

28. (3 points) Calculate the odds ratio of the food with the highest risk. Show your work.

OR = AD/BC=(43*18)/(11*3) = 23.5

29. (3 points) If all of the people who were controls happened to be, by chance, immune to stomach illness, what would this be called?

Random Error

30. (2 points) If the person who interviewed the cases encouraged the cases to report additional food consumed, what would this be called?

Interviewer Bias

31. (3 points) If the disease caused the memory of the cases to be hazy which led the cases to incorrectly characterize their exposure status, what would this be called?

Recall bias

32. (2 points) Based on these results, what would you consider the cause of the outbreak?

Ice cream

33. (4 points) List Koch's Postulates

1. The microorganism must be found in abundance in all organisms suffering from the disease, but should not be found in healthy animals. 2. The microorganism must be isolated from a diseased organism and grown in pure culture. 3. The cultured microorganism should cause disease when introduced into a healthy organism. 4. The microorganism

must be reisolated from the inoculated, diseased experimental host and identified as being identical to the original specific causative agent.

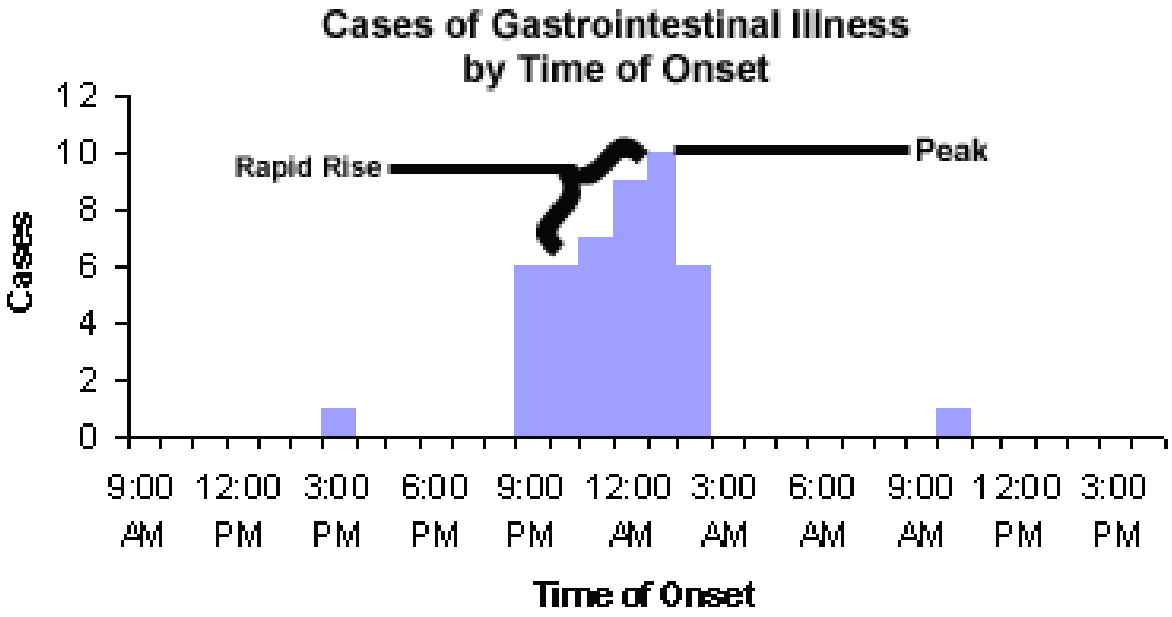
34. (2 points) If the food that caused the outbreak contained *E. coli* and the other food did not, would this satisfy Koch's Postulates?

Yes

35. (2 points) If the *E. coli* from the food was cultured an placed into a healthy organism but did not cause disease, would this satisfy Koch's Postulates?

No

You created an Epi Curve of a subset of your cases.



36. (2 points) What is the time of onset of the first case?

3:00pm

37. (2 points) What type of epi curve is this and why?

Point source- there is a rapid rise and only one peak

38. (1 point) What is the peak number of cases?

10 cases

39. (2 points) Using the epi curve, estimate the incubation period.

6-12 hours

Part 3: Happy Farms Drug Company

You're a member of the Happy Farms Drug Company testing team and you need to see how effective recent tests have been in curing Advanced Dermatitis. You conduct an experimental study.
Your test was administered to 100 people and these 100 people were enrolled in a study to test your drug, but only 50 people in the end provided you with data.

40. (3 points) What is the gold standard for clinical trial tests?

Randomized Controlled Trials

41. (3 points) What are the people called who did not give your data?

Lost to follow up

Out of the 50 people that gave you data, 15 people were diagnosed with the disease and correctly identified by the test as having the disease. 25 people overall had the disease.
10 people without the disease were marked by your test as having the test.

42. (5 points) Construct a 2x2 table with these numbers.
The table can be flipped, so that the positive/negative appears on the left hand side.

	Positive	Negative
Disease	15 True Positive	10 False Negative
No Disease	10 False Positive	15 True Negative

43. (1 point) What are the people that do not have the disease but marked positive by the test called?

False Positive

44. (1 point) What are the people that do have the disease but were not marked positive by the test called?

False Negative

45. (1 point) What are the people that do have the disease and were correctly marked positive by the test called?

True positive

46. (1 point) What are the people that do not have the disease and were marked negative by the test called?

True negative

47. (2 points) Please label your 2x2 table with the terms from 28-31

48. (3 points) Calculate the positive predictive value; show work and explain what this number means.

True positive/ (true positive + false positive) = 15/(15+10) = 0.6
Number of true positives over total positives, how effective the test is.

49. (3 points) Calculate the negative predictive value; show work and explain what this number means.

True negative / (true negative + false negative) = 15/(10+15) = .6
Number of true negatives over total negatives, how effective the test is in identifying correct negatives.

50. (6 points) Calculate the sensitivity and specificity; show work and explain what these numbers mean.

Sensitivity = true positive/(true positive + false negative) = 15/(15+15) = .5
Probability that the test says a person has the disease when in fact they do have the disease ability of test to identify positive results

Specificity = true negative/(true negative + false positive) = 15/(15+10) = .6
Probability that the test says a person does not have the disease when in fact they are disease free

51. (6 points) Is this test a good indicator of the disease? Why or why not? Explain using your numbers from above.

No, the sensitivity and specificity are rather low and the PPV and NPV are low as well. It only has a .5 probability that the disease will recognize true positive results and a .6 probability that it will recognize true negative results.

52. (3 points) If a disease has a low cost of treatment but is very harmful to health, would you rather have a lot of false negatives or false positives? Why?

False positives, more people would be treated and less people would be sick. Since the treatment is cheap and the disease is really detrimental, you would want to treat more people than rather let some people who are ill go free.

Part 4: General Epidemiology Knowledge and Advanced Statistics

For 53-58, use the following information.
A suspected outbreak of *E. coli* occurred at the Right State University campus. On September 1st, 15 people were sick. On each of the following days, 3 more people got sick until September 5th.

53. (2 points) Calculate the average number of people who got sick per day. Show your work.

Mean = (sum of all terms) / number of terms = (15+18+21+24+27)/5 =21

54. (2 points) Calculate the variance. Show your work, the formula, and label your variables.

$$S^2 = \frac{\sum (x - \bar{x})^2}{n-1}$$

X = each score
$\bar{X}$ = the mean or average
n = the number of values
Σ means we sum across the values

=22.5

55. (4 points) Calculate the standard deviation. Show your work, the formula, and label your variables.

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$$

X = each score
$\bar{X}$ = the mean or average
n = the number of values
Σ means we sum across the values

Sum$(27-21)^2$ $(24-21)^2$ etc /4
=4.74
OR just square root the variance.

56. (2 points) Calculate the standard deviation of the mean. Show your work, the formula, and label your variables.

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

S = standard deviation, n= number of terms
=2.12

57. (3 points) Write the Z scores for the 90%, 95%, and 99% Confidence Interval

90% 1.645
95% 1.96
99% 2.576

58. (5 points) Calculate the 90% confidence interval of the mean. Show your work, the formula, and label your variables.

$$\overline{X} \pm Z\frac{s}{\sqrt{n}}$$

X bar = mean, Z = z score, right hand side is standard deviation of the mean.

=21±3.49

59-73: Match the following vocabulary words to the definitions. (1 point each)

G 59. Agent	A. a widespread outbreak of an infectious disease in a specified community
K 60. Bacteria	B. system of doctors and health officials collecting and comparing data on various diseases or infections within communities
A 61. Epidemic	C. ultramicroscopic infectious agent that replicates itself only within cells of living hosts
L 62. Outbreak	D. an organism that transmits a pathogen
C 63. Virus	E. a disease causing agent
O 64. Symptom	F. invasion of the body by pathogenic agents
B 65. Public Health Surveillance	G. a substance that causes an effect
N 66. Exposure	H. an inanimate object that can be used in the transmission of diseases
F 67. Infection	I. time of the appearance of the first symptoms of an illness
E 68. Pathogen	J. any disease of animals communicable to humans
M 69. Pandemic	K. microscopic, single-celled organisms that lack chlorophyll and nuclei
I 70. Onset	L. a sudden occurrence of disease in two or more people during a specified period of time
D 71. Vector	M. a disease that occurs over a wide geographic area and affecting an exceptionally high proportion of people
J 72. Zoonosis	N. to come in contact with an infectious agent in a manner that promotes transmission and the likelihood of disease
H 73. Fomite	O. evidence or sign of disease or infection

74. (2 points) Who is the father of modern epidemiology?

John Snow