

**BE.104 Spring '05**  
**Problem Set 2 Answers – 10pts Total**  
**Sherley**

1. After visiting each household in their defined public health sector (A) during the months of July-August, your team of healthcare monitors documents 40 cases of an unusual skin rash associated with headaches. Sector A is comprised of 500 households located in one section of the city of Boston.

What information does your team need so that they can develop estimates of the prevalence and incidence of skin rashes in sector A?

**Answer: The number of individuals at risk (i.e., population) is needed for both estimates. 0.25 pt**  
**For incidence, you also need to know how many of the cases are new within the defined period. 0.25 pt**

*Insight: If you knew that each household had only one case, you could develop an analogous prevalence estimate based on households at risk: e.g., “cases/household from July-August”*

*However, in the absence of this knowledge the “average # of cases/household from July-August” could be very misleading. For example, they may all have occurred in a small number of households and someone getting the prevalence data wouldn't know it. Same analysis for new cases/household and incidence.*

2. The team leader for sector B calls you because you are team leader of sector A. The B leader wants to know if your team has observed any case of an unusual skin rash associated with headaches. Leader B reports that they have observed 60 cases in their 1000 households. You both also relate that 95% of the observed cases are children.

What information do you need to determine whether your sector has a greater problem than sector B?

**Answer: It appears that kids are the primary at risk group. You need to know the age distribution of sectors A and B so that you can either compare the “kid-specific” prevalence and incidence (per answer to number 1) or adjust for differences in the percentage of kids in the two sectors' populations. 0.5 pt**

3. Further investigation reveals that the rash/headache illness in sectors A and B last for only three days. What feature of your data would you evaluate to determine the current state of this epidemics progression?

**Answer: Incidence versus time. 0.5 pt**

*Insight: Incidence is always the best measure of the populations risk for developing the disease. However, with an acute illness that has no background prevalence in the absence of a specific exposure, prevalence can be a good indicator of epidemics progression, too.*

4. Based on your team’s astute observations, you recognize that kids who drink from playground fountains show an increased incidence of the illness. You develop a test to identify kids who are at risk because of this activity. From national surveys or childhood behavior, you learn that for kids who always drink water at playgrounds, 80% will visit the playground  $\geq 3$  days a week; and for kids who never drink playground water, 80% never go to the park more than 3 days a week.

Design the highest quality test for detecting kids exposed to playground water based on asking kids (or their parents) how many days they go to the playground each week.

Determine the following properties of your test:

- Sensitivity
- Specificity
- Positive Predictive Value
- Negative Predictive Value

Answer: The percentage data define the distributions of kids who drink and kids who don’t drink in terms of the test parameter, days at the playground.

**Test: Ascertain from parents (kids) if their kids spend  $\geq 3$  days at the park or  $< 3$  days at the park.**

**If  $\geq 3$  days, call them drinkers**

**If  $< 3$  days, call them non-drinkers**

**2 pts**

Goal of test is to detect drinkers, i.e., a positive test is calling a kid a drinker (whether they actually do or not)

Based on the distributions:

		<b>Actual Drinking Behavior</b>	
		<u>Drinks</u>	<u>Doesn’t Drink</u>
<b>Test result</b>	Drinker	TP = 0.80	FP = 0.20
	Non-Drinker	FN = 0.20	TN = .80

Therefore:

**Sensitivity =  $TP / (TP + FN) \times 100\% = 80\%$**

**Specificity =  $TN / (TN + FP) \times 100\% = 80\%$**

**PPV =  $TP / (TP + FP) \times 100\% = 80\%$**

**NPV =  $TN / (TN + FN) \times 100\% = 80\%$**

**1 pt; 0.25 pts each**

5. Using your test, you discover that of the kids affected in the two sectors, 15 tested as no exposure to playground water. The total number of kids in the two sectors is 1115. Based on a sample of 200, you estimate that unaffected kids are equally likely to have been exposed or not exposed to playground drinking water.

What is the relative risk for the illness with respect to drinking playground water?  
 What is the odds ratio?

**Answer:**

Considerations:

Only 95% of the cases are kids.

Therefore, the total number of affected kids = 0.95 x 100 total cases = 95

**15** of these **affected** were **non-exposed**.

It follows that 95-15 = **80** of the **affected** kids were **exposed**.

Based on the rate of exposure in the sample of 200 (Let's hope it is a good random sample!), we can estimate that the number of **exposed** kids among the **unaffected** ~ 0.5 x (1115-95) = **510**

It follows that the number of **unexposed** among the **unaffected** also ~ **510**

**2 x 2 Contingency table:**

	Affected	Not Affected
Exposed	80	510
Not Exposed	15	510

$$RR = [80/(80 + 510)] / [15/(15 + 510)] = 4.7 \quad \boxed{3 \text{ pts}}$$

$$OR = (80 \times 510) / ((15 \times 510)) = 5.3 \text{ (over-estimates the risk of exposure)} \quad \boxed{0.5 \text{ pt}}$$

6. How much of the risk of skin rash/headache in kids in these communities can be attributed to their drinking playground water?

**Answer:**

Consideration: About 50% of the at risk population is exposed to playground water (drinks it)

$$\text{Attributable risk} = [0.5 (4.7 - 1)] / \{[0.5 (4.7 - 1)] + 1\} = 0.65, 65\% \quad \boxed{0.5 \text{ pt}}$$

Given that this is a "new illness," how do you explain the fraction that is not attributable to drinking playground water?

**Answer:**  $\boxed{1.5 \text{ pts; } 0.5 \text{ pt each}}$

**Errors in testing. You missed people who actually were exposed.**

**Other risk factors that you are now going to have to consider searching for.**

**Depending on the incidence determinations, pre-existing illness due to other causes.**