Elementary Epidemiology:
Disease Investigation Using Basic Math Skills

By: Sara Howard, MPH

Grade: 7th Grade

Time Required: 1 hour is suggested. However, the educator should adjust the time and number of activities to best fit the class needs.

Standards *(based on Tennessee Health Education)*:

- **Standard 9:** The student will understand attitudes and behaviors for preventing and controlling disease.
  - **9.1:** identify and define common pathogens
  - **9.2:** describe ways pathogens and diseases are spread, prevented, and managed;
  - **9.3:** describe signs, symptoms, and risk factors related to communicable and non-communicable diseases

Standards *(based on Tennessee 7th Grade Mathematics)*:

- **7.RP.A.3**: Use proportional relationships to solve multi-step ratio and percent problems. Examples: simple interest, tax, markups and markdowns, gratuities and commissions, fees, percent increase and decrease, percent error.
- **7.EE.A.2**: Understand that rewriting an expression in different forms in a contextual problem can provide multiple ways of interpreting the problem and how the quantities in it are related.

Lesson Objectives:

The student will:

- Understand the basic principles of epidemiology and how the principles related to skills they already possess
- Conduct a simulated epidemiologic investigation by integrating epidemiologic principles and their skills

Central Focus:

The main focus of this lesson is to allow students to apply their skills to a real world problem in epidemiology. After an introduction to the science of epidemiology and an introduction to new

terms, students will work through a disease investigation. Using example data, groups of students (or individual students) will analyze a set of data to determine the cause of an infectious disease outbreak. If time permits, the students will also determine how best to communicate their results to others.

Materials:

- *Elementary Epidemiology: Disease Investigation Using Basic Math Skills* 7th grade Presentation
- Disease Investigation Data – One set per group
- Group Activity Worksheet 7th Grade
- Paper
- Pencil

Background Information:

Epidemiology combines science and mathematics to study the distribution of disease within a population and the factors that influence the disease. At its core, epidemiology uses basic math skills to determine the distribution and cause of diseases. This lesson will provide students with an introduction to the science while allowing them to apply their skills to a real world problem. To complete the activity, students will need to familiar with adding, elapsed time, fractions, reducing fractions, and proportional relationships. Additionally, the lesson will introduce the students to new terms that they will need to complete the activity.

Presentation:

Hook:

Who here likes detective or cop shows? Does anyone like the mystery books or movie?

Allow time for the students to respond. Then introduce the topic of epidemiology.

Well today we are going to learn how a specific type of science, called epidemiology, can be used to solve different mysteries of diseases such as the cause of a particular disease or factors that could make a disease worse or prevent a disease from occurring. In epidemiology, science and math are combined to answer questions about diseases and their factors.

Introduction:

Begin by introducing the lesson. Tell the students that they are going to be learning how the math skills they already know are used by epidemiologists every day to help them understand causes of diseases.

Slide 2:

Introduce epidemiology by stating the definition, which Gordis defines as "the study of how diseases are distributed in a population and the factors that influence or determine the distribution." In this definition, population means a group of people. This can be a group of people of any size or characteristics you choice. However, this is discussed on a later slide. Distribution in this definition means how widely spread the disease is among the population.

Ask students:

> *Why would we want to study why diseases occur or how they are distributed?*
> *Why do we want to prevent or control the diseases?*

Tell students that epidemiologists are the scientists who work to understand diseases so they can prevent and control their impact on people. They are sometimes referred to as the "detectives of the medical field." The students will get to practice being these medical field detectives today.

Slide 3:

Lead into the next slide by asking:

> *What types of diseases do you think epidemiologists study?*

The answer is every kind of disease of course. Like all sciences, epidemiology has its own specialties. Although the list on the slide is not exhaustive, it shows the most common types of epidemiology. Please clarify each type of epidemiology for the students. Additional information is provided below.

> *Infectious Disease Epidemiology: Infectious diseases are diseases that are caused by other organisms such as bacteria and viruses. Typically, these are passed from person to person or animal/organism to person. Examples would include the flu, strep throat, pneumonia, and HIV.*
>
> *Chronic Disease Epidemiology: Chronic diseases are disease that last for an extended period (3 months at least), and they generally do not have a cure. Examples include cancer, diabetes, cardiovascular disease, HIV, Alzheimer's disease, and Parkinson's disease.*
>
> *Environmental Epidemiology: This looks at how the environment affects health. Certain diseases can be caused by exposure to things in the environment. This can include*

second hand smoking, chemical exposures, air quality, or waterborne illnesses. Examples of diseases would be lung cancer, asthma, and anthrax.

Cancer Epidemiology: Cancer epidemiologists study the causes of different types of cancer and what influences them. For example, the chemicals in cigarette smoke increases the risk for different types of cancer including lung cancer.

Genetic Epidemiology: Genetic epidemiologists study how diseases that are inherited (passed down to family members) and how a person's genes influence disease. Examples of genetic diseases would be Down's syndrome, Muscular Dystrophy, and Cystic Fibrosis.

Occupational Epidemiology: Occupational epidemiologist study diseases that affect groups of workers. Occupations, and even different jobs within those occupations, have different health risks. In this field, epidemiologist work to understand, control, and prevent the diseases from occurring. For examples, occupational epidemiologists may study the effect of pesticide exposure on farmers or how stress levels affect teachers' health.

When reviewing the types of epidemiology, you may want to ask the students if they can provide examples to the different types of diseases that would be covered in each category. For example, asking if they can name a chronic disease may help them stay engaged with the material.

Lead into the next slide by asking:

What skills do you think will be required for this job?

Slide 4:

After you allow a few minutes for the students to answer, go to the next slide. This infographic shows some of the groups of skills need for epidemiology. Explain what each area is and why they are important.

Research Design & Analysis: Epidemiologists need to know how to design a study so that it is ethical and scientifically meaningful. It should also be defensible, meaning that they should be able defend against scientific scrutiny.

Medical, Biological, and Mathematical Knowledge: Epidemiologists must have knowledge of medicine, biology, and math. This helps them design their studies, collect their data, and understand the disease processes.

Data Analysis, Presentation, and Communication: Epidemiologist must know different techniques to analyze data both by hand and by software. Then, they need to know how

to communicate the result effectively, both visually and orally. Visually here means graphs and figures as well as written documents. Orally in this sense means public speaking as well as one-on-one conversations with people.

Public Policy and Cultural Customs: Epidemiologists need to know the laws that govern the population they are studying as it may influence their health outcomes. For example, different countries have different healthcare laws. These can influence the health of a population. Epidemiologists also need to know cultural customs because culture will also influence health. For example, the southeastern part of the United States has culture that can be very different from the northeastern part. A strategy to improve health that works well in the northeastern US may not work as well in the southeastern US.

After explaining the infographic, ask students:

What subjects do you study that are used by epidemiologists?

Slide 5:

Now that you have a few basic principles of epidemiology we are going learn some terms that epidemiologists use every day. These terms will be the basis for the group activity. Review the terms with the students and provide examples for each definition. You can use any examples that fit the definitions, but examples are provided below.

Example Population: If we wanted to study everyone in 7th grade at this school, who would be in our population?

- *Answer: All the students in seventh grade at this school. We would not include any other 7th grade students at a different school, and we would not include students at this school who are not in 7th grade.*

Example Sample: Sometimes it is not possible to study everyone in a population. For example, it would be difficult to study everyone in the United States (not enough time & not enough money), so epidemiologists will use a sample of the population. If our study population is the students at this school, we may decide to study the people in this classroom as our sample. This means that only the people in the classroom that will be included in the study. If someone is in the hallway, they are no included in the sample.

Example Case & Control: A case can be anyone within your study group (sample or population) who has the disease you are studying. A control is anyone else left in the sample. For example, if our sample is this classroom, we decide that the "disease" we are studying is people wearing classes. Our cases would be people in this classroom who

wear classes, and our controls would be people in this classroom who do not wear glasses. Raise your hand if you are case. Raise your hand if you are a control.

Slide 6:

Use this slide to provide students with another example of population, sample, case, and control. Walk them through the terms using the images in the PowerPoint.

If the population is the entire picture, what would be an example of a sample?

- *People in the left half, right half, top half, or bottom half of the photo.*
- *People in color or people in black & white*

If we make our sample the people on the left side of the page, what could our cases and control be?

- *Cases could be the people in color*
- *Controls could be the people in black and white*

If the students understand the concepts of cases, controls, population, and sample well from the slide 5, then slide 6 may not be needed. However, these concepts are important to their understanding of epidemiology and the activity.

When you click 'enter' on slide, the division line for the sample will appear. Additionally, another click will reveal the cases if the left side of the image is the sample.

Slide 7:

This slide introduces the scenario for the epidemiologic investigation. *Tip: The name of the school in the scenario can be changed to fit your school. If you change the school name, please change slide 11 to match.* Review this slide with the students. This is the basis for the activity.

Slide 8:

This slide has examples of a few things they need to find out during their investigation.

Ask the students:

Based off what you have learned, what are somethings you need to figure out?

- *Number of cases*
- *Number of controls*
- *Who ate the school lunches on Monday*
- *What foods were served on Monday*

- *When did people become sick*
- *What were their symptoms*

Click the space bar the answers to appear on the slide.

Slide 9:

To answer these questions, we need a little more information. We need to define what makes a case, so we need to know the symptoms the *salmonella* and the incubation period. Introduce the term incubation period and its definition.

Incubation period is a new term. Does anyone have any idea what it might be?

The incubation period is the time period from when a person is exposed to an organism to when symptoms are first developed. In this case, the incubation period begins when a person ingested the organization (ate school lunch) to when the person first developed symptoms.

So why do care about the incubation period?

- *Different organizations have different incubation periods, so this can help epidemiologists determine if a person has the illness or not.*

Once you introduce incubation period, reveal the symptoms and incubation period for *salmonella*.

For the activity, the students will need to find people who are sick within the correct incubation period (Tuesday, Wednesday, or Thursday) and who have at least 3 of the symptoms listed.

Slide 10:

Epidemiologists use a specific kind of math, called statistics, to help them understand and find the cause of diseases. We are going to use some of those same math skills to help us today. Click to show the statistics they will be using. Go through each of the statistics the students will calculate. When reviewing the terms, emphasize how they can calculate each term. For example, to find the incident rate, the students only need the number of cases and the number of controls.

Note: While the attack rate and incident rate are different statistics, the strategy to finding them is the same. In the attack rate, the 'population' becomes the number of students who ate a specific food. Therefore to find the attack rate the students need to find the number of cases and controls in the new population.

The students will have to use the information they have learned about fractions to create the attack rate and the incident rate. If the students can, they should try to reduce the fractions.

Slide 11:

This slide shows the next steps for the students. Review the information and divide the students into teams. Ideally, the teams will be 3-4 people.

Note: Since each team will receive one set of data, it can be useful to print the data on colored paper (each team gets their own color). This allows the teacher to easily identify that each group received only one set of data.

Slide 12:

Slide 12 shows the main analyses of the activity. Based on the information provided in the presentation, the students should be able develop a strategy to calculate the analyses.

Some classes may need more guidance on where to begin with the investigation. Typically, it is best to start broad with the analysis and work towards more specific information. Thus, students should begin by dividing the data into cases and controls, so they can calculate the incident rate. When the students are dividing the data into cases and controls, they may notice that one food is more common in the cases than others or that a food is more often in the controls. This observation can help guide them when they begin calculating the attack rates.

Slide 13:

This slide marks the difference between the introductory material and the bonus information. The bonus information mainly meets the 7^{th} grade mathematic standards. The bonus material allows the students to see how some of the principles they learn as a part of their math education are used in real world problems.

Slide 14:

Slide 14 introduces the concept of generalizability. Generalizability is the ability to make inferences about a population based on the results of a sample. Basically, generalizability allows the results to be extended to the larger population. For this to occur, the results must be both internal and externally validity.

Internal validity assesses reliability, or repeatability, of the assessments. It answers the question 'Did we do things right?' In other words, were the methods correct? Could this method be repeated and generate the same (or similar) results? Internal validity allows scientists to determine if the observed outcomes are results of chance or of an actual association.

External validity determines if the results can be extended to a larger or different population. External validity is met if the sample is representative of the population. It answers the question 'did we do things the right way?'

Together internal and external validity are required for the results to generalized to a population. If the study methods were flawed, the study will not be internally valid. Therefore, the result cannot be externally valid. If the methods are repeatable and valid (internal validity) but the study has different characteristics than the population, the results are not externally valid. Therefore, they cannot be extended to the population.

Slide 15:

A study must be both internally and externally valid for the results to be generalized.

If the study is generalizable, what does that mean? The results in the sample should be proportionate to the results in the population.

Why does that matter? When results are generalizable, we can make inferences about the total population. This can be used to help identify outbreaks, direct policy, initial interventions strategies, and dictate future research.

Slide 16:

This slide is a transition to the word problem. It reiterates the sentiment that sample results can allow scientists to be estimate the effect in the population. The second bullet introduces the idea of using the sample population incident rate to estimate the number of expected cases in the sample population.

Slide 17:

The first bullet is an example word problem.

> *"If the sample has 73 cases and 77 controls, how can we find the expected number of cases in a population of 356, 897?"*
>
> *What information in the problem is needed to answer question?*

Click and the important information will be underlined. The next step is to set up the proportion. The proportion will be the incident rate in the sample equal to the incident rate in the population.

To find the sample incident rate, the numerator will be the number of cases (73), and the denominator will be the total number in the sample. This can be found by adding the number of cases and controls (73 +77 = 150).

Since the number of cases is missing for the population, an 'x' is used in the numerator. The denominator is the total in the population (356,897). To find the number of cases, solve of x.

The answer will be 173,689.8733. Since you cannot have half a person, the number should be rounded up to the next whole number. Therefore, the number of cases expected in the population is 173,690.

Clicking in PowerPoint will reveal the information step by step.

Slide 18:

This slide transitions to disease surveillance and background levels of disease occurrence. Epidemiologist monitor levels of disease within a population to help them identify any changes. This is called disease surveillance. Disease surveillance can take place at a federal, state, or local level, depending on the disease. Each population has its own level of disease that normally occurs. This is known as the background level of disease. A spike in disease occurrence can signify an outbreak, and a decline can indicate the effectiveness of an intervention. *Note: both can signal other things such as changes in screening programs or monitoring data. However, this is outside the scope of this lesson.*

Slide 19:

Slide 20 allows students to figure out if the results exceed the background level of disease occurrence. To determine if expected number of cases in the population is greater than the background level, students need to find the background level. To do this, they must multiply the background level range (7%-15%) by the total population (356,897). The students can then compare the numbers to the expected number of cases.

Clicking on the PowerPoint will walk the students step by step through the process.

Disease Investigation Activity:

Instructions:

Before beginning the activity, the students should be familiar with the basic principles of epidemiology that are provided in the presentation. For the activity, the students should be divided into groups, ideally with 3-4 students each. Each group will be given a set of the data, which can be found on the ORISE website, to analyze.

The data are a group of 50 forms that mimic some of the information gathered during a disease investigation. The forms contain name, gender, symptoms, foods eaten at lunch on Monday, and first day symptoms appeared. Some of the items (name and gender) will not be needed by students; however, students will decide information is useful and what information is extraneous.

> *Note: The names of the students in the data are all famous scientists. Although this information is not used in the analysis, it may be worth pointing out some of the scientists. An explanation of each scientist is provided in the materials on the ORISE website.*

A worksheet is available for the group activity to help the students work through the disease investigation process. However, this worksheet is not required for the activity. It is merely a tool to provide structure to critical thinking needed for this activity. If not using the worksheet, the students should record the process they used to solve the disease outbreak scenario on another piece of paper. The students should record the number of cases, number of controls, total population, the incident rate, the attack rate for each food, and the mean incubation period. If the lesson includes the bonus information on generalizability, the worksheet is required for the activity.

After the students receive their data, the students should begin sorting their data. Since they need to know how many people are in the sample as well as how many people are cases and controls, the students may want to begin by sorting the data into cases and controls (cases + controls = sample). They do not necessarily have to start with this step, although it will be more time efficient. Some students may want to begin by counting the total number in the sample. The students should arrive at the same answer either way.

Others may want to start by sorting the data by foods consumed. For the activity, please try to persuade them from this. Typically, it is better to begin broad and work towards more specific analysis. In this case, the calculating the incident rate is the broad starting point. To do this, the students need the number of cases and controls in the data. When they are sorting the data, the students may make observations regarding which foods are common in the cases and controls. This observation can be used into inform their starting point for the attack rates.

In this lesson, a case is defined as someone who at the school lunch on Monday, has at least 3 of the symptoms of *salmonella*, and became sick 24-72 hours after lunch on Monday. A control is defined as someone who at the school lunch on Monday but did not get sick within the correct timeframe or did not have at least 3 symptoms.

After the students have found the number of cases, controls, and the number in the sample, they can calculate the incident rate. If possible, the students can reduce the fraction. Once the

students have calculated the incident rate, they may have made some observations about the data. The students should, then, make a prediction of which food they think could be the culprit.

The students will then calculate the attack rates. Since the attack rates will be calculated by food item, the student may want to start with the food they think is responsible for the outbreak. Regardless of where they start, each student should calculate the attack rate for each food. If possible, they can reduce the fractions.

After the attack rates are calculated, the students should be able to identify which food is responsible for the outbreak. They should also be able to explain why they have selected that food.

The students will then be asked to graph the results of the data analysis. They students have many options for graphs, which include the following:

- Bar Graph – By Attack Rate
- Line Graph – By Number of Cases on Each Day
- Bar Graph – By Number of Cases on Each Day

The types of graphs listed above are certainly not an exhaustive list. The students can choose what type of graph or graphs they would like to use.

Once the students have graphed the results, they need to decide how best to communicate the results of the investigation. Options could include a written report, a presentation to the community, phone calls to all the parents, press release to the news outlets, and social media posts.

The three bonus questions allow the students to work through proportional problems. The step-by-step answers are provided in the answer key.

Formulas to Remember:

Incident Rate: number of cases / the total population

Attack Rate: number of people who ate a specific food and became ill / number of people who at the specific food

Mean incubation period: sum of the number of hours of the incubation period for each case / total number of cases

Closing:

To close the lesson, ask each group to present their findings to the class. The groups should cover which food they think is responsible for the outbreak and how they arrived at that conclusion. If time allows, the students should also present how they graphed the results and how they planned to communicate the results to the community.

Assessment:

To assess the level of understanding among the students during the activity, the teacher could walk and ask questions as they work. Since groups will work at varying paces, the teacher can gage how far along in the activity the students are and address the questions accordingly. For groups that are struggling, the teacher should provide aid to help them better understand.

In the initial walk around, the teacher should check for the understanding of cases and controls. Since the rest of the activity is involves builds on the understanding of cases and controls, the students must understand these concepts to successfully proceed.

For a more formal assessment of understanding, the worksheet is used. If the worksheet is not used for the activity, each group should turn in a sheet of paper documenting their methods for investigating the disease outbreak and the results of the investigation. This should include the statistics they calculated and an explanation as to why they came to their conclusion. An answer key to the worksheet is provided. If the worksheet is not used in the activity, the teacher can still find the answers to the number of cases and controls, the incident rate, and the attack rates by food in the worksheet answer key.