Module 2

The investigations in Module 2 of the *Detectives in the Classroom* curriculum prepare students to answer the second of five Essential Questions:

> Is there an association between the hypothesized cause and the disease?

When students understand how to answer this question, they will be developing the second of five Enduring Understandings that provide the structural framework for the curriculum. The second Enduring Understanding is:

> Causal hypotheses can be tested by observing exposures and diseases of people as they go about their daily lives. Information from these observational studies can be used to make and compare risks and identify associations.

By completing the Module 1 investigations, students learned how health-related conditions and behaviors are not distributed uniformly in a population, that each has a unique descriptive epidemiology that can be discovered by identifying how it is distributed in a population in terms of person, place, and time. This descriptive epidemiology provides clues for formulating hypotheses. Students realized that descriptive epidemiology is only the first step and that formulating hypotheses is not the same as proving hypotheses.

The Module 2 investigations develop students' hypothesis-testing skills. These hypothesis-testing skills are called analytical epidemiology. While descriptive epidemiology looks at the way a disease is distributed in a group of people in order to formulate a hypothesis, analytical epidemiology looks at the way an exposure and a disease are distributed in a group of people in order to test a hypothesis.

One tool that epidemiologists use to study how an exposure and a disease are distributed is a 2x2 table. A 2x2 table has two columns, one for the people who have the disease and the other for the people who do not have the disease and two rows, one for the people who were exposed to the hypothesized cause of the disease and the other for the people who were not exposed.

	Disease	No Disease	
Exposure	a	b	
No Exposure	c	d	

The 2 rows and 2 columns cross each other and form 4 cells. These cells are labeled a, b, c, and d as shown above.

For any combination of an exposure and a disease, for example, using a cell phone and brain cancer, every person belongs in one of the four cells:

Cell	Exposure (Uses a Cell Phone)	Disease (Has Brain Cancer)
a	Yes	Yes
b	Yes	No
c	No	Yes
d	No	No

In **Investigation 2-1: The 2 x 2 Table,** students examine the results of a hypothetical study to determine if a medication prevents acne. In doing so, students uncover that a scientific study needs to examine not only people who are exposed to the medication (cells a and b of the 2x2 table), but also people who are not exposed to the medication (cells c and d), called a control group. Students construct 2x2 tables and realize that when they study how an exposure and a disease are distributed, every person in the study population fits into one of the four cells of the 2x2 table. Students also realize that the way an exposure and a disease are distributed in a group of people can be expressed mathematically.

In **Investigation 2-2: Compared to What?,** students continue to examine the need for a control group and have more experience in using a 2x2 table. First, students learn how risks are calculated from the 2x2 table. They will calculate two risks: 1) the risk among people who are exposed to the hypothesized cause (a / a + b), and 2) the risk among people who are not exposed to the hypothesized cause (c / c + d), the control group.

Students then compare the risks of disease among the exposed and unexposed by dividing one risk by the other risk. Because one risk is being compared to another, this calculation is called a relative risk. From the relative risk, inferences can be made based on the similarities or differences between the two risks. Epidemiologists calculate risks, compare risks, and based on the degree of similarity or difference between the risks, make inferences about whether or not an exposure and a disease are associated with each other.

Three examples describe the range of possible inferences:

1. If the risk of getting the disease among the people who were exposed is 5 times as likely as the risk of getting the disease among the people who were not exposed, a strong association between exposure and disease has been found and students infer that the exposure may cause the disease.
2. Conversely, if the risk of getting the disease among the people who were exposed is 1/5 times as likely as the risk of getting the disease among the people who were not exposed, a strong association between exposure and disease has been found and students infer that the exposure may prevent the disease.
3. Finally, if the risk of getting the disease among the people who were exposed is the same or very similar to the risk of getting the disease among the people who were not exposed, no association has been found between exposure and disease and the students may infer that the exposure is unrelated to the disease.

This is summarized in the table below:

Difference (Relative Risk)	Possible Inference
5 (Five times as likely)	Exposure and disease are associated because exposed people appear more likely to have the disease.
5 (1/5 times as likely)	Exposure and disease are associated because exposed people appear to be protected from disease.
1 (No difference)	Exposure and disease are not associated.

Students will realize that it is only when the risks of getting disease in exposed and unexposed groups are compared that a hypothesis can be tested. The same risk of getting a disease in the exposed group (for example 10%) may lead to different inferences depending on the risk of disease in the unexposed group (10%, 50%, or 25%).

This investigation concludes with the observation that CDC is not only an acronym for the Centers for Disease Control and Prevention, but also stands for what epidemiologists do: count, divide, and compare.

The remaining Module 2 investigations provide opportunities for students to see the challenges and issues involved in designing and conducting different analytical epidemiological studies.

When an exposure is hypothesized to prevent disease, it may be ethical to test the hypothesis in a trial. For example, trials have been done to see whether vitamins reduce the risk for cancer, whether aspirin lowers risk of heart attack, and whether a vaccine is effective in preventing an infectious disease.

In **Investigation 2-3: What's Wrong with This Picture?,** students uncover the circumstances under which it is ethical to conduct human experiments (trials). The investigation begins by asking students to list principles that would assure that human experiments are done ethically. Next, students learn about the history of the development of ethical standards for human experimentation by reading articles about human rights abuses in human experimentation (Nazi medical experiments and the Tuskegee Study). Based on the readings, students develop a set of principles for conducting human experiments, compare their principles to the Belmont Report, and learn to distinguish between the ethical and unethical conduct of human experiments. Students also learn about informed consent, including the main elements of consent and when it is needed.

Human experiments may be appropriate for studying a hypothesis about an exposure that is hypothesized to prevent a disease. However, to test hypotheses about exposures that might cause a disease, epidemiologists rely on observational studies or “natural experiments.” These are studies of free-living people and the things they are exposed to and the diseases they contract as they go about their daily lives. Unlike a trial, where people are intentionally exposed to something, epidemiologists do not actually perform “natural experiments.” People perform these experiments on themselves. They are going on regardless of the epidemiologist’s efforts. The epidemiologist takes advantage of these opportunities by creating a research design/observational study to examine what is happening anyway.

Observational studies are therefore the only way we can test hypotheses in humans about possible disease-causing exposures because it would be unethical to intentionally expose human study subjects to something that may cause harm. For example, to test the hypothesis that cigarette smoking causes lung cancer, only observational studies were done. Researchers did not randomly assign people to experimental and control groups and ask the experimental group to smoke a pack of cigarettes a day for 20 years. Rather, people decided for themselves whether or not to smoke cigarettes and, in doing so, put themselves into either the exposed or unexposed group. The epidemiologist then calculated the risk of getting lung cancer among the people who smoked and compared it to the risk of getting lung cancer among the people who did not smoke. No one was forced to do anything.

In **Investigation 2-4: Backpacks and Back Pain,** students test the hypothesis that carrying a heavy backpack causes back pain. Using a 2x2 table, students perform an in-class observational study by collecting and analyzing data from the "natural experiment" that they perform on themselves by carrying their backpacks. This first-hand experience in conducting an observational study helps students understand procedures involved in creating definitions of an exposure and an outcome, displaying data in a 2x2 table, calculating risks, comparing risks (calculating relative risk), and making inferences.

In **Investigation 2-5: Observational Studies,** students work in their Epi Teams to select their own hypothesis to test, and then design and conduct an in-class observational study. Students create questions that will allow them to count and classify their "exposure" and "outcome," assure that participants have given their informed consent, collect data, display data in a 2x2 table, calculate risks and relative risks, and make inferences. Epi Teams conclude by planning a presentation, according to the IMRAD format (Introduction, Methods, Results, And Discussion), and presenting to the class.

A hypothesis can be tested using different epidemiologic study designs. A challenge to epidemiologists is to decide which design is best for a given set of circumstances.

In ***Investigation* 2-6: The Journey,** students explore the four basic analytical epidemiological study designs including the fundamental experimental design used in a controlled trial and the three designs used to observe "natural experiments": the cohort study, case-control study, and cross-sectional study. A "train analogy" is used to depict the journey between exposure and disease and the different ways exposure and outcome data are collected during the journey. The train analogy helps students uncover how the study designs differ according to when during the journey the epidemiologist determines study participants' exposure and disease status.

The train analogy for the four designs is described below:

Experimental Study Design	**Trial**: The epidemiologist is on the train during the entire journey and "assigns" passengers to either receive or not receive a certain "exposure." The exposure is carefully administered to those who were assigned to receive it and is not given to the others. The researcher keeps checking the exposure status of the participants and whether or not each person has developed the disease of interest.
Observational Study Designs	**Cohort Study:** Just as in the trial, the epidemiologist is also on the train during the entire journey. But there is an important difference. The epidemiologist is not assigning passengers to an exposed or not exposed group, but rather, is just observing them and counting. Passengers are not being told to have, or not have, an exposure. They are just living their normal lives as passengers. The epidemiologist, on the train for the whole journey, just keeps observing the passengers' exposures and whether or not they develop the disease during the journey.
	Case-Control Study: This design is much different than the first two study designs because the epidemiologist is not on the train. Rather, he/she is waiting at the train station at the end of the journey. As passengers get off the train, the epidemiologist selects all the passengers who developed the disease during the journey (cases), and selects a few other passengers who are similar but did not develop the disease during the journey (controls). The epidemiologist then asks each person in the case group and control group questions about their exposures during the train ride. The epidemiologist has to rely on passengers' memories of exposures. This design differs from the trial and the cohort design because, instead of starting with healthy people, collecting data about exposure, and waiting to see if they get sick, it starts with groups of people with and without a disease and looks back to try to collect exposure data to figure out what made them sick.
	Cross-Sectional Study: The epidemiologist, who is not on the train during the journey, stops the train somewhere during the trip (kind of like a train robbery). The epidemiologist gets on board and takes a "snapshot" by asking everyone two kinds of questions about how things are at the same time: 1) about disease status today, and 2) about exposures encountered today. Then the epidemiologist leaves the train to analyze the data gathered for that particular day. Epidemiologists call this design a cross-sectional study because data are collected about exposure and disease during a "slice" in time.

To summarize, the timing of three events distinguishes each of these designs:

- When during the journey does the epidemiologist become involved?
- When during the journey is exposure status determined?
- When during the journey is disease status determined?

In **Investigation 2-7: Epi Team Challenge,** students practice distinguishing between the four study designs by participating in an Epi Team challenge. Epi Teams are given clues about a study design and asked to hold up cards to show which design is being described.

In **Investigation 2-8: Which Design Is Best?,** students become more familiar with the epidemiological study designs, uncover the strengths and weaknesses of each, and realize the circumstances under which each design is "best." In order to do so, students rank order the four basic epidemiological study designs according to different characteristics, such as the speed with which a study design can be implemented, its expense, and the accuracy of the data.

For example, we are not as certain about associations that are identified in "natural experiments" as we are of the results of controlled trials. This is because in observational studies of natural experiments, exposure and disease status may not be able to be measured as accurately, and because the people who are exposed to something may be different in other ways from the people who are not exposed.

Given these challenges, the epidemiologist must carefully select a sample of people whose “natural experiments” will be studied, apply the appropriate research design to this sample, measure exposure and disease status accurately, recognize the limitations of these investigations, and make inferences cautiously.

The table below outlines some of the strengths and limitations of the four designs:

Study Designs: Strengths and Limitations		
Study Design	Main Strengths	Main Limitations
Trial	• Closest to laboratory experiment • Has control over exposure dose	• Unethical to do human experiment with possible harmful exposures
Cohort Study	• Starts with healthy people, time order of exposure and disease is evident • Accurate exposure data	• Can take a long time to get answer • Can cost a lot of money
Case-Control Study	• Can get answers quickly • Is less expensive	• Exposures might be forgotten • Exposures might be remembered differently by people with and without disease
Cross-Sectional Study	• Can get answers quickly • Is less expensive	• Unknown time order of exposure vs. outcome

In **Investigation 2-9: Designs, Diagrams, and Tables,** students identify, for each of the four basic epidemiological study designs, where data from a study design flow diagram "fit" into a 2x2 table. Students appreciate how, for any study design, the 2x2 table is a useful tool for:

- Classifying disease occurrence according to exposure status
- Calculating risks of disease among exposed and unexposed
- Calculating the relative risk
- Making inferences

In **Investigation 2-10: Concept Connections,** students identify the important concepts that need to be understood to answer the second Essential Question: "Is there an association between the hypothesized cause and the disease?" Each Epi Team then creates a Concept Map that depicts and explains how the concepts connect to each other. At the conclusion of this investigation, students will realize that they have developed the second Enduring Understanding of *Detectives in the Classroom*: “Causal hypotheses can be tested by observing exposures and diseases of people as they go about their daily lives. Information from these observational studies can be used to make and compare risks and identify associations.”