CHAPTER 3

Research and Global Health

Population health research provides essential information about the prevalence of diseases around the world, risk factors for those diseases, and the effectiveness of interventions. A basic knowledge of the methods used to collect, analyze, and synthesize global health data allows anyone to read and understand a vast array of resources for evidence-based global health practice and policy.

3.1 THE IMPORTANCE OF GLOBAL HEALTH RESEARCH

Health research at its broadest encompasses everything from molecular and cellular biology to clinical research to population health. Most global health research focuses on the public health end of the spectrum. The goals of public health research include identifying and classifying new health problems, determining risk factors for disease, developing and testing new interventions for preventing or treating illness, evaluating the impact of health policies on health outcomes, and synthesizing existing knowledge.

Epidemiology is the study of the distribution and determinants of morbidity, mortality, and disability in populations. Epidemiologists and other public health researchers collect and disseminate data about the health-related conditions that occur in particular populations and the characteristics of the people who are most at risk for developing those conditions. This information helps clinicians to diagnose illnesses, prescribe appropriate therapies, and encourage healthy lifestyles for their patients. It also helps communities to set their own public health priorities and design and evaluate programs to address these issues, especially when a process called **community-based participatory research (CBPR)** is used. Research reports also inform the development of evidence-based policies and programs.

Several significant global health reports are released by major international organizations each month, and thousands of new global health-related

academic and professional journal articles are published each month. A basic understanding of health research methods makes all of these resources—including the more than 500 references cited in this book—accessible to those seeking additional information about a particular health topic.

3.2 THE RESEARCH PROCESS

Health research follows a fixed set of steps (**Figure 3–1**).[1] Researchers start by identifying a focused study question and selecting an appropriate study design. They then work out the logistics of the study and collect data, which could take the form of interviewing people, running laboratory tests, acquiring documents for review, or other methods. After analyzing the collected data, the findings are disseminated through oral presentations and written publications.

The basic unit of population health research is the **primary study** that collects new data from individuals drawn from a well-defined population, such as the students at one school or a sample of residents of one suburban area. Most primary studies are observational and simply ask participants to complete a questionnaire. Some primary studies are experimental ones in which researchers assign at least some of the participants to do something new, perhaps to start taking a daily multivitamin or to take a new drug for their health condition. When the results of primary studies are published, they are said to add to the literature on a particular topic. (A **secondary study** also contributes to the body of knowledge on a topic by analyzing and reporting on existing data that someone else collected.) The summaries of these articles are often indexed in abstract databases such as PubMed that allow the contents of the manuscripts to be searched.

For global health, it is often important to have a worldwide perspective on disease incidence and prevalence rates, risk factors for disease, and other health information. A **tertiary study** seeks to identify all the primary (and secondary) studies that have been published on a particular topic and to summarize what those studies say. These systematic reviews and

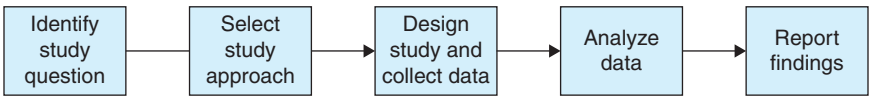

Figure 3–1 The research process.

Source: Reproduced from Jacobsen KH. *Introduction to health research methods: a practical guide*. Sudbury MA: Jones & Bartlett; 2012.

meta-analyses provide a comprehensive picture of what is known about a particular issue, and the findings and estimates from these studies can be used to predict health status in populations for which no data are currently available and to forecast future situations. Most of the global health reports published by the World Health Organization and other international agencies are based on primary studies (including country-level surveillance reports) and on meta-analyses that synthesize primary studies. These reports provide comparable data from around the world that can be used by policymakers, public health professionals, and others to foster improved health in their communities and countries.

3.3 OBSERVATIONAL STUDY DESIGNS

Most population-based public health research uses an observational study design. An **observational study** simply observes what people are doing or asks about what they have done in the past. No intervention is assigned to participants. The goal is to learn about a population as it is. **Descriptive studies** seek to describe the members of a population, the prevalence of risk factors within that population, or the rate of disease within that population. Descriptive studies often seek to answer questions about person (who?), place (where?), and time (when?). **Analytic studies** aim to understand the associations between risk factors and disease within a population and to answer "why?" questions.

3.3.A Prevalence Surveys

A **prevalence survey**, also called a **cross-sectional survey**, can be used to get a snapshot of a population's health status at one point in time. The research plan is fairly simple: recruit a representative sample of the population the researcher wants to know about, ask the participants a series of questions, and then analyze the collected data to see what proportion of the population reported various characteristics. The questionnaire can cover a wide variety of topics, including demographics (such as the age, sex, household income, and educational level of the participant), risk behaviors and other risky exposures, and illnesses and disabilities. The survey instrument can also include "KAP" questions about *k*nowledge, *a*ttitudes/beliefs, and *p*ractices/behaviors.

Prevalence surveys are one of the most common study designs used in public health research. They are often used as part of community needs

assessments, and they are also used for conducting program evaluations. They are especially useful when there are time and budget constraints, because data can be collected quickly and inexpensively.

There are two key cautions about conducting and critically assessing cross-sectional surveys. First, it is very important for prevalence studies to recruit participants who are truly **representative** of the population the researchers say they want to examine. For example, a study about the health of women in a community should not be limited to including only women who are currently pregnant, because that recruiting strategy would systematically exclude older women who are of post-reproductive age. And a study about the health of college students should not recruit only students who are members of sports teams at the school, because those students are likely to be fitter than the general student body. Second, no conclusions about causality can be made from cross-sectional data, because all the questions about exposures and diseases are asked at the same time. For example, a cross-sectional survey of chewing tobacco use and dental cavities among a group of 1000 high school students might find a significantly higher prevalence of cavities among people who use chew, but that would not prove that chew caused cavities nor would it prove that cavities cause people to chew. This type of survey cannot show whether the chew or the cavities happened first.

3.3.B Case Series

A **case series** looks at the characteristics of a group of people who all have the same disease (or all had the same exposure). (A **case study** is a description of one patient. A case series describes two or more patients.) Most case series studies are written by and for clinicians, and most summarize the information in the medical records of people who were treated at a particular hospital for a particular condition. The goal of a case series may be to understand the demographic and other characteristics of people with a particular disease, to describe an unusual presentation of a disease, or to clarify the typical progression of a disease. Because a case series does not include a comparison group of healthy people, it is not possible to examine risk factors for the disease.

3.3.C Case-Control Studies

Case-control studies recruit people with a disease (**cases**) and similar people who do not have that disease (**controls**) so that their past exposures can be compared. After confirming that a participant has the disease of

interest or does not have the disease, the participant is asked about his or her health behaviors (such as diet, physical activity, tobacco use, and alcohol use now and in the past), environmental exposures, and health history. After a sufficient number of cases and controls have completed the questionnaire, statistical analysis is used to identify the exposures that were reported more often by cases than by controls.

Case-control studies are ideal for learning about rare diseases. They can also be helpful for identifying past exposures that might increase the risk of disease, but the results of case-control studies must be interpreted cautiously because participants may have difficulty accurately recalling exposures that took place years or even decades before the study.

The most common way to look at the association between an exposure and a disease outcome is to create a **2 × 2 table** that has two rows for exposure status and two columns for disease status. Each individual in the study population is classified into one of the four groups created by the 2 × 2 table—exposed and diseased, exposed but not diseased, not exposed but diseased, and not exposed and not diseased. The count of the number of individuals in each of the four groups is filled into the cells of the 2 × 2 table, and various measures of association can then be calculated from those values.

The typical measure of association between an exposure and an outcome in a case-control study is the odds ratio. This is the same type of measurement used in betting (**Figure 3–2**). If someone thinks that a horse has a 25% chance of winning a race (and a 75% chance of losing), then the **odds** on the horse are 25:75, which can be simplified to 1:3 or 1/3 or 0.33—one chance of winning compared to 3 chances of losing. The **odds ratio (OR)** compares the odds of a case having a history of a particular exposure to the odds of a control having been exposed to the potential risk factor (**Figure 3–3**). An OR near 1 means that there was no association between the disease and the exposure in the study population, because cases and controls were equally likely to report the exposure. An OR greater than 1 indicates that people with the disease were more likely than people without disease to have a history of the exposure, which implies that the exposure was risky. An OR less than 1 indicates that cases were less likely than controls to have a history of the exposure, which implies that the exposure was protective.

Because a relatively small number of people sampled from a larger population cannot exactly describe the population as a whole, ORs and other statistical measures are often reported using confidence intervals. A **95% confidence interval** for an OR can be interpreted as saying "based on the sample of people the researchers took from the larger population, we can be 95% confident that the true OR in the

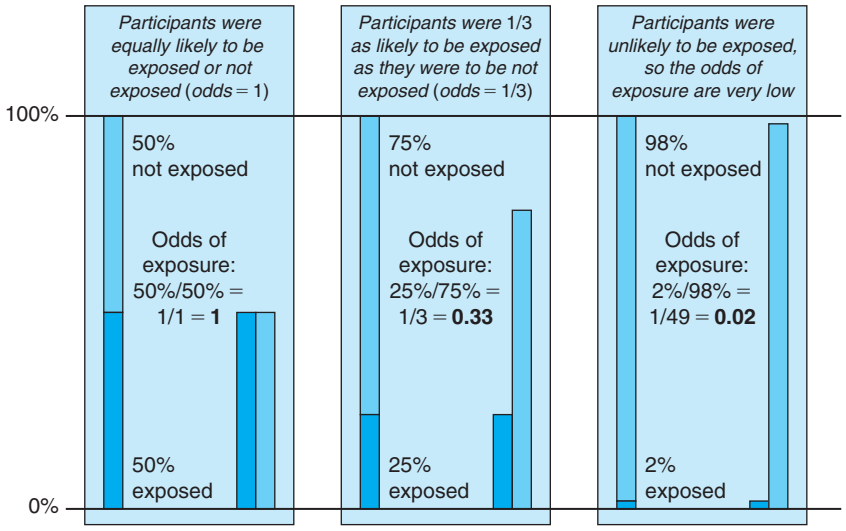

Figure 3–2 Odds.

Source: Reproduced from Jacobsen KH. *Introduction to health research methods: a practical guide.* Sudbury MA: Jones & Bartlett; 2012.

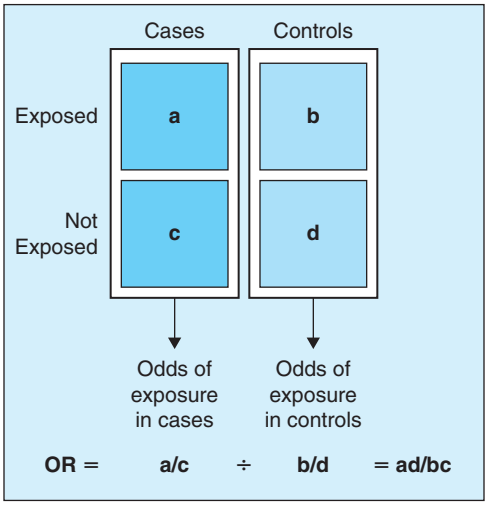

Figure 3–3 Case-control study analysis: odds ratio (OR).

Source: Reproduced from Jacobsen KH. *Introduction to health research methods: a practical guide*. Sudbury MA: Jones & Bartlett; 2012.

population as a whole is somewhere within this range of possible ORs" (**Figure 3–4**). If the entire confidence interval is greater than 1, the result is said to be statistically significant, and the conclusion is that the exposure appears to be risky. If the entire confidence interval is less than 1, the result is also statistically significant, and the conclusion is that the exposure appears to be protective. If the confidence interval overlaps 1, it means that there is not strong evidence that the exposure is risky or protective, and the conclusion is that there is no statistically significant association between the exposure and the outcome in the study population. In the example shown in **Figure 3–5**, the OR and 95% confidence interval is 0.56 (0.32, 0.93). Because the entire range is less than 1, the association between the exposure and disease is said to be statistically significant, and the conclusion is that cases were much less likely than controls to have had the exposure.

3.3.D Cohort Studies

A **cohort** is a group of similar people, and **cohort studies** recruit a group of similar people and follow them forward in time. At the start of the study, the researchers ask all of the participants about a variety of health behaviors and other exposures and characteristics, and they confirm that

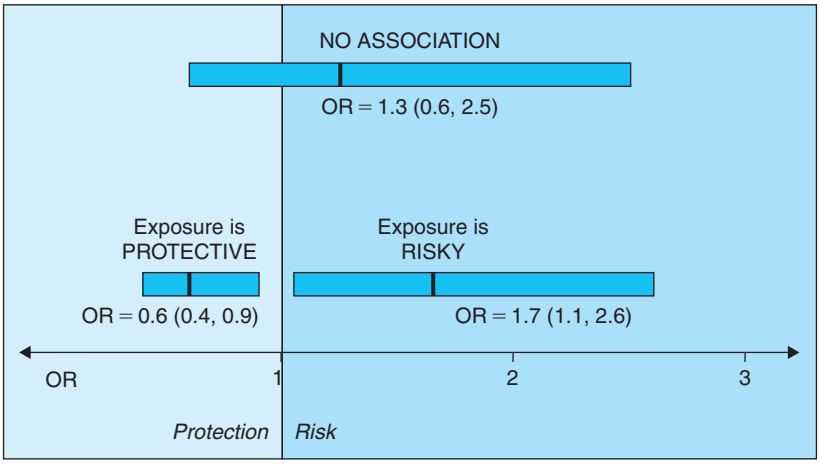

Figure 3–4 Case-control study analysis: 95% confidence interval for the OR.

Source: Adapted from Jacobsen KH. *Introduction to health research methods: a practical guide*. Sudbury MA: Jones & Bartlett; 2012.

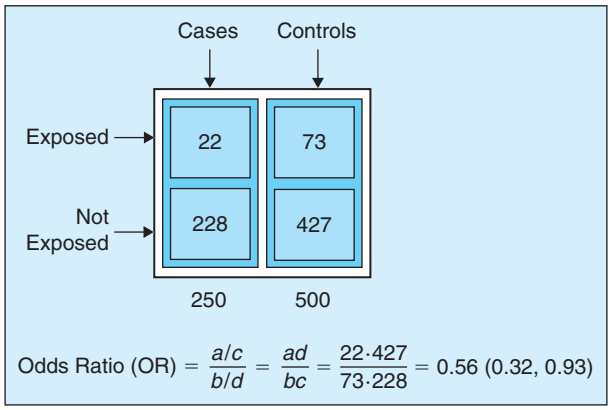

Figure 3–5 Example of a case-control study.

no one enrolled in the study already has the disease outcome of interest. The participants are then tracked for months or years, so that researchers can count the number of people who develop the disease or disability of interest. Statistical analysis is used to compare the rate of incident (new) disease among those with a particular exposure and those without that exposure.

Because the data collected at the start of the study can prove that an exposure existed before the onset of disease, cohort studies are very helpful for establishing whether an exposure causes a disease. Cohort studies are also useful for measuring the incidence of new disease in a population. The population studied can be a representative sample of a whole community or even a whole country. For example, the Framingham study has been following thousands of residents of one town in Massachusetts since 1948,[2] and the Whitehall studies have followed British civil servants from all occupational classes since the 1960s.[3] Another option is for people with an unusual exposure, such as an exposure to a particular industrial chemical, to be recruited and tracked for a long time so that researchers can study the impact of that rare exposure on the participants' future health status.

Two of the most common measures of association for a cohort study are the rate ratio and the attributable risk. The incidence **rate ratio** (also called the **risk ratio** or **relative risk**, or simply shortened to **RR**) is calculated by dividing the rate of incident disease in the exposed cohort by the rate in the unexposed cohort (**Figure 3–6**). An RR near 1 means that exposed and unexposed participants were equally likely to develop the disease during the study period. An RR greater than 1 indicates that the exposure was associated with increased risk of disease. An RR less than 1 indicates that the exposure was protective. Confidence intervals can be used

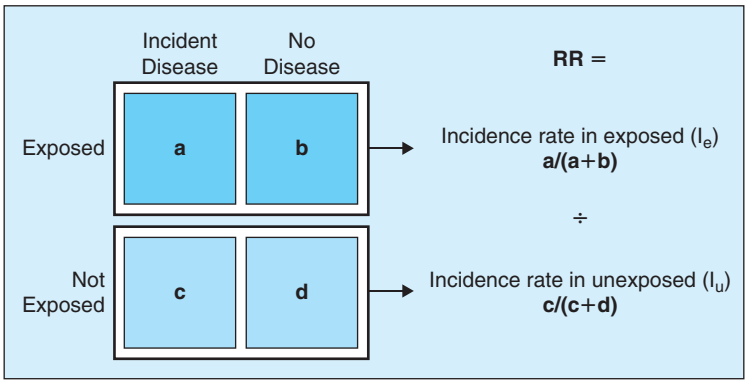

Figure 3–6 Cohort study analysis: rate ratio (RR).

Source: Reproduced from Jacobsen KH. *Introduction to health research methods: a practical guide*. Sudbury MA: Jones & Bartlett; 2012.

to show the level of certainty about the RR in the larger population from which participants were drawn (**Figure 3–7**). The **rate difference** (also known as **excess risk** or **attributable risk**) subtracts the rate of disease in the unexposed from the rate of disease in the exposed. If the exposed and unexposed populations were similar except for their exposure status, then this difference in disease rates represents cases of disease among exposed people that would not have occurred if they had not been exposed (**Figure 3–8**). In the example shown in **Figure 3–9**, the RR and 95%

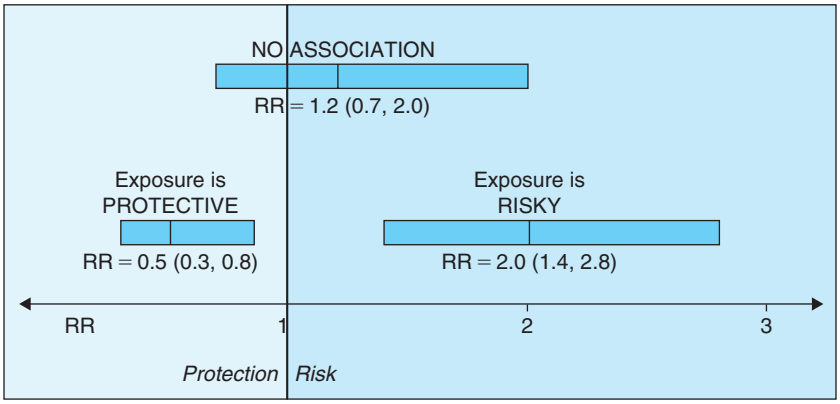

Figure 3–7 Case-control study analysis: 95% confidence interval for the RR.

Source: Reproduced from Jacobsen KH. *Introduction to health research methods: a practical guide*. Sudbury MA: Jones & Bartlett; 2012.

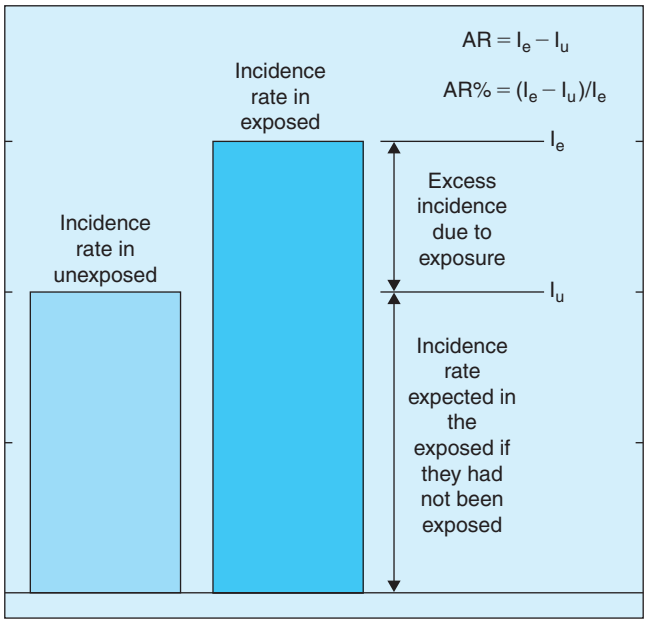

Figure 3–8 Cohort study analysis: attributable risk (excess risk).

Source: Reproduced from Jacobsen KH. *Introduction to health research methods: a practical guide*. Sudbury MA: Jones & Bartlett; 2012.

confidence interval is 2.00 (1.46, 2.74). Because the entire range is greater than 1, the association between the exposure and disease is said to be statistically significant, and the conclusion is that the exposure is a risk factor for the disease. The attributable risk percent is 50%, which means that half of the cases of disease among the exposed participants could have been prevented by removing the exposure.

3.4 EXPERIMENTAL STUDIES

Experimental studies, sometimes called **intervention studies**, are studies in which the researchers assign participants to receive a particular exposure. Experimental trials are the best study design for assessing causation, because the researchers intentionally subject participants to an exposure and then see what happens afterward. But because the researchers may be placing participants at risk of unexpected and potentially serious adverse outcomes, there are some special ethical concerns associated with experimental studies.

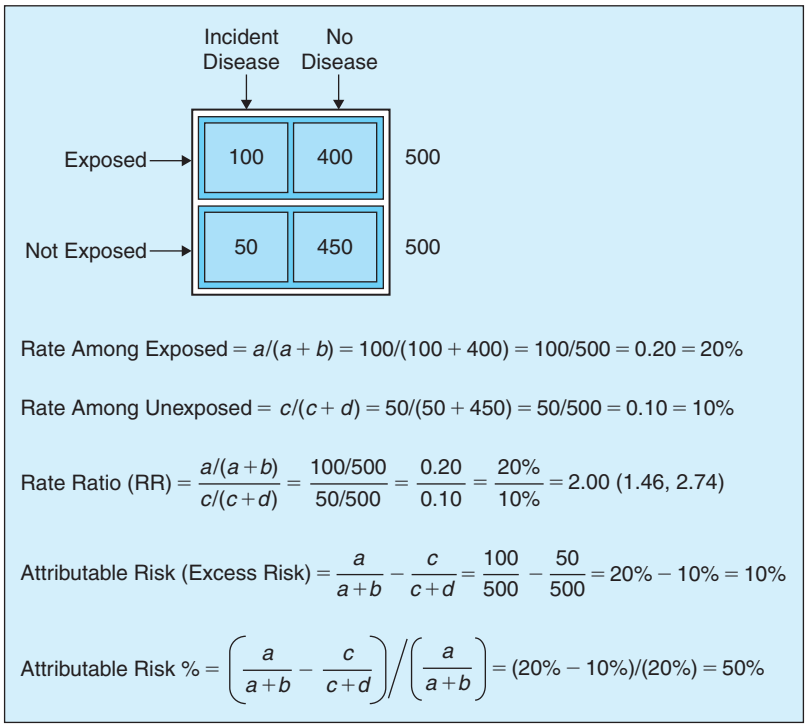

Figure 3–9 Example of a cohort study.

Some studies are deemed too risky to be conducted. Those that are approved are closely monitored by research ethics committees.

Some experimental studies are **clinical trials** of a new medication, a new vaccine, another new medical product, or some other intervention. Most clinical trials use a **randomized controlled trial (RCT)** design in which some people are assigned by chance to the active intervention and others are assigned by chance to a comparison group. The comparison may be a placebo, like a sugar pill or saline injection, or may be an active control such as the best drug already on the market or a lower dose of the new medication being tested. Most clinical trials are **double blind**, which means that neither the participants nor the people assessing the participants' health outcomes know whether a participant is receiving the trial drug or a placebo. That way neither the participants nor the examiners will be tempted, even subconsciously, to find a better outcome in a patient who they know is taking the new drug.

The most common outcome measure for an RCT is the **efficacy** of the intervention, which measures the ability of the intervention to produce

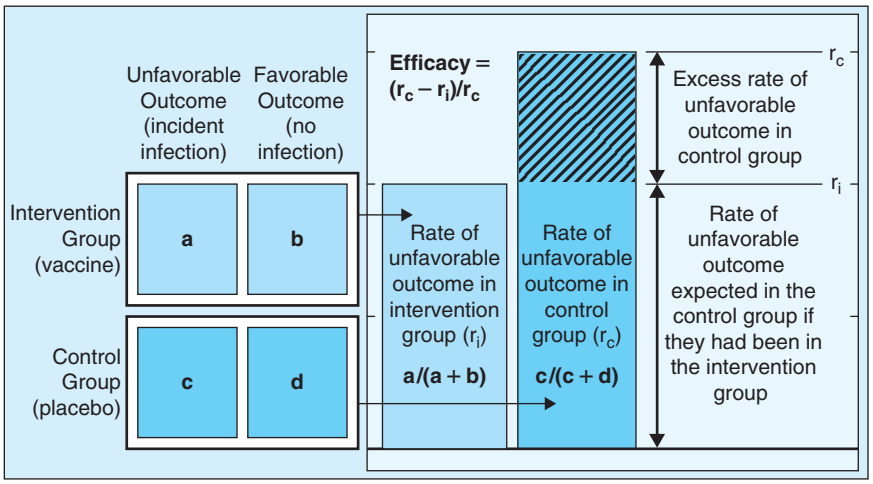

Figure 3–10 Experimental study analysis: efficacy.

Source: Adapted from Jacobsen KH. *Introduction to health research methods: a practical guide*. Sudbury MA: Jones & Bartlett; 2012.

the desired effect. For example, a vaccine trial with a placebo control will evaluate how well the vaccine prevented infection by comparing the rates of infection in the vaccine group and the placebo group (**Figure 3–10**).

Most new vaccines, medications, and other pharmaceutical agents undergo several rounds of testing before the product is released to the public. The first phases of the study test the safety of the product in small numbers of people. Later phases recruit hundreds or thousands of people to ensure the safety and effectiveness of the new product. Ongoing safety monitoring continues after the product is in wide use.

3.5 RESEARCH ETHICS

Nearly all health research projects that involve contact with people or access to identifiable personal information are supervised by ethics committees, commonly called **Institutional Review Boards (IRBs)** or **Research Ethics Committees (RECs)**. Review boards will not approve studies that do not meet the three main ethical considerations of health research: beneficence, respect for persons, and distributive justice.

Beneficence means that the study should be beneficial for the participants and for their communities. This call to do good is often paired with

nonmaleficence, from the root words for nonbadness, which call for the study to do no harm.

Respect for persons demands that all potential participants have the **autonomy** to choose whether they want to volunteer to participate in a study and that all potential participants are given all the information they need to be able to make an informed decision about whether to participate. Candidates for a research study should be told about the goals of the study, the potential risks and benefits of participation, the study procedures, the time requirements of participation, and the process for withdrawing from the study if they change their mind about participating. The process of sharing information and agreeing to participate is called **informed consent**. No one should feel pressured to participate in a research study. Respect for persons also requires researchers to keep the safety of participants as their top priority and to protect the privacy of participants and the confidentiality of the information participants choose to disclose.

Distributive justice aims to ensure that the populations that bear the risks of research participation have access to the benefits of that research. For example, this means that a community that took on the risk of volunteering to test a new medication or a new vaccine should also have the benefit of having access to that product after it is approved and marketed.

Adherence to these standards prevents the types of research misconduct that occurred in the mid-20th century when medical experiments were often conducted without participant consent. One of the most widely known examples from the United States is the Tuskegee Syphilis Experiment, which was conducted by the U.S. Public Health Service in Alabama for 40 years beginning in 1932.[4] Nearly 400 African-American men with late-stage syphilis were offered free medical care by doctors who were conducting a study on the progression of the disease. The men were not told that they had syphilis and many were not treated with antibiotics even after penicillin became the standard cure for the infection in 1947. Treatment was provided only after a major newspaper reported on the study in 1972. By that time many of the men had died of syphilis and many of their wives had been infected.

In order to protect the "human subjects" of research, new and ongoing research projects must be approved by an independent IRB in each country where study data will be collected. Research ethics committees will not approve studies that they deem to be unreasonably dangerous, poorly planned, or unnecessarily targeting members of a vulnerable population. After approval, the IRBs monitor ongoing studies. Researchers must seek prior approval for any changes they want to make to their research

protocols, and they must immediately report any adverse event to the IRBs that are overseeing the project. These rules help researchers to design and implement high-quality research plans, and they help to ensure the protection and safety of all research participants.

3.6 SYNTHESIS STUDIES

Some research investigations synthesize the results of dozens or hundreds of previous primary studies. By combining and analyzing the results of multiple similar research studies from different parts of the world or various points in time, these tertiary analysis studies provide a comprehensive summary of the scientific literature on a particular topic, a foundation for projections about health problems and needs, and new insight about risk factors for particular diseases.

3.6.A Correlational Studies

A **correlational study**, sometimes called an **ecological study**, uses numeric data about a particular exposure and a particular health outcome from several populations to look for trends. The results of correlational studies are often displayed using a scatterplot. For each population, a point is placed on the graph by using the value for the exposure in that population as the x-coordinate and the value for the health outcome in that population as the y-coordinate. After all the points are plotted, a line that represents the best fit to the points is added to the graph.

The value of the **correlation** coefficient, r (which is often reported as r^2), measures how well the line predicts the location of the points (**Figure 3–11**). An r near 1 means that all of the points fall almost exactly on a line, so if the exposure level in a population is known, the outcome can be predicted with a high level of certainty. An r near 0 is extremely weak and means that the line has no predictive value. In other words, when r is close to 0, the exposure is not associated with the disease outcome. An r near 0.5 indicates a moderately strong correlation.

The slope of the line shows the direction of association. A positive value for r indicates a positive slope (one that goes up from left to right) and signifies that an increase in the rate of the exposure is associated with an increase in the rate of the outcome. For example, a study of 37 countries found that countries with greater income inequalities also had greater rates of bullying among middle school students ($r = 0.62$).[5] A negative value for

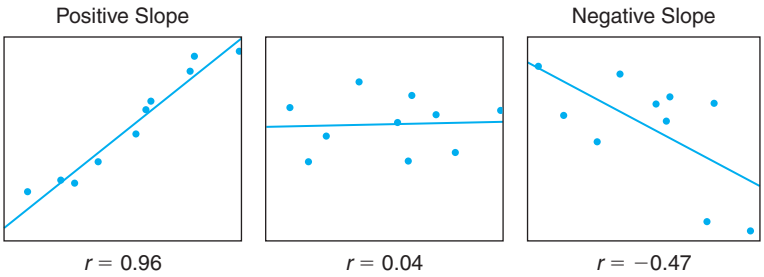

Figure 3–11 Analysis for a correlational study: correlation.

r indicates a negative slope and shows that an increase in the exposure rate is associated with a decrease in the outcome rate. For example, a study of 43 African countries found that countries with a higher proportion of adults who could read had a significantly lower proportion of women who died in childbirth ($r = -0.52$).[6]

These two examples highlight a key aspect of ecological study design: most ecological studies are designed to examine population-level exposures and outcomes, like income inequality, literacy rates, air quality, and maternal mortality rates. Ecological studies are not used to test individual-level correlations. Thus, the results of ecological studies that use population-level data can be applied to populations but cannot necessarily be applied to individuals. For example, if an ecological survey using data from dozens of cities shows a strong positive correlation between the number of tanning beds per 1000 adults and the number of new skin cancer diagnoses each year per 10,000 adults, this finding would not prove that the individuals who used tanning beds were the ones who were diagnosed with skin cancer. It is possible (though unlikely) that none of the individuals who use tanning beds developed skin cancer. The **ecological fallacy** describes when population-level correlations are incorrectly interpreted to be measures of individual risk. However, even with this limitation, ecological surveys can be a very helpful first step in testing a hypothesis about a possible risk factor for disease.

3.6.B Systematic Reviews and Meta-Analyses

Systematic reviews identify as many articles and reports about a particular topic as can be found, then check each one to see if it meets the predefined criteria for inclusion in the analysis. Information from each eligible article is extracted and compared with the other studies in order to paint a comprehensive picture of what is known (and what is not known) about the topic.

For example, a systematic review might show strong agreement in the literature about a particular exposure being a risk factor for a specific disease. Or the review might suggest that a particular exposure does not appear to increase the prevalence of a disease. Or a review might determine that the previous studies have mixed results and no consensus can be reached about the association based on the current scientific literature. (To ensure a fair conclusion, systematic reviews usually consider the potential effects of publication bias on their findings. **Publication bias** occurs when studies that find a statistically significant result are more likely than "null result" studies to published.)

When the study designs and the statistics used for each of the included studies in a systematic review are quite similar, it is sometimes possible to pool the results from the independent studies to create one summary statistical measure. This combined statistical analysis is called a **meta-analysis**.

3.6.C Forecasting and Modeling

Mathematical models can be used to estimate disease rates in populations lacking good data and to predict future health trends. For example, global burden of disease studies generally start with a systematic review of the disease of interest, and usually find incomplete data for many countries in the world. A mathematical model that incorporates the information that is available for a country (such as the distribution of the population by age) and estimates of morbidity and mortality from other countries with similar geographic and socioeconomic profiles can provide a good foundation for understanding the likely health profile of the understudied country. Similarly, if health and demographic data from several points in time in one country or world region are added to a model, researchers can create projections about the likely health situation in that area in 10, 25, or even 50 years. Models can also be modified to simulate the short-term and long-term effects of public health interventions and other population-level changes.

3.7 INTERPRETING STATISTICS

The only way to have a 100% accurate measure of health in a population is to collect data from every individual in that population. However, this is rarely done for large populations because of time and money constraints. Instead, a small proportion of the population is recruited for the study, and statistical tests are used to provide estimates of the health status in the whole population based on measures taken from the sample population.

Figure 3–12 shows an example of sampling. Of the 100 people in the population from which participants will be sampled, 32 are obese. If 10 individuals are sampled at random from that population, it is likely that the prevalence of obesity in the sample population will be 20% or 30% or 40%, something relatively close to 32%. However, some samples will, by chance, have a prevalence of 80% or 90% or even 100%. So a sample of 10 will not allow a high level of precision about the prevalence rate in the total population, even though it can provide a reasonable rough estimate.

The uncertainty about what the sample measure says about health in the larger population from which participants were drawn can be captured in a 95% confidence interval (95% CI) similar to the ones used for ORs and RRs. For example, the obesity prevalence estimate and its 95% CI for a sample of 10 individuals may be 30% (8%, 62%). This means that based on one sample of 10 people, we are 95% confident that the true prevalence in the larger population is somewhere between 8% and 62%. That range does indeed capture the 32% that is the true prevalence. (There is a 5% chance that this sample happened to be extreme, and that the 95% CI does not include the true value of the prevalence.) If a narrower CI is desired, then a larger sample population must be used. For example, if 50 individuals participate in the study instead of 10, then the estimate of the prevalence might be 34% (22%, 48%). That would mean that based on one sample of 50 people we are 95% confident that the true prevalence is somewhere between 22% and 48%. If more people are sampled, the 95% CI would become even narrower and closer to the true value of 32%.

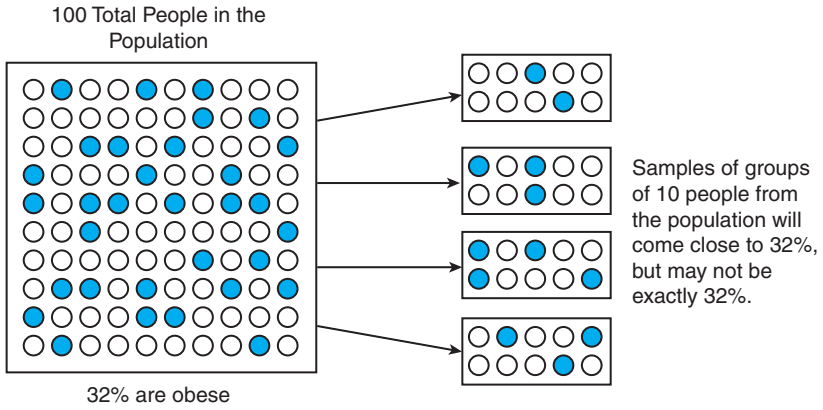

Figure 3–12 Sampling from a population.

A **p-value** (or probability value) is another measure of uncertainty. P-values are usually used with statistical tests of difference to indicate how likely it is that a difference exists between two or more populations. For example, a **t-test** can be used to see if there is a difference in the mean age of men and women who participated in a cross-sectional study, and a **Chi-squared test** can be used to see if there is a difference in the proportion of adults with diabetes in four different neighborhoods in one city. The p-value for these tests gives the estimated probability that, given the number of people in the sample, an even bigger difference between the groups than the one found for the sample might occur by chance even if there really was no difference between the groups being compared. Just as a larger sample size results in a narrower confidence interval, a larger sample size makes it easier to have a small p-value. A larger sample gives the test greater statistical **power**, a better ability to detect a difference between two or more groups when the groups really are different.

A small p-value (usually less than 0.05, or 5%) means that the statistical test found that it is unlikely that a larger difference between groups would occur by chance. Tests that produce p-values less than 0.05 are said to have statistically significant results and are deemed to show a difference between the groups being compared. For example, if a t-test comparing the mean ages of men and women has a p-value of 0.02, there is only a 2% likelihood that such an extreme difference in mean age would be observed by chance if there was really no difference between the populations. Because it is so unlikely that the apparent difference in mean ages of men and women was due to chance, the conclusion for this test is that the mean ages for men and women in that population are different. If the t-test has a p-value of 0.68, there is a 68% likelihood that an even bigger difference could be observed by chance if samples were drawn from populations with the same mean age. In this situation, the conclusion is that the means are not different. Other examples of p-value interpretation are shown in **Table 3–1**.

Knowing the meaning of CIs and p-values allows a reader to interpret almost all statistical results. A p-value of less than 0.05 means that a difference exists. A larger p-value means that no difference has been observed. A CI provides a range of likely values for a measure in a population, and that range gets narrower as the sample size increases. The CIs for ratios (like ORs and RRs) are centered around 1; if the CI does not include 1, then there is a difference between the two populations being compared by the ratio.

Table 3–1 Examples of p-value interpretation.

Goal of Test	*p-value*	*Is the p-value "extreme" ($p < 0.05$)?*	*Conclusion*
To compare mean ages of men and women	0.13	No	The mean age of the populations is **not different**.
To compare mean scores on a test for children in grades 1, 3, and 5	0.002	Yes	The mean test scores by grade are significantly **different**.
To compare the proportion of employees in different divisions of a company who walk or bike to work each day	0.43	No	The distribution of responses by the various groups is **not different**.
To compare the prevalence of diabetes in two cities	0.03	Yes	The proportions of people with diabetes in the two cities are significantly **different**.

3.8 CRITICAL READING

Several characteristics of good public health and medical research reports are listed in **Table 3–2**. Beyond these factors, there are several other important considerations for readers who are assessing and applying the results of published studies. These cautions include bias, measurement validity, and target populations.

Bias is a systematic error in study design, data collection, or data analysis that might create a difference between what the study intended to measure and what it actually measured. Bias of any type can lead to an overestimation or an underestimation of the association between an exposure and an outcome. **Selection bias** occurs when the people who participate in a study are not representative of the intended sample population. One example of selection bias is **volunteer bias**, which occurs when people who volunteer to be part of a research study turn out to be different from the desired sample population, perhaps because they are systematically healthier or less healthy than the population as a whole.

Table 3–2 Characteristics of good public health and medical research reports.

- The article has been peer-reviewed and published in a respected journal or by a trusted organization.
- The population studied is clearly defined and the number of participants in the study was reasonably large.
- An appropriate epidemiological study design was used (with a control group, if required).
- The methods used to measure exposures/interventions and health outcomes are explained in detail.
- The results of statistical tests are presented using easy-to-read charts, graphs, and tables.
- The relationship of the new study to previous studies is discussed, and many other articles are cited.
- The strengths and limitations/biases of the study are acknowledged and discussed.
- The conclusions seem reasonable and are based on the new results.
- The article is well written and follows a logical outline (usually Introduction/Background, Methods, Results, Discussion/Conclusion).
- The article states that the study was approved and overseen by an ethics review committee, and there are no obvious conflicts of interest that may have influenced the findings.

Information bias occurs when incorrect information is given to researchers. For example, **recall bias** may happen when participants do not accurately recall past events. When there is differential recall—say, when people with cancer strain to recall any potentially harmful past exposure but people in the control group of a case-control study are not similarly motivated to remember past exposures—the results of a study may be inaccurate. Bias can be avoided or minimized when a study is carefully designed, conducted, and analyzed. The discussion sections of articles usually include an explanation of the limitations of the study and the possible sources of bias (as well as potential issues related to confounding, effect modification, or other scenarios that could affect the results). Readers can also make their own evaluations about whether bias in a study may have influenced the findings.

Validity asks how well a test measures what it is supposed to measure (internal validity) and how well a study measures the true situation in a population (external validity or **generalizability**). A test should be **accurate** (valid), which means that it gives the actual values of height, blood pressure,

or some other measure. A test should also be **precise** (**reliable**), which means that when the test is given several times to the same person the results are consistent. Most global health research articles provide details in their methods sections about their survey instruments (questionnaires), clinical and laboratory tests, and other assessments. Readers should check to be sure that the questions used for a survey appear to accurately capture the exposure of interest. Self-reported measures (like "how many calories did you eat today?" and "how many miles did you walk today?") may not be as accurate as observed measures (such as having a researcher quantify the actual portions of various foods eaten during the day and using a pedometer to record the number of steps taken). Additionally, different types of questions may yield different responses (such as "how many servings of vegetables did you eat today?" versus "how many servings of vegetables do you eat in a typical day?").

A third consideration relates to the populations to which the results of a study can be applied. The conclusions of a study that included only male participants between 20 and 24 years of age should not be applied to women ages 80 to 89. The conclusions of a study of women living in California probably do not apply to women living in Malawi. The conclusions of a study of nonsmokers should probably not be applied to smokers. Results from tests done in rats do not necessarily apply to humans. Caution should be used when applying the results of any study conducted in one place at one point in time to another population.

3.9 EVIDENCE-BASED GLOBAL HEALTH

Clinicians often use a process called **evidence-based medicine (EBM)** to guide them as they seek to make the best decisions about how to care for their patients. The goal of EBM is to use facts rather than anecdotes or ideology to make clinical decisions. Like EBM, evidence-based public health and evidence-based global health require a careful and critical review of the literature prior to implementing any intervention (**Figure 3–13**).[7] The goal of this review process is to learn what has worked for others and what has not worked. There is no need to "reinvent the wheel" when the global health literature is full of examples of successful public health programs and policies. There is also a substantial benefit to learning from the mistakes others have made and not repeating them. Ideally, evidence-based global health promotes the cost-efficient and effective use of limited health resources to solve public health problems. The tools in this chapter

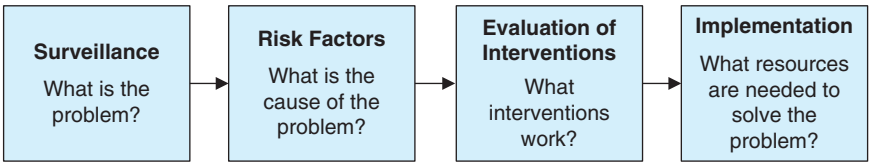

Figure 3–13 The public health approach.

Source: Adapted from Holder Y, Peden M, Krug E, Lund J, Gururaj G, Kobusingye O, editors. *Injury surveillance guidelines*. Geneva: WHO; 2001.

provide the foundation for conducting a valuable analysis of any global health concern or intervention. It is not an overstatement to say that health research saves lives.

3.10 DISCUSSION QUESTIONS

1. Look at the health webpage for a popular Internet news site. What topics are covered? How many of the stories present the results of a research project?
2. Would you participate in a health research study? Why or why not? Would your answer be different for an observational study and an experimental study?
3. Use PubMed or another abstract database to find an academic journal article on a health topic of interest to you. Read the article to find the answers to these questions: (a) What was the main study question?, (b) Who participated in the study, where did it take place, and when was it conducted?, (c) What study design was used?, and (d) What was the answer to the main study question?
4. Look up the most recent issue of the *State of the World's Children* or another United Nations report and examine the statistical annex. What sources of data contributed to these tables?
5. Find a recent news story from the popular press about a newly released health research report. Look up and read the scientific article on which the news report was based. Was the news story accurate? Did it leave out any critical information?
6. How have the results of health research studies contributed to improving your quality of life?
7. How do the results of health research studies contribute to improving global health?

References

1. Jacobsen KH. *Introduction to health research methods: a practical guide*. Sudbury MA: Jones & Bartlett; 2012.
2. Splansky GL, Corey D, Yang Q, et al. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol*. 2007;165:1328–1335.
3. Marmot M, Brunner E. Cohort profile: the Whitehall II study. *Int J Epidemiol*. 2005;34:251–256.
4. White RM. Unraveling the Tuskegee Study of Untreated Syphilis. *Arch Intern Med*. 2000;160:585–598.
5. Elgar FJ, Craig W, Boyce W, Morgan A, Vella-Zarb R. Income inequality and school bullying: multilevel study of adolescents in 37 countries. *J Adolesc Health*. 2009;45:351–359.
6. Alvarez JL, Gil R, Hernández, Gil A. Factors associated with maternal mortality in sub-Saharan Africa: an ecological study. *BMC Public Health*. 2009;9:462.
7. Buekens P, Keusch G, Belizan J, Bhutta ZA. Evidence-based global health. *JAMA*. 2004;291:2639–2641.
8. Holder Y, Peden M, Krug E, Lund J, Gururaj G, Kobusingye O, editors. *Injury surveillance guidelines*. Geneva: WHO; 2001.