introduction

LEADING CAUSES OF DEATH

Acknowledgments

Leadership for this project was provided by CDC's Division of Health Promotion Statistics, National Center for Health Statistics and Division of Media and Training, Public Health Practice Program Office. The training package would not have been possible, however, without the contributions of many individuals dedicated to the application of data to improve public health practice. We thank them for their patience and devotion throughout the development process.

The following individuals shared their technical expertise in the use of public health data and reviewed successive drafts:

Centers for Disease Control and Prevention

Ralph Caraballo, PhD
Sally Curtin, MA
Pat Drehobl, RN, MPH
Donna Hoyert, PhD
David Nelson, MD, MPH
Michael Schooley, PhD
Stephanie Ventura, MA
Chris Zahniser, RN

State Health Departments

Miebi Akah, PhD (Ohio)
Kathy Blue (North Carolina)
Carol Friedman, DO (Texas)
Ned Hayes, MD (Maine)
Louise Lex, PhD (Iowa)
Stephen Pickard, MD (Kansas)
Bob Rolfs, MD, MPH (Utah)
Cathie Schumacher, MD, MPH, (Utah)
Jennifer Woodward, PhD (Oregon)

Six health agencies graciously hosted focus groups and field tests of the training materials:

Cobb County Health Department, Georgia
Georgia Department of Public Health
Iowa Department of Health
Ohio Department of Health
Oregon Department of Health
North Carolina Department of Health

A special thanks to two CDC staff, Ken Keppel, PhD, NCHS, and Nancy Gathany, PHPPO, who provided leadership, policy guidance, and overall coordination for this project. Together with our training consultant, Susan B. Toal, MPH, and two video/communications experts, Katie Baer, MPH, and Jim Rothschild, they merged creativity with scientific integrity to produce a product that we hope will benefit a broad public health audience.

These training materials are being made available through the Public Health Training Network (PHTN).

Why was this course developed?

Effective use of data has always been an essential component of public health practice. Its value has increased considerably in recent years, however, as health departments across the country are called upon to perform the three core functions of public health — **assessment, policy development, and assurance** — as outlined in *The Future of Public Health.*[1]

The goal of this course is to improve the ability of public health professionals to understand and use data routinely in their jobs. The course is intended primarily for health practitioners working in local and state health departments or community-based health agencies. It will be most useful for individuals who have limited experience in using and interpreting data.

What is covered in the course?

We have anticipated that participants in the course will come from a variety of disciplines, occupations, and organizational affiliations. Additionally, we have assumed that the course will be used in a variety of ways:

- As a classroom course with an on-site instructor;
- As a distance-learning course in a self-study format; and
- As a stand-alone package for reference or self-study.

Thus, we have divided the course content into a sequence of modules, each of which can be completed during a half-day (2-4 hour) session. The modules build on each other and are meant to be completed sequentially.

- **Module 1** illustrates how a state health department responded to the Board of Health's concerns about teen violence.

- **Module 2** focuses on a local health agency's experience with an erroneous newspaper article about teen pregnancy.

- **Module 3** highlights the efforts of a local health department and a managed care organization to estimate the impact of smoking in their service area.

---

[1] Institute of Medicine. The Future of Public Health. Washington, D.C.: National Academy Press, 1988.

Each module consists of a video case study and a written workbook. The video case studies dramatize public health problems that can be addressed using data. The situations portrayed in the case studies are based on real experiences that have been adapted for instructional purposes. The aim of this approach is to create dramatic interest, help you connect the health problems with your own professional concerns, and make the course more realistic and engaging.

At specific points, the video will instruct you to stop the tape and open your workbook. Here you will read supplementary text on key issues in data collection and analysis, and have an opportunity for reflection and discussion. The workbook also contains structured exercises which require you to independently practice or apply what you have learned. In most of the exercises, you will be working through the same problems that the characters in the video are facing. This hands-on practice will enable you to acquire the knowledge and skills addressed in the course.

What are the learning objectives for each module?

Module 1:

Introduces you to the concept of quantitative data and its value for public health. The module describes the major types of data available for analysis, then focuses specifically on one type: data on deaths (or mortality data). You will explore where these data originate, how they are coded, and how they can be used to answer important questions.

At the end of Module 1, you will be able to:

- Define population-based data.
- Explain data typically included on a death certificate, and discuss the value and limitations of those data.
- Explain ICD codes.
- Describe the difference between numbers of deaths (counts) and death rates.
- Identify and analyze leading causes of death.
- Calculate general death rates and specific types of death rates, including age-specific, cause-specific and sex-specific death rates.
- Use rates to examine differences in causes of death by sex.
- Analyze changes in rates over time (trends) and changes in underlying causes over time.
- Determine effective ways to present data.

Module 2:

Builds on the first module and introduces you to other sources of quantitative data: birth certificates, reports of fetal death, and abortion reports. You will focus on the subject of teen pregnancy and explore how these data are generated, and how they can be used (or misused) to answer important questions.

At the end of Module 2, you will be able to:

- Describe the respective roles of health departments and the media in providing information to the public on health-related issues.
- State the definitions of teen pregnancy and live birth, and describe the difference.
- Describe the difference between occurrence and residence data.
- Identify sources of data on teen pregnancy, and discuss the strengths and limitations of those data.
- Calculate age-specific pregnancy and live birth rates.
- Determine percentage changes in rates over time.
- Use confidence limits and statistical significance to interpret variability.

Module 3:

Takes you even further into the world of data by exposing you to a rich data source on personal health behaviors. You will explore how these data are collected and analyzed, and how they can be used to estimate a person's risk of developing certain chronic diseases.

At the end of Module 3, you will be able to:

- Describe the Behavioral Risk Factor Surveillance System (BRFSS), including its purpose and potential uses, its methodology, and its strengths and limitations.
- Explain prevalence rates and how they differ from incidence rates.
- Compare aggregate smoking prevalence data across years in a region (multi-county) with data for the state as a whole.
- Define the concepts of relative and attributable risk.
- Estimate the number of deaths attributable to smoking in a state and region.

What is the role of the instructor?

If you are working on this course in a classroom or group setting, you will most likely have an instructor. He or she is responsible for guiding you through the course and keeping the training on track. The instructor will most likely:

- Open the course and welcome participants.
- Ask you to introduce yourself to other participants.
- Handle administrative and logistic details.
- Set a pace for the course that is compatible with participants' learning styles and individual needs.
- Provide additional explanation of concepts covered in the course.
- Help apply the course concepts to local situations and nuances.
- Lead group discussions on key issues.
- Answer questions to clarify concepts or resolve confusion.
- Conclude the training with an end-of-course evaluation questionnaire.
- Arrange for participants to receive appropriate continuing education credits for course completion.

☞ ***Now, some background information to provide you with a foundation for this course . . .***

How can this course improve your use of data?

As public health workers, you are concerned with preventing or controlling disease, disability and injury, and with improving the quality of life for the residents of your state or community. How do you know what the major health problems are in your area? How do you know where to focus your efforts and resources? How do you know if you are making progress in solving your community's health problems?

Data can help answer all of these questions — and more. Valid, reliable data on the communities we serve can help us to:

- assess the health of a community or population;
- search for causes of disease, injury and disability;
- plan programs to meet community needs; and
- measure progress in prevention and control efforts.

What are population-based data?

Quantitative health data can focus on *individuals* or entire *populations*. A classic example of individualized data is a patient's medical record. Each record is devoted exclusively to one person and contains information about his or her unique illnesses, injuries, behaviors, etc. The data are used primarily to improve the health of that one individual.

In public health, however, our focus is primarily on populations (communities, cities, counties, states). We are concerned with population-based data which tell us about the overall "health" of that population. If there are more people without a regular source of primary care in our community compared to other communities, we need to ask ourselves what can be done to increase access to care. If there are fewer youth in our community wearing bicycle helmets than in other communities, we need to ask ourselves what we can do to increase helmet use.

How have data been used historically to guide public health policy?

A few historical highlights in using data to improve public health:

- ***John Gaunt, 1662***: The first statistician to quantify patterns of birth, death and disease occurrence. He also noted male-female differences in occurrence, high infant mortality, urban-rural differences, and seasonal variations.

- ***John Snow, 1854***: Deemed the father of "field epidemiology," Snow determined that cholera was transmitted by contaminated water from a (Broad Street) water pump used by many community residents 30 years before the identification of the cholera vibrio by Robert Koch.[2]

- ***After World War II*:** The past 50 years have seen an explosion in the development of research methods and the application of statistics to the entire range of health problems. Classic studies include the 1950 study[3] linking smoking and lung cancer, and the Framingham Study[4] of cardiovascular disease, begun in the 1950s and still continuing.

---

[2] Snow J. Snow on cholera. London: Humphrey Milford: Oxford U Press, 1936.

[3] Doll R, Hill AB. Smoking and carcinoma of the lung. Br Med J 1950; 1:739-48.

[4] Dawber TR, Kannel WB, Lyell LP. An approach to longitudinal studies in a community: The Framingham study. Ann NY Acad Sci 1963; 107:539-56.

- ***Lead in gasoline:*** The discovery of the association between highways and lead poisoning in children led to the formulation of unleaded gasoline and subsequently to a recognizable reduction in the levels of blood lead in children.

- ***1992:*** Numerous reports from other countries showed an association between a prone sleeping position for infants and a higher incidence of Sudden Infant Death Syndrome (SIDS). Based on this research, the American Academy of Pediatrics recommended that healthy infants be placed on their sides or back to sleep.[5] Combined with a subsequent National Back to Sleep Campaign in 1994, SIDS rates fell 30% from 1992 to 1995.[6]

What is the difference between quantitative and qualitative data?

Quantitative data provide the *what, who, when,* and *where* of health-related events. They are measurable and tangible. Quantitative data involves the counting of people, behaviors, conditions, or other discrete events; classifying those events into categories; and using math and statistics to answer questions. Numbers of deaths can be used to identify leading causes of death (*what*). Numbers of smokers and nonsmokers by gender can be used to determine whether men are more likely to smoke than women (*who*). Keeping track of the number of people with flu can identify the beginning of the flu season (*when*). Comparing the proportion of women who began prenatal care after the first trimester in various counties will provide an indication of *where* access to prenatal services may be a problem.

Qualitative data can be used to explain the *why* and the *how* of health-related events. Qualitative data involve observing people in selected places and listening to discover *how* they feel and *why* they might feel that way. A focus group of teenage girls could provide valuable insights concerning *why* they do or don't use contraceptives. A visit to a local clinic might indicate *how* people might feel as they enter the waiting area.

This course focuses almost entirely on quantitative data.

---

[5] American Academy of Pediatrics Task Force on Infant Positioning and SIDS. Positioning and SIDS. Pediatrics 1992; 89:1120-6.

[6] American Academy of Pediatrics Task Force on Infant Positioning and SIDS. Positioning and SIDS: Update. Pediatrics 1996; 98(6):1216-8.

What are the major types of health data?

Health outcome (or health status) data measure the presence or absence of disease, injury, physical disability, or death. Death certificate data, for example, provide information about death as an outcome for specific age, race, and gender groups. Birth certificates, on the other hand, provide information about the outcome of a pregnancy in terms of gestational age and birth weight. Regular use of a prescription drug, restricted activity days, and absence from work or school due to illness can be viewed as health outcomes. Outcome variables may also consist of health indicators such as lung function, blood pressure, cholesterol levels, and mental status.

Risk factor data are associated with or explain a particular health outcome. The term *risk factor* encompasses direct causes or disease agents, personal characteristics and environmental factors that make individuals more or less prone to a particular disease or injury. Examples of a direct cause or disease agent would be high blood pressure and high total serum cholesterol, both of which predispose an individual to coronary heart disease. Personal characteristics include sociodemographic factors such as age, gender, and race; behavioral factors such as exercise, diet, and use of alcohol and other drugs; and genetic defects and genes that predispose individuals to disease or disability. There may also be risk factors in the environment, such as air pollution, contaminated ground water, the absence of fencing around backyard pools, and blocked emergency exits.

Resource data describe the resources available in communities to treat diseases or alleviate risk factors. Resources are frequently measured in terms of health care facilities and health care professionals; however, they also can include smoking cessation classes, exercise facilities, emergency shelters, and other resources that promote health. These facilities may also generate program data in the course of providing a program or service. Examples include: information about the number of times pregnant women were seen at a particular clinic; the proportion of people completing a smoking cessation class; the proportion of Medicaid-eligible women enrolled; and the number of volunteer hours required to operate an immunization program. Information about the cost of providing a program or service can also be very important. Such data make it possible to determine the cost effectiveness of one intervention compared to another or the cost of the same intervention in two different places.

Demographic data can also be important in understanding the population of interest. A retirement community, for example, will have greater needs for programs to measure blood pressure and cholesterol, interventions to promote physical activity, and long-term care facilities. A community with a large military facility may have many young families needing safe housing and affordable day care. It is also important to look at health outcome, risk factor, and resource data according to the demographic characteristics of individuals. Causes of death (outcomes) and risk factors may differ by age, race, gender, etc. Resources may also be designed to serve specific populations (e.g., day care centers for pre-schoolers, school-based clinics for adolescents, and respite care for the elderly).

What can quantitative data <u>not</u> tell us?

Unfortunately, quantitative data cannot tell us everything that we would like to know about a public health problem. Quantitative data cannot predict which individuals with a given risk factor will eventually contract the related condition. As an example, while smokers are more likely to get lung cancer, not all smokers will get lung cancer. Furthermore, lung cancer will also occur in some nonsmokers.

Quantitative data seldom tell us how to "fix" a particular public health problem. They may tell us, for instance, what proportion of people are wearing seat belts but they cannot, by themselves, tell us how to increase seat belt use. Even if we had quantitative data indicating that most people do not use seat belts because they find belts uncomfortable, we could not guarantee an increase in seat belt use with a more comfortable design.

Quantitative data can help us rank problems based on their size, their seriousness, or the effectiveness of available solutions. But they cannot tell us which problem is the most "important" or which problem to tackle first.

Quantitative data alone cannot tell us what to do as public health professionals. They are but one ingredient in a decision making process that must also incorporate our values as professionals and the values of the communities that we serve.

Now it's time to begin the first module. Return to the videotape and begin viewing Module 1.