

# Chapter 3

## Descriptive Epidemiological Methods

### Learning Outcomes

After completing this chapter, you will be able to:

1. Construct rate measures to identify population subgroups at high risk for a health problem. Provide the information in tabular and graphic forms.
2. Select a health problem and using the appropriate descriptive epidemiologic measures to characterize it in terms of person, place, and time.
3. Construct and interpret odds ratios and relative risks to identify potentially modifiable risk factors.
4. Determine which risk factors, identified from a set of data, could be targeted for a prevention intervention.

**Keyterms** Adjusted rate • Crude rate • Incidence rate • Life expectancy • Mortality rate • Odds ratio • Population attributable risk • Prevalence rate • Relative risk • Specific rate

### Introduction

The previous chapter discussed principles for ensuring the accuracy of the information obtained through the measurement process. These principles apply to both raw (primary) data and that which is calculated from the raw data (secondary data). Proper management of the documentation of the data collected, prevention of errors through well-designed data collection protocols, staff training in data collection procedures, and complete and detailed operations manuals support the reliability of the information obtained and ultimately help to ensure the validity of the information analyzed. These measurement fundamentals provide the basis for constructing the descriptive epidemiologic measures described in this chapter to aid in accurately assessing population health, potential for adverse exposures, and determining priorities for action.

Descriptive epidemiological methods characterize health events, health problems, and exposures in terms of person, place, time, and variability. These measures help answer the questions relative to the events, problems, and exposures: “Who is affected?” “Where?” “When?” and “How does this change over time?” There are numerous applications of descriptive epidemiological measures. Descriptive epidemiology measures are used to: identify populations and population subgroups or areas at particular risk, prioritize health problems, establish a baseline of the level of exposure of health problem for evaluating trends, and evaluating effectiveness of programs and policies in terms of progress toward achieving health goals.

This chapter will focus on those descriptive epidemiological measures commonly used in characterizing the health of a population. Application of the measures will be illustrated.

## Types of Rate Measurements

To quantify the magnitude of disease, injury, disability, other health event, or exposure it must be measured. In epidemiology, the measurement of events is expressed in terms of occurrence relative to a referent population instead of only raw numbers. Descriptive epidemiologic measures of the frequency of health events in population are called rates. Table 1 displays some of the common rate measures used in assessing the health of populations. The basics of computing rate measures follow.

### *Incidence and Mortality Rate*

An *incidence rate* is the occurrence of new or incident cases of a health event diagnosed within a defined period of time occurring in a population at risk for the event. A *mortality (or death) rate* is another descriptive measure representing the number of deaths occurring within a specified time in a population at risk for the death event. Both the incidence rate and the mortality rate are descriptive measures of risk. The generalized format for the computation of these rate measures is:

$$\text{Incidence rate (or mortality rate) per } 10^k = E_{(t)} / P_{(t)} \times 10^k$$

where  $E$  = number of events (or deaths) occurring in the population at risk during a specified period of time (usually a year) ( $t$ );  $P$  = average population at risk (or mid-year July 1 population) in the same geographic area at the same time  $t$  in which the events (or deaths) occurred;  $10^k = k$  factor of 10 to represent the occurrence of the event as a whole number relative to the population from which it was derived.

The *numerator* may refer to the onset of an illness or represent all new cases of a health event (morbidity) or a death (mortality) occurring in a population. To identify a new or incident case of an illness or condition, knowledge of the time of onset of

**Table 1** Common rate measures

Age-specific death rate per 100,000	=	$\frac{\text{No. of deaths in age group during a specific time period}}{\text{Midyear resident population * in age group during same time period} \times 100,000}$
Birth rate per 1,000 population	=	$\frac{\text{No. of live births in a population}}{\text{Midyear resident population} * \times 1,000}$
Case fatality rate per 100 persons	=	$\frac{\text{No. of deaths from a disease in a year}}{\text{No. diagnosed with disease in year} \times 100}$
Death rate per 100,000 population	=	$\frac{\text{No. of deaths in a population in a year}}{\text{Midyear resident population} * \times 100,000}$
Fertility rate per 1,000 women aged 15–44 years	=	$\frac{\text{No. of live births regardless of age of mother}}{\text{No. of women 15–44 years} \times 1,000}$
Fetal death rate per 1,000 live births plus fetal deaths	=	$\frac{\text{No. of fetal deaths with stated or presumed gestation of } > 20 \text{ weeks}}{\text{Sum of live births plus fetal deaths} \times 1,000}$
Hospital occupancy rate as a percent	=	$\frac{\text{Total annual no. patients}/365 \text{ days}}{\text{No. hospital beds, cribs, and pediatric bassinets on last day of reporting period} \times 100\%}$
Infant mortality rate per 1,000 live births	=	$\frac{\text{No. of infant } (< 1 \text{ year old}) \text{ deaths during a calendar year}}{\text{No. of live births during same year} \times 1,000}$
Maternal mortality rate per 100,000 live births	=	$\frac{\text{No. of maternal deaths}}{\text{No. of live births} \times 100,000}$
Neonatal mortality rate per 1,000 live births	=	$\frac{\text{Death of children under 28 days of age}}{\text{live births during some year} \times 1,000}$
Nursing home occupancy rate as a percent	=	$\frac{\text{No. of residents at the facility reported on the day of the interview}}{\text{No. of reported beds} \times 100\%}$
Perinatal mortality rate per 1,000 live births plus late fetal deaths	=	$\frac{\text{No. of late fetal deaths plus infant deaths within 7 days of birth}}{\text{Sum of live births plus late fetal deaths} \times 1,000}$
Postneonatal mortality rate per 1,000 live births	=	$\frac{\text{No. deaths of children under 28 days of age}}{\text{No. live births during same year} \times 1,000}$
Relative survival rate expressed as a percent	=	$\frac{\text{Observed survival rate for a patient group}}{\text{Expected survival rate for persons in the general population of same age, race, sex, and calendar year of observation} \times 100\%}$

\*July 1 in noncensus years, April 1 in census years

the event is essential. This can only be achieved if the population in a geographic area is already being monitored for changes in health status or will be involved in such monitoring. The ability to identify and define an incident event varies depending upon the condition. Death events and the cause(s) of death are identified from a death certificate (Fig. 1). For acute onset conditions, such as gastroenteritis, acute myocardial infarction, or nonfatal trauma, the identification of a new case from a medical record or registry can be pinpointed. Even with good clinical definition of a case, an incident case may be difficult to identify through administrative or medical records. An alternate source of morbidity incident cases is through a registry. Registries to maintain incidence information, be it for diseases (e.g., tumor registry), trauma cases, or for other conditions (e.g., birth defects) are difficult and costly to maintain. The emerging possibility of biomarkers used to identify accurately new cases, such as troponin testing for the presence of myocardial damage, offers the ability to track and predict disease trends which may impact health utilization and decrease mortality from heart disease (Sanfilippo et al., 2008). When the onset is not readily apparent, as is the case with chronic conditions, mental disorders, and behavioral problems, surrogate measures are used. For example, an incident (or new) case of cancer is defined by the date of histological confirmation. The incidence rate of addiction to an illegal drug, such as heroin, may be approximated from the date of first presentation to a treatment center, adjusting for a time lag in first use (Hickman et al., 2001). In some circumstances, multiple new events may occur within the same population in the time period of observation. This is exemplified in the calculation of nonfatal injury rates. The numerator may consist of either the number of persons who experienced an injury or the number of injuries occurring, depending upon the intent of the reporting. Reasons for not identifying an incident case or properly classifying a death event include: varying in the severity of a condition, lack of access to health care, lack of availability of diagnostic services, miscoding of the death event, new syndrome not previously recognized, or increased reporting of cause unknown/unspecified in the case of deaths (e.g., Sudden Infant Death Syndrome [SIDS]) (Shapiro-Mendoza et al., 2006). Regardless of the definition or ascertainment method for the event, the numerator must be derived from the same population comprising the denominator.

The *denominator* should not include those who already have the condition or who are not susceptible to it by virtue of immunity, immunization, surgery, or other factors that excludes the potential for being exposed. For example, in computing the incidence rate of prostate cancer neither females nor individuals who had their prostate removed should be included in the denominator. The choice of a denominator for a rate measure depends on the manner in which the time of observation for the events is considered. The denominator for constructing rate measures of events occurring in community populations is derived from the decennial census or from interim population projections based on census data. In this circumstance, the identification of persons who are truly at risk for the health event is not practical. The average population at risk, the population at mid-year (July 1) for noncensus years or April 1 for census years, is used for denominator in computing the rate. For rates calculated for smaller populations, such as an outbreak, corrections are

**U.S. STANDARD CERTIFICATE OF DEATH**

LOCAL FILE NO.		U.S. STANDARD CERTIFICATE OF DEATH			STATE FILE NO.	
1. DECEDENT'S LEGAL NAME (include AKA's if any) (First, Middle, Last)				2. SEX	3. SOCIAL SECURITY NUMBER	
4a. AGE-Last Birthday (Years)		4b. UNDER 1 YEAR Months      Days		4c. UNDER 1 DAY Hours      Minutes		5. DATE OF BIRTH (mo/day/yr)
7a. RESIDENCE-STATE			7b. COUNTY		7c. CITY OR TOWN	
7d. STREET AND NUMBER			7e. APT. NO.	7f. ZIP CODE		7g. INSIDE CITY LIMITS? <input type="checkbox"/> Yes <input type="checkbox"/> No
8. EVER IN US ARMED FORCES? <input type="checkbox"/> Yes <input type="checkbox"/> No		9. MARITAL STATUS AT TIME OF DEATH <input type="checkbox"/> Married <input type="checkbox"/> Married, but separated <input type="checkbox"/> Widowed <input type="checkbox"/> Divorced <input type="checkbox"/> Never Married <input type="checkbox"/> Unknown		10. SURVIVING SPOUSE'S NAME (if wife, give name prior to first marriage)		
11. FATHER'S NAME (First, Middle, Last)			12. MOTHER'S NAME PRIOR TO FIRST MARRIAGE (First, Middle, Last)			
13a. INFORMANT'S NAME		13b. RELATIONSHIP TO DECEDENT		13c. MAILING ADDRESS (Street and Number, City, State, Zip Code)		
14. PLACE OF DEATH (Check only one: see instructions)						
IF DEATH OCCURRED IN A HOSPITAL: <input type="checkbox"/> Inpatient <input type="checkbox"/> Emergency Room/Outpatient <input type="checkbox"/> Dead on Arrival			IF DEATH OCCURRED SOMEWHERE OTHER THAN A HOSPITAL: <input type="checkbox"/> Hospice facility <input type="checkbox"/> Nursing home/Long term care facility <input type="checkbox"/> Decedent's home <input type="checkbox"/> Other (Specify):			
15. FACILITY NAME (if not institution, give street & number)			16. CITY OR TOWN, STATE, AND ZIP CODE		17. COUNTY OF DEATH	
18. METHOD OF DISPOSITION: <input type="checkbox"/> Burial <input type="checkbox"/> Cremation <input type="checkbox"/> Donation <input type="checkbox"/> Entombment <input type="checkbox"/> Removal from State <input type="checkbox"/> Other (Specify):			19. PLACE OF DISPOSITION (Name of cemetery, crematory, other place)			
20. LOCATION-CITY, TOWN, AND STATE			21. NAME AND COMPLETE ADDRESS OF FUNERAL FACILITY			
22. SIGNATURE OF FUNERAL SERVICE LICENSEE OR OTHER AGENT			23. LICENSE NUMBER (of Licensee)			
ITEMS 24-28 MUST BE COMPLETED BY PERSON WHO PRONOUNCES OR CERTIFIES DEATH				24. DATE PRONOUNCED DEAD (Mo/Day/Yr)		25. TIME PRONOUNCED DEAD
26. SIGNATURE OF PERSON PRONOUNCING DEATH (Only when applicable)			27. LICENSE NUMBER		28. DATE SIGNED (Mo/Day/Yr)	
29. ACTUAL OR PRESUMED DATE OF DEATH (Mo/Day/Yr) (Spell Month)			30. ACTUAL OR PRESUMED TIME OF DEATH		31. WAS MEDICAL EXAMINER OR CORNER CONTACTED? <input type="checkbox"/> Yes <input type="checkbox"/> No	
<b>CAUSE OF DEATH (See instructions and examples)</b>						
32. PART I. Enter the chain of events—diseases, injuries, or complications—that directly caused the death. DO NOT enter terminal events such as cardiac arrest, respiratory arrest, or ventricular fibrillation without showing the etiology. DO NOT ABBREVIATE. Enter only one cause on a line. Add additional lines if necessary.						
IMMEDIATE CAUSE (Final disease or condition) resulting in death: a. _____ Due to (or as a consequence of): _____						
Sequentially list conditions, if any, leading to the cause listed on line a. Enter the UNDERLYING CAUSE (disease or injury that initiated the events resulting in death) LAST: c. _____ Due to (or as a consequence of): _____						
PART II. Enter other significant conditions contributing to death but not resulting in the underlying cause given in PART I						
35. DID TOBACCO USE CONTRIBUTE TO DEATH? <input type="checkbox"/> Yes <input type="checkbox"/> Probably <input type="checkbox"/> No <input type="checkbox"/> Unknown			36. IF FEMALE: <input type="checkbox"/> Not pregnant within past year <input type="checkbox"/> Pregnant at time of death <input type="checkbox"/> Not pregnant, but pregnant within 42 days of death <input type="checkbox"/> Not pregnant, but pregnant 43 days to 1 year before death <input type="checkbox"/> Unknown if pregnant within the past year		37. MANNER OF DEATH <input type="checkbox"/> Natural <input type="checkbox"/> Homicide <input type="checkbox"/> Accident <input type="checkbox"/> Pending Investigation <input type="checkbox"/> Suicide <input type="checkbox"/> Could not be determined	
38. DATE OF INJURY (Mo/Day/Yr) (Spell Month)		39. TIME OF INJURY		40. PLACE OF INJURY (e.g., Decedent's home; construction site; restaurant; wooded area)		41. INJURY AT WORK? <input type="checkbox"/> Yes <input type="checkbox"/> No
42. LOCATION OF INJURY: State: _____ City or Town: _____ Zip Code: _____						
43. DESCRIBE HOW INJURY OCCURRED:				44. IF TRANSPORTATION INJURY, SPECIFY: <input type="checkbox"/> Driver/Operator <input type="checkbox"/> Passenger <input type="checkbox"/> Pedestrian <input type="checkbox"/> Other (Specify):		
45. CERTIFIER (Check only one): <input type="checkbox"/> Certifying physician-To the best of my knowledge, death occurred due to the cause(s) and manner stated. <input type="checkbox"/> Pronouncing & Certifying physician-To the best of my knowledge, death occurred at the time, date, and place, and due to the cause(s) and manner stated. <input type="checkbox"/> Medical Examiner/Coroner-On the basis of examination, and/or investigation, in my opinion, death occurred at the time, date, and place, and due to the cause(s) and manner stated.						
Signature of certifier: _____						
46. NAME, ADDRESS, AND ZIP CODE OF PERSON COMPLETING CAUSE OF DEATH (Item 32)						
47. TITLE OF CERTIFIER		48. LICENSE NUMBER		49. DATE CERTIFIED (Mo/Day/Yr)		50. FOR REGISTRAR ONLY- DATE FILED (Mo/Day/Yr)
51. DECEDENT'S EDUCATION-Check the box that best describes the highest degree or level of school completed at the time of death. <input type="checkbox"/> 8th grade or less <input type="checkbox"/> 9th - 12th grade; no diploma <input type="checkbox"/> High school graduate or GED completed <input type="checkbox"/> Some college credit, but no degree <input type="checkbox"/> Associate degree (e.g., AA, AS) <input type="checkbox"/> Bachelor's degree (e.g., BA, AB, BS) <input type="checkbox"/> Master's degree (e.g., MA, MS, MEng, MEd, MSW, MBA) <input type="checkbox"/> Doctorate (e.g., PhD, EdD) or Professional degree (e.g., MD, DDS, DVM, LLB, JD)		52. DECEDENT OF HISPANIC ORIGIN? Check the box that best describes whether the decedent is Spanish/Hispanic/Latino. Check the "No" box if decedent is not Spanish/Hispanic/Latino. <input type="checkbox"/> No, not Spanish/Hispanic/Latino <input type="checkbox"/> Yes, Mexican, Mexican American, Chicano <input type="checkbox"/> Yes, Puerto Rican <input type="checkbox"/> Yes, Cuban <input type="checkbox"/> Yes, other Spanish/Hispanic/Latino (Specify) _____		53. DECEDENT'S RACE (Check one or more races to indicate what the decedent considered himself or herself to be) <input type="checkbox"/> White <input type="checkbox"/> Black or African American <input type="checkbox"/> American Indian or Alaska Native (Name of the enrolled or principal tribe) _____ <input type="checkbox"/> Asian Indian <input type="checkbox"/> Chinese <input type="checkbox"/> Filipino <input type="checkbox"/> Japanese <input type="checkbox"/> Korean <input type="checkbox"/> Vietnamese <input type="checkbox"/> Other Asian (Specify) _____ <input type="checkbox"/> Native Hawaiian <input type="checkbox"/> Guamanian or Chamorro <input type="checkbox"/> Samoan <input type="checkbox"/> Other Pacific Islander (Specify) _____ <input type="checkbox"/> Other (Specify) _____		
54. DECEDENT'S USUAL OCCUPATION (Indicate type of work done during most of working life. DO NOT USE RETIRED).						
55. KIND OF BUSINESS/INDUSTRY						

Fig. 1 (a) Sample death certificate

Examples of properly completed medical certifications

<p><b>CAUSE OF DEATH (See instructions and examples)</b></p> <p>32. PART I. Enter the <u>chain of events</u>—diseases, injuries, or complications—that directly caused the death. DO NOT enter terminal events such as cardiac arrest, respiratory arrest, or ventricular fibrillation without showing the etiology. DO NOT ABBREVIATE. Enter only one cause on a line. Add additional lines if necessary.</p> <p>IMMEDIATE CAUSE (Final disease or condition → resulting in death) a. <u>Rupture of myocardium</u> Due to (or as a consequence of)</p> <p>Sequentially list conditions, if any, leading to the cause listed on line a. Enter the UNDERLYING CAUSE (disease or injury that initiated the events resulting in death) LAST</p> <p>b. <u>Acute myocardial infarction</u> Due to (or as a consequence of)</p> <p>c. <u>Coronary artery thrombosis</u> Due to (or as a consequence of)</p> <p>d. <u>Atherosclerotic coronary artery disease</u></p>		<p>Approximate interval: Onset to death</p> <p><u>Minutes</u></p> <p><u>6 days</u></p> <p><u>5 years</u></p> <p><u>7 years</u></p>
<p>PART II. Enter other <u>significant conditions contributing to death</u> but not resulting in the underlying cause given in PART I</p> <p><u>Diabetes, Chronic obstructive pulmonary disease, smoking</u></p>		<p>33. WAS AN AUTOPSY PERFORMED? <input checked="" type="checkbox"/> Yes <input type="checkbox"/> No</p> <p>34. WERE AUTOPSY FINDINGS AVAILABLE TO COMPLETE THE CAUSE OF DEATH? <input checked="" type="checkbox"/> Yes <input type="checkbox"/> No</p>
<p>35. DID TOBACCO USE CONTRIBUTE TO DEATH? <input type="checkbox"/> Yes <input type="checkbox"/> Probably <input checked="" type="checkbox"/> No <input type="checkbox"/> Unknown</p>	<p>36. IF FEMALE:</p> <p><input type="checkbox"/> Not pregnant within past year</p> <p><input type="checkbox"/> Pregnant at time of death</p> <p><input type="checkbox"/> Not pregnant, but pregnant within 42 days of death</p> <p><input type="checkbox"/> Not pregnant, but pregnant 43 days to 1 year before death</p> <p><input type="checkbox"/> Unknown if pregnant within the past year</p>	<p>37. MANNER OF DEATH</p> <p><input checked="" type="checkbox"/> Natural <input type="checkbox"/> Homicide</p> <p><input type="checkbox"/> Accident <input type="checkbox"/> Pending Investigation</p> <p><input type="checkbox"/> Suicide <input type="checkbox"/> Could not be determined</p>
<p><b>CAUSE OF DEATH (See instructions and examples)</b></p> <p>32. PART I. Enter the <u>chain of events</u>—diseases, injuries, or complications—that directly caused the death. DO NOT enter terminal events such as cardiac arrest, respiratory arrest, or ventricular fibrillation without showing the etiology. DO NOT ABBREVIATE. Enter only one cause on a line. Add additional lines if necessary.</p> <p>IMMEDIATE CAUSE (Final disease or condition → resulting in death) a. <u>Aspiration pneumonia</u> Due to (or as a consequence of)</p> <p>Sequentially list conditions, if any, leading to the cause listed on line a. Enter the UNDERLYING CAUSE (disease or injury that initiated the events resulting in death) LAST</p> <p>b. <u>Complications of coma</u> Due to (or as a consequence of)</p> <p>c. <u>Blunt force injuries</u> Due to (or as a consequence of)</p> <p>d. <u>Motor vehicle accident</u></p>		<p>Approximate interval: Onset to death</p> <p><u>2 Days</u></p> <p><u>7 weeks</u></p> <p><u>7 weeks</u></p> <p><u>7 weeks</u></p>
<p>PART II. Enter other <u>significant conditions contributing to death</u> but not resulting in the underlying cause given in PART I</p>		<p>33. WAS AN AUTOPSY PERFORMED? <input checked="" type="checkbox"/> Yes <input type="checkbox"/> No</p> <p>34. WERE AUTOPSY FINDINGS AVAILABLE TO COMPLETE THE CAUSE OF DEATH? <input checked="" type="checkbox"/> Yes <input type="checkbox"/> No</p>
<p>35. DID TOBACCO USE CONTRIBUTE TO DEATH? <input type="checkbox"/> Yes <input type="checkbox"/> Probably <input checked="" type="checkbox"/> No <input type="checkbox"/> Unknown</p>	<p>36. IF FEMALE:</p> <p><input type="checkbox"/> Not pregnant within past year</p> <p><input type="checkbox"/> Pregnant at time of death</p> <p><input type="checkbox"/> Not pregnant, but pregnant within 42 days of death</p> <p><input type="checkbox"/> Not pregnant, but pregnant 43 days to 1 year before death</p> <p><input type="checkbox"/> Unknown if pregnant within the past year</p>	<p>37. MANNER OF DEATH</p> <p><input checked="" type="checkbox"/> Natural <input type="checkbox"/> Homicide</p> <p><input type="checkbox"/> Accident <input type="checkbox"/> Pending Investigation</p> <p><input type="checkbox"/> Suicide <input type="checkbox"/> Could not be determined</p>
<p>38. DATE OF INJURY (Mo/Day/Yr) (Spell Month)</p> <p>August 15, 2003</p>	<p>39. TIME OF INJURY</p> <p>Approx. 2320</p>	<p>40. PLACE OF INJURY (e.g., Decedent's home, construction site, restaurant, wooded area)</p> <p>roadside near state highway</p>
<p>42. LOCATION OF INJURY: State: Missouri City or Town: near Alexandria</p> <p>Street &amp; Number: mile marker 17 on state route 46A Apartment No.:</p>		<p>41. INJURY AT WORK? <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No</p>
<p>43. DESCRIBE HOW INJURY OCCURRED:</p> <p>Decedent driver of van, ran off road into tree</p>		<p>44. IF TRANSPORTATION INJURY, SPECIFY:</p> <p><input checked="" type="checkbox"/> Driver/Operator</p> <p><input type="checkbox"/> Passenger</p> <p><input type="checkbox"/> Pedestrian</p> <p><input type="checkbox"/> Other (Specify)</p>

Fig. 1 (b) Sample properly completed cause of death section on death certificate (From Heron et al., 2008)

made to the denominator. The denominator may also be formed in consideration of how long the population at risk was observed. This denominator is expressed in terms of person time-units and an example is computed below:

*Given a group of four persons observed for the following duration:*

- Person A is observed for 5 days
- Person B is observed for 1 day
- Person C is observed for 1 week
- Person D is observed for 1 month (assumes the average month is 30.5 days)
- Denominator = 43.5 person-days

The *factor of* ( $10^k$ ) used in the computation of a rate enables the relationship between the smaller number of events relative to a larger denominator to be more meaningful by converting the dividend to a whole number with one or two decimals to the right. The factor of 10 selected depends upon convention and the size of the

population from which the events were derived. A factor of 100,000 is used for computing rates for community populations. Generally, a power of 10 is selected such that the lowest level of the variable for which the rate is calculated is one digit to the right of the decimal.

The events for the numerator must come from the same *time period* in which the denominator was observed. Incidence rates can be computed for short time periods – hours, days, or weeks, as for epidemics of infectious disease – or over longer periods. The time period is by convention 1 year for common community-based events. If the number of events in a given year is low (<10), events and the denominator are averaged over a period of time, usually 3–5 years.

Data from Barker et al. (2006) illustrate the calculation of incidence rates for two population subgroups using a person-years denominator:

Incidence rate of heart failure (HF) per 1,000 person years (PY) of observstion for 1990–1994 by age group 70–74 years =

$$\frac{56 \text{ HF cases}_{70-74\text{y.o.}}}{6,641 \text{ person-years}_{70-74\text{y.o.}}} \times 1,000 = 8.4 \text{ HF cases per 1,000 PY}$$

Incidence rate of HF per 1,000 PY for age group 75–79 years =

$$\frac{48 \text{ HF cases}_{75-79 \text{ y.o.}}}{4,093 \text{ person-years}_{75-79\text{y.o.}}} \times 1,000 = 11.7 \text{ HF cases per 1,000 PY}$$

The same factor of 10 must be used when calculating rates for all levels within a variable.

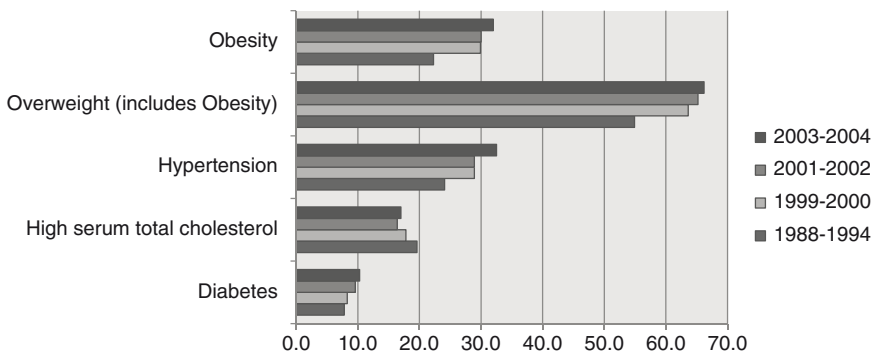
The use of the decimal in the representation of the rate value must also be consistent throughout the levels of the variables for which the rates are calculated.

## ***Prevalence Rate***

Prevalence rate is the total number of living persons (new or incident cases and preexisting cases) identified within a defined time period with the health condition divided by the average number of persons in the population or mid-year population in the same time period in which the total number of cases were selected. The prevalence rate, when referring to a disease in a specific period of time, is a measure of morbidity or disease burden in a population. The number of individuals with the health condition or factor examined who constitute the numerator also must be derived from the denominator and both should come from the same geographic area and in the same time period. Mathematically, this is represented as:

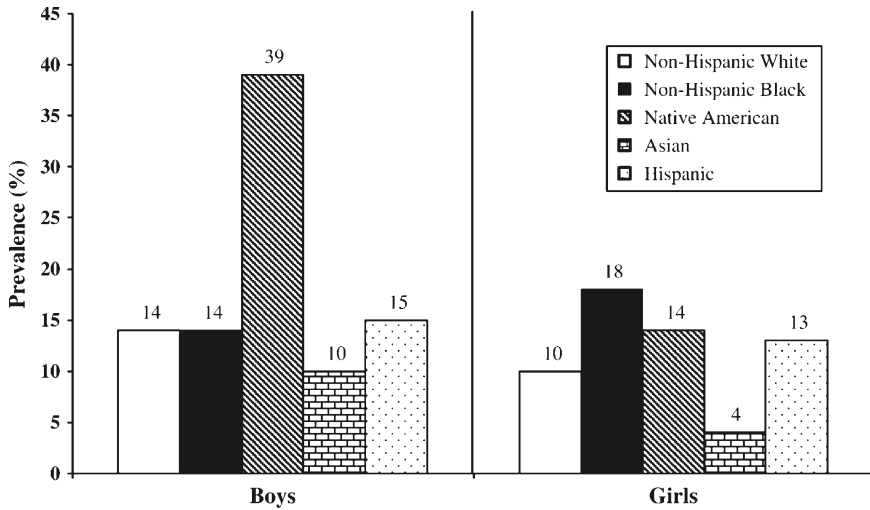
$$\text{Prevalence rate per } 10^k = \frac{\text{Total number of cases}_{(t)}}{\text{Total population at mid-year}_{(t)}} \times 10^k$$

The referent time ( $t$ ) can be short, one-point in time (point prevalence), or over time (period prevalence). The prevalence rate is influenced by how many individuals develop the condition in a particular time frame (incidence rate), how long the condition lasts (duration), and/or how long individuals with the condition survive. Factors that influence the development of an incident case (e.g., changes in exposures that result in disease, changes in the way individuals access health care, or changes in diagnostic methods) also affect the prevalence. The validity of the prevalence rate is also a function of the size of the population studied, the precision in establishing the boundaries (neighborhood, community, geopolitical area) of the population at risk (e.g., by verifying address where a person usually sleep at night in the case of the prevalence of IV drug users, residency in a school district in the case of autistic children), the thoroughness of efforts to ascertain all living cases of the conditions (e.g., clinics, schools, self-reports, field case-finding, etc.), and the standardization of the procedures or accuracy of testing to classify the case. A prevalence rate of exposure can also be computed. In this context, the number of persons in a population exposed to a particular risk factor in a given time period is the numerator and the criteria for the selection of the denominator follow as above. The factor of 10 in computing this prevalence rate is commonly 100 or 100%. Figure 2 displays trended prevalence rates expressed as per 100 persons for selected health conditions and risk factors. Prevalence rates are useful in guiding planning, program evaluation, and for resource allocation because they represent the number of cases (or “burden of disease”) relative to the number of persons with a particular condition in a geographic area. Figure 2 shows that the most dominant risk factor, and one that is rapidly increasing, among the vast majority of the US population is “overweight.” The high prevalence of obesity in Black females and Native American boys in Fig. 3 illustrates the urgent need for intervention also in these younger population subgroups. A decreasing prevalence of a disease may reflect either a decrease in exposure to a factor is decreasing or that control measures are increasing in effectiveness. Data from the National Health and Nutrition Examination



**Fig. 2** Prevalence rate per 100 persons, selected conditions and risk factors, persons aged 20+ years, civilian noninstitutionalized population, United States, 1988–2004 (National Center for Health Statistics, 2007)





**Fig. 3** Prevalence of obesity (body mass index  $\geq 30$  kg/m<sup>2</sup>) among US adolescents, by ethnicity, National Longitudinal Survey of Adolescent Health, 1994–1996 data, Wang, Y., et al. The obesity epidemic in the United States—gender, age, socioeconomic, racial/ethnic, and geographic characteristics: a systematic review and meta-regression analysis, *Epidemiol. Rev.* 2007;29:6–28, by permission of Oxford University Press

Survey 1999–2004 revealed that despite an increase in the prevalence of diabetes, the prevalence of the population glycemic (HbA<sub>1c</sub> <7.0%) and blood pressure (<130/80 mm Hg) control in the diabetic population has also increased (Ong et al., 2008). This illustrates how the prevalence rate of disease or exposure establishes a baseline level to better understand if greater awareness of a condition, availability of or access to health care services (e.g., introduction of insurance coverage for mammography), or changes in the criteria for diagnosing the condition influence the emergence of trends observed.

Prevalence rates may be difficult to compute directly, particularly for health conditions that may not come to diagnostic attention because of lack of culturally appropriate initiatives for prevention services and opportunities for testing, lack of access to treatment, or because local surveillance programs have not fully implemented confidential reporting. Because prevalence is in part determined by the incidence rate of the disease, any limitation associated with the accuracy of defining a case also affects the accuracy of the prevalence rate. For example, estimates of human immunodeficiency virus (HIV) infection prevalence may be underestimated as cases may be initially asymptomatic and are more likely to occur among those who do not routinely seek medical care (Campsmith et al., 2008).

Prevalence rates may also be difficult to determine for emerging conditions for which specific diagnostic tests do not exist. Such is the case with autism which can include a spectrum of disorders including deficits in communication and socialization, developmental disabilities, and unusual behaviors such as rituals or

preoccupation with parts of objects (Nicholas et al., 2008). In circumstances such as the psychosocial disorders, a prevalent case is determined by trained clinician diagnosticians, with expertise in the disorder under review descriptive data, as appropriate to the condition, from sources such as special education service evaluations, psychometric or neurobehavioral tests, description or observations of behaviors, and developmental histories.

## Formats of Rates

### *Crude Rate*

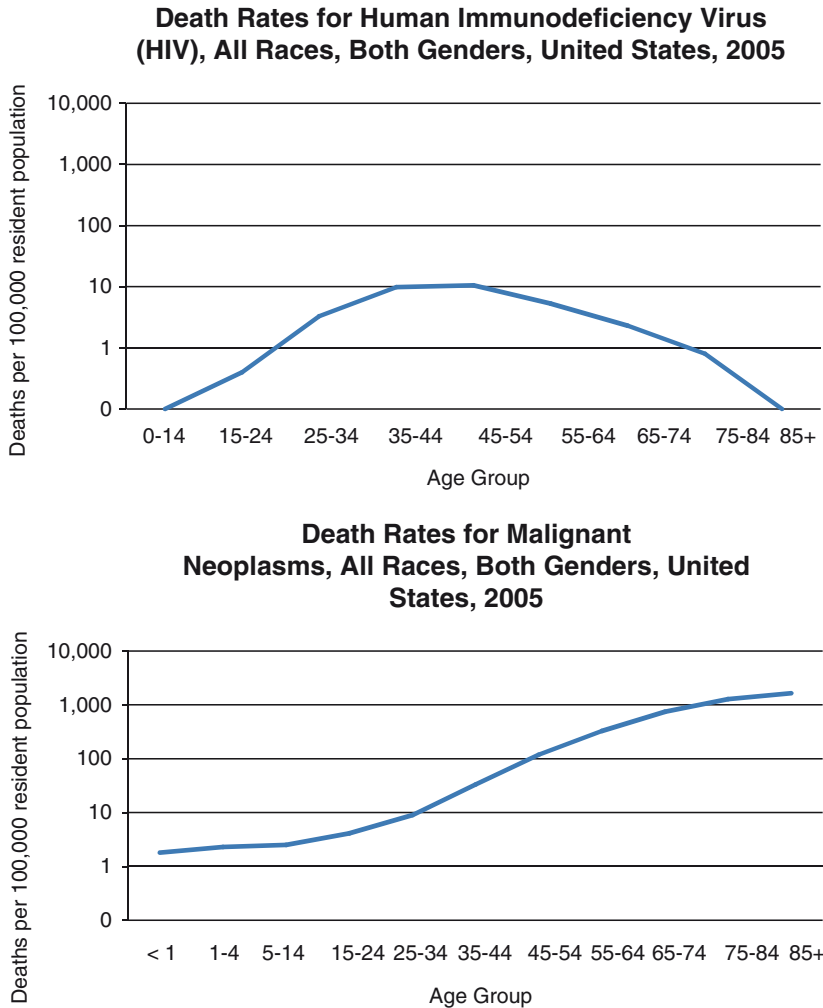
A *crude rate* is the total count of persons with the event or exposure in a population at risk for the event or exposure. This rate is not adjusted for any factors.

### *Specific Rate*

A *specific rate* refers to a count of persons with an event in a segment of a population in relation to the total number of persons in that same segment of the population. The specific rate is commonly restricted to subgroups of the population by age, race, sex, or causes of death. Commonly health problems have unique patterns of specific rates, at the very least, by age. Figure 2 displays age-specific mortality. For HIV, the rates increase during middle age and then declines (Fig. 4a); for malignant neoplasms, the rates exhibit a steep increase with age (Fig. 4b); for diseases of the heart, the rates exhibit a J-shaped curved (Fig. 4c); for motor vehicle accidents, the rates demonstrate bimodality with peaks in the 18–24-year age groups and then increasing over age 75 years (Fig. 4d).

### *Summary Rate Measures*

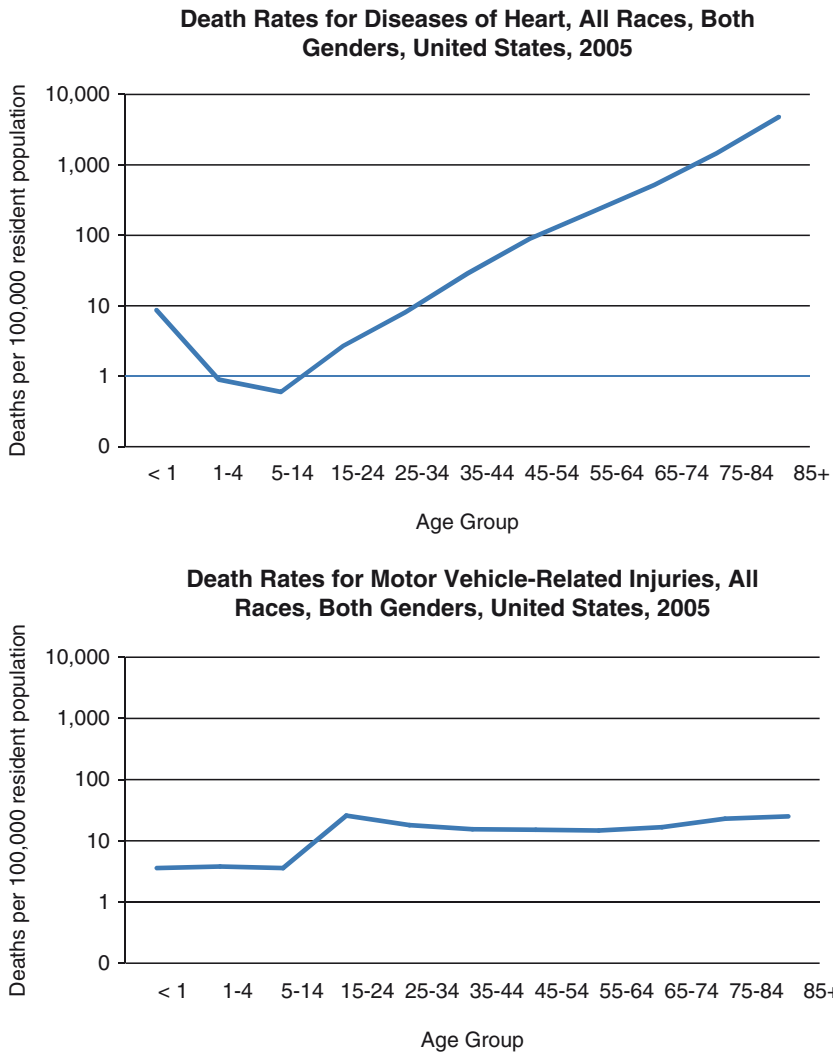
An *adjusted rate* is a summary descriptive epidemiologic measure which takes into account the variables that may emphasize or diminish the detection of the frequency of the health problem in a population examined. Adjusted rates are used to compare rates across time periods, geographic areas, or organizational units where differences in population characteristics are expected. The variables used in the adjustment are most commonly age, gender, and race. As illustrated in Fig. 2, age has a pronounced influence on the emergence and variation of health problems whether they are acute or chronic conditions or injuries. Age would, thus, be an important factor to account



**Fig. 4** (a) Age-specific mortality rates for human immunodeficiency virus (HIV) infection, all races, both genders, United States, 2005. (b) Age-specific mortality rates for malignant neoplasms, all races, both genders, United States, 2005

for when comparing annual injury rates between cities or states. Such comparisons can only be made when rates are adjusted for an important confounding factor. Adjustment to rates can be accomplished through arithmetic weighting procedures or statistical procedures.

The most common method for adjusting rates is the *direct standardization method*. The direct standardization method weights specific rates within strata (or groups) by the variable targeted to control for confounding, usually age. The weights are derived from the proportions of persons from a standard population or a standard



**Fig. 4** (c) Age-specific mortality rates for diseases of the heart, all races, both genders, United States, 2006. (d) Age-specific mortality rates for motor vehicle accidents, all races, both genders, United States, 2006 (National Center for Health Statistics, 2008)

million population in the corresponding stratum (or group) (*i*). Any standard population can be used for the weighting, US population (1970, 1990, 2000), a 1991 Canadian population, a European population, or a world population depending upon the comparison intended. The US 2000 population is currently commonly in use for national comparisons over time or across geographic area. The formula for computing a one-factor adjusted rate using the direct method is as follows:

$$\text{Adjusted rate} = \sum_{i=x}^y \left[ \left( \frac{\text{event counts}(i)}{\text{population}(i)} \right) \times 100,000 \times \left( \frac{\text{standard population}(i)}{\sum_{j=x}^y \text{standard population}(j)} \right) \right]$$

where *i* represents the level (or interval) of the variable targeted for adjustment. This formula is applied to the data Table 2. The stratum (age group)-specific rates of breast cancer mortality per 100,000 in the two race groups are given in this example. The breast cancer mortality rates in each age group (*i*) are then respectively weighted (i.e., multiplied) by the proportion of individuals in that age group by the US population in each age-group. The sums of these products (viz., columns “weighted rate...”) are summed for each race group yielding the final summary or adjusted (age-adjusted in this example) rate.

The standardized adjusted rate computed in Table 2 is interpreted to mean that the mortality from breast cancer is higher among Black women than White women, even adjusting for differences in the distribution of age between the two groups. If adjustment for more than one factor is desired, the reader should consult the methods in Fleiss et al. (2003). An adjusted rate is a summary rate and plotted as an individual data point may not allow easy visualization of trends.

The leading causes of death in the U.S. are displayed in Table 3. The crude death rate is used to rank order the causes of death as a means of prioritizing target areas for research, intervention, or policy formulation (see chapter “Epidemiology and the Public Policy Process”). The age-adjusted rates are used to calculate the percentage change for a particular cause over time.

In addition to this arithmetic method of adjusting rates, there are *statistical methods for adjusting rates* for comparing them across geographic areas and times including ordinary least squares regression and Poisson regression models. Ordinary least squares and Poisson regression models are used in small area analysis; it is important to adjust for a wide variety of variables including age, sex, race, primary chronic condition, and multiple chronic conditions which could influence the utilization of health care (Wennberg et al., 2008). Joinpoint regression (National Cancer Institute, 2008) model is used to fit a series of joined straight lines on a log scale to estimate trends in adjusted rates to examine temporal trends. The resultant values from the regression are represented as annual percentage changes which equal to  $100 \times (e^b - 1)$ , where *b* is the slope of the regression line and the significance is evaluated using the appropriate statistical test, two-tailed.

*Years of potential life lost (YPLL)* is a summary measure of premature mortality. To compute the YPLL, the number of deaths for each age group, with eight groups (under 1 year, 1–14, 15–24, 25–34, 45–54, 55–64, 65–74 years) being the convention, is tallied. The midpoint for each group is determined (e.g., 0.5, 7.5, 19.5, 29.5, 39.5, 49.5, 59.5, 69.5) and subtracted from the age that they would have expected to live, which then represents the number of years of life lost (YLL). The average life expectancy in the USA is 75 years and the mid-point for each age group is subtracted from 75 years. For example, if a death from homicide occurred in the 15–24 years age group, 55.5 years of life would be considered lost (75 years – 19.5 (midpoint

**Table 2** Calculation of age-adjusted mortality rate for breast cancer using the direct method

Age group (years)	White (W) female breast cancer mortality rate per 100,000 <sup>a</sup> <i>r</i> W	Black (B) female breast cancer mortality rate per 100,000 <sup>a</sup> <i>r</i> B	2000 US standard population by age group <sup>b</sup> <i>n</i>	Proportion of 2000 US standard population by age group <i>p</i>	Weighted rate white females <i>p</i> × <i>r</i> W	Weighted rate black females <i>p</i> × <i>r</i> B
Under 5 years	0	0	18,986,520	0.069134	0	0
5–9 years	0	0	19,919,840	0.072532	0	0
10–14 years	0	0	20,056,779	0.073031	0	0
15–19 years	0	0	19,819,518	0.072167	0	0
20–24 years	0.1	0.1	18,257,225	0.066478	0.006648	0.006648
25–29 years	0.6	1.8	17,722,067	0.06453	0.038718	0.116154
30–34 years	2.7	6.2	19,511,370	0.071045	0.191822	0.440479
35–39 years	6.9	15.1	22,179,956	0.080762	0.557258	1.219506
40–44 years	12.9	25.6	22,479,229	0.081852	1.055891	2.095411
45–49 years	20.9	39.5	19,805,793	0.072117	1.507245	2.848622
50–54 years	33.7	58.1	17,224,359	0.062718	2.113597	3.643916
55–59 years	48.9	72.4	13,307,234	0.048454	2.369401	3.508070
60–64 years	61.1	85.8	10,654,272	0.038794	2.370313	3.328525
65–69 years	75.9	88.1	9,409,940	0.034264	2.600638	3.018658
70–74 years	89.1	103.6	8,725,574	0.031772	2.830885	3.291579
75–79 years	110.8	121.4	7,414,559	0.026998	2.991378	3.277557
80–84 years	133.8	135.6	4,900,234	0.017843	2.387393	2.419511
85+ years	179.2	196.5	4,259,173	0.015509	2.779213	3.047519
Totals			274,633,642	1.000000	23.8004	32.26215

Age-adjusted breast cancer mortality rate per 100,000 population =

$$\sum (p \times r_i) \times 100,000 =$$

= 23.8 breast cancer deaths among White females per 100,000 population

= 32.3 breast cancer deaths among Black females per 100,000 population

<sup>a</sup>Source: Surveillance, Epidemiology, and End Results Program of the National Cancer Institute (2004)

<sup>b</sup>Source: U.S. Bureau of the Census, Census of Population (2000)

**Table 3** Deaths and death rates for 2006 and age-adjusted death rates and percentage changes in age-adjusted rates from 2005 to 2006 for the 15 leading causes of death: United States, final 2005 and preliminary 2006 (Heron et al., 2008)

Rank <sup>a</sup>	Cause of death	ICD Code on the International Classification of Diseases, Tenth Revision, Second Edition, 2004	Number	Crude death rate	Age-adjusted death rate		
					2006	2005	Percent change
...	All causes		2,425,901	810.3	776.4	798.8	-2.8
1	Diseases of heart...	(I00-I09,I11,I13,I20-I51)	629,191	210.2	199.4	211.1	-5.5
2	Malignant neoplasms...	(C00-C97)	560,102	187.1	180.8	183.8	-1.6
3	Cerebrovascular diseases...	(I60-I69)	137,265	45.8	43.6	46.6	-6.4
4	Chronic lower respiratory disease	(J40-J47)	124,614	41.6	40.4	43.2	-6.5
5	Accidents (unintentional injuries)	(V07-X59,Y85-Y86)2	117,748	39.3	38.5	39.1	-1.5
6	Alzheimer's disease...	(G30)	72,914	24.4	22.7	22.9	-0.9
7	Diabetes mellitus...	(E10-E14)	72,507	24.2	23.3	24.6	-5.3
8	Influenza and pneumonia...	(J10-J18)	56,247	18.8	17.7	20.3	-12.8
9	Nephritis, nephrotic syndrome and nephrosis...	(N00-N07,N17-N19,N25-N27)	44,791	15	14.3	14.3	0.0
10	Septicemia...	(A40-A41)	34,031	11.4	10.9	11.2	-2.7
11	Intentional self-harm (suicide)	(U03,X60-X84,Y87.0)	32,185	10.7	10.6	10.9	-2.8
12	Chronic liver disease and cirrhosis	(K70,K73-K74)	27,299	9.1	8.7	9	-3.3
13	Essential hypertension and hypertensive renal disease	(I10,I12,I15) <sup>b</sup>	23,985	8	7.6	8	-5
14	Parkinson's disease...	(G20-G21)	19,660	6.6	6.3	6.4	-1.6
15	Assault (homicide)...	(*U01-U02,X85-Y09,Y87.1)	18,029	6	6	6.1	-1.6
...	All other causes...	(Residual codes)	455,333	152.1	...	...	...

<sup>a</sup>Rank based on number of deaths

<sup>b</sup>Cause-of-death title has been changed in 2006 to reflect the addition of Secondary Hypertension (ICD-10 code I15)

<sup>c</sup>Beginning 2001 includes acts of terrorism

value) years = 55.5 YLL). The number of deaths in an age group is then multiplied by the number of YLL in that age group. The summary YPLL is then derived by summing the products of the YLL multiplied by the number of deaths in each age group over all the age groups. The YPLL can be calculated for all causes of death combined or for specific causes of death. The YPLL may be represented as a rate per 100,000 population under 75 years. The YPLL is a useful summary measure when examining or prioritizing health problems in communities that disproportionately affect younger persons or for assessing the impact of community-based interventions in geographic areas with a large minority or immigrant population which are typically younger than the population of the community as a whole. Table 4 displays the age-adjusted YPLL of major causes of death in the USA for 2000 and 2005. Note that mortality from unintentional injuries is ranked higher as a health concern using the YPLL method than its ranking computed using direct standardization. Additionally, the direction of the trends is different with the direct standardization method showing that unintentional injuries are declining whereas the YPLL method identifies an increasing trend in these categories of deaths. The YPLL rankings in Table 3 indicate that additional policies and/or public and community interventions are warranted aimed at violence prevention and injury control, particularly in the types of

**Table 4** Percentage change in age-adjusted<sup>a</sup> years of potential life lost before age 75 (YPLL-75) for selected causes of death, all persons, US, 2000 and 2005 (Heron et al., 2008)

Cause of death	YPLL(75) per 100,000 population under 75 years of age		Percentage change 2000–2005
	2000	2005	
All causes	7,578.1	7,299.8	-3.67
Diseases of heart	1,253.0	1,110.4	-11.38
Ischemic heart disease	841.8	701.8	-16.63
Cerebrovascular diseases	223.3	193.3	-13.43
Malignant neoplasms	1,674.1	1,525.2	-8.89
Trachea, bronchus, and lung	443.1	392.9	-11.33
Colorectal	141.9	124.7	-12.12
Prostate	63.6	55.1	-13.36
Breast	332.6	296.2	-10.94
Chronic lower respiratory diseases	188.1	181.2	-3.67
Influenza and pneumonia	87.1	83.6	-4.02
Chronic liver disease and cirrhosis	164.1	152.6	-7.01
Diabetes mellitus	178.4	179.9	-0.84
Human immunodeficiency virus (HIV) disease	174.6	133.6	-23.48
Unintentional injuries	1,026.5	1,132.7	10.35
Motor vehicle-related injuries	574.3	564.4	-1.72
Suicide	334.5	347.3	3.83
Homicide	266.5	276.8	3.86

<sup>a</sup>Age-adjusted to the year 2000 standard population



communities previously described. The concept of premature mortality or YLL can be extended to include years lost due to a state less than full health (termed “disability”) (YLD). The YLL and YLD when summed equal a Disability Adjusted Life Year (DALY). The DALY is a composite summary measure which is the sum of the years of life lost due to premature mortality (YLL) and the years lost due to disability for incident cases of a health condition. One DALY represents the loss of 1 year of full health. The DALY then is used to represent a “burden of disease” in a population which is a particularly good measure for health conditions which have multi-dimensional impairments (e.g., limitations of physical functioning, social functioning impairments, etc.) associated with them, such as mental illness (Ustun, 1999).

*Life expectancy* is the average number of years remaining to a person at a particular age who was born at a particular time. The remaining years of life are based upon the age-specific death rates in effect at the time the measure was constructed. The death rates are applied to survivors in each age or age group to obtain a probability of dying with the resultant tabulations presented as a life table from which the life expectancy measure is calculated. Thus, life expectancy is a measure of the mortality pattern across all age groups, infants, children, adults, and the elderly. The primary drivers of the summary life expectancy values are the magnitude of rates occurring early in life. Populations with high infant mortality and/or high mortality during adolescent or young adult years will have markedly lower life expectancies. Life expectancy is a proxy summary measure of overall mortality level in a population or a population subgroup as well as is a proxy measure of the quality of life of a nation. Table 4 displays the life expectancies for various member countries of the World Health Organization according to country income level.

## Ratio Measures

The *risk ratio* compares two rate measures of risk (e.g., incidence or mortality rates). Generally, the group with the higher risk (rate) for the health problem evaluated or the group in which exposure to the risk factor is greatest is placed in the numerator. The rate of the lower-risk group or the group with less or no exposure is in the denominator. The rate in the numerator and the rate in the denominator typically will be an adjusted rate so that the effect of the major confounding variable is eliminated and the two populations can be appropriately compared. The male–female risk ratio for mortality due to HIV infection is computed using the age-adjusted death rates from Table 5:

$$\begin{aligned}
 \text{Risk ratio} &= \frac{\text{rate in the higher risk group or exposed group}}{\text{rate in the lower risk or non-exposed group}} \\
 &= \frac{19.4 \text{ age-adjusted HIV death rate per } 100,000 \text{ among blacks}}{2.2 \text{ age-adjusted HIV death rate per } 100,000 \text{ among whites}} \\
 &= 8.82
 \end{aligned}$$

**Table 5** Life expectancy and selected morbidity measures and risk factors for World Health Organization (WHO) member nations by region and income group (WHO, 2008)

WHO group	Life expectancy <sup>a</sup> at birth (years) both sexes	Prevalence of tuberculosis <sup>a</sup> (per 100,000 population)	Prevalence of HIV among adults aged ≥15 years (per 100,000 population) <sup>b</sup>	Infant mortality rate per 1000 live births, both sexes <sup>c</sup>	Low birth weight newborns (%) <sup>d</sup>	Access to improved sanitation (%) <sup>e</sup>
Region						
African region	51	547	4,459	84	14	33
Region of the Americas	75	44	442	18	9	87
South-East Asia Region	64	289	289	52	26	37
European Region	74	54	342	14	8	92
Eastern Mediterranean Region	64	152	192	62	17	60
Western Pacific Region	74	199	89	20	8	63
Income group						
Low income	59	362	1,039	73	22	78
Lower middle income	71	188	239	27	9	88
Upper middle income	69	121	1,484	22	10	95
High income	80	17	249	6	7	100
Global	67	219	662	49	16	60

<sup>a</sup>2006

<sup>b</sup>2005

<sup>c</sup>2007

<sup>d</sup>2000–2002

<sup>e</sup>2006

See World Health Organization (2008) for list of countries within region and income groupings

The computed risk ratio means that blacks are 8.82 times more likely to die from HIV infection than whites. Since this ratio compares two age-adjusted rates, the risk ratio is not due to differences in the age distribution between blacks and white. A risk ratio of twofold or greater for a health problem between the two groups compared indicates disparity of particular concern. Rank-ordered, a twofold greater risk of death is noted as follows: in blacks relative to whites for HIV (8.82), homicide (5.70), and diabetes (2.08); and for males relative to females for suicide (4.09), homicide (3.84), HIV (2.70), motor vehicle accidents (2.44), and chronic liver disease (2.14) (Table 6).

The *relative risk (RR)* is computed from a cross-tabulated format. In Table 8 the members of the sample studied are cross classified into exposure-outcome cells whereby the rows represent exposure levels, usually a nominal level of measurement, exposed vs. not exposed, and columns represent the outcome, usually also a nominal level of measurement, outcome present vs. absent.

The true measures of risk are typically only derived from studies with a longitudinal component or where the population at risk is known. The RR can be estimated from an *odds ratio (OR)*. When estimating the RR, the OR is calculated as the ratio of the odds of an event occurring given exposure to the odds of an event when exposure is absent. The RR and the OR computed from the same sample are approximately equal when the prevalence of the event is low. Note that in Table 9 the computed RR and the OR are in the same positive direction and are approximately equal.

The RR and the OR are both descriptive ratio measures used to appraise the strength of an association between exposure and an event. As ratio measures, the valid values of the RR and the OR range from 0 to  $+\infty$ . An RR or an OR of 1 means that the chances of the event (or exposure) are the same between the two groups compared, the exposed and the nonexposed.

**Table 6** Age-adjusted death rates per 100,000 population for selected causes of death, by gender and race, United States, 2005 (Heron et al., 2008)

Cause of death	Age-adjusted death rate per 100,000 population			
	Males	Females	Blacks	Whites
All causes	951.1	677.6	1,016.5	785.3
Diseases of heart	260.9	172.3	271.3	207.8
Malignant neoplasms	225.1	155.6	222.7	182.6
Cerebrovascular disease	46.9	45.6	65.2	44.7
Chronic lower respiratory disease	51.2	38.1	30.6	23.4
Influenza and pneumonia	23.9	17.9	21.7	20.2
Chronic liver disease and cirrhosis	12.4	5.8	7.7	9.2
Diabetes mellitus	28.4	21.6	46.9	22.5
Human immunodeficiency virus (HIV) disease	6.2	2.3	19.4	2.2
Motor vehicle-related injuries	21.7	8.9	14.5	15.6
Suicide	18.0	4.4	5.2	12.0
Homicide	9.6	2.5	21.1	3.7

**Table 7** Risk ratios of age-adjusted rates per 100,000 population for selected causes of death by gender and race, United States, 2005

Cause of death	Risk ratio of	
	Male to female	Black to white
All causes	1.40	1.29
Diseases of heart	1.51	1.31
Malignant neoplasms	1.45	1.22
Cerebrovascular disease	1.03	1.46
Chronic lower respiratory disease	1.34	0.67
Influenza and pneumonia	1.34	1.07
Chronic liver disease and cirrhosis	2.14	0.84
Diabetes mellitus	1.31	2.08
Human immunodeficiency virus (HIV) disease	2.70	8.82
Motor vehicle-related injuries	2.44	0.93
Suicide	4.09	0.43
Homicide	3.84	5.70

**Table 8** Formulae for computation of relative risk and odds ratio

	Event		Total
	Yes	No	
Exposure			
Yes	a	b	a + b
No	c	d	c + d

Relative risk of event = Incidence rate in high risk group/Incidence rate in low risk group  
 $= [a/(a+b)]/[c/(c+d)]$   
 Odds ratio of exposure yielding an event =  $(a/b)/(c/d)$

**Table 9** Relative risks and odds ratios of medical error from shifts (Rogers et al., 2004)

Average hours worked on scheduled day	No. of shifts	No. of errors (%)	Odds ratio	Relative risk
<8.5	771	12 (1.6)	1.00 (referent)	1.00 (referent)
8.5–12.5	2,484	77 (3.1)	2.02	1.99
12.5 or more	2,057	103 (5.0)	3.33	3.22
Total	5,312	192 (3.5)	–	–

Using formula from Table 8 to compute odds ratio and relative risk of medical error:

Average hours worked	Error		Total
	Yes	No	
8.5–12.5 h	77(a)	2,407 (b)	2,484 (a + b)
<8.5	12 (c)	759 (d)	771 (c + d)

Odds ratio =  $(a/b)/(c/d) = (77/2,407)/(12/759) = 2.02$

Those odds of those 8.5–12.5 h reporting a medical error are 2.02 times those working less than an 8.5 h day

Relative risk =  $(77/2,484)/(12/759) = 1.99$

The risk of medical error is 1.99 times greater in those working a 8.5–12.5 h than those working less than 8.5 h

Because the RR and the OR are estimates of a true value calculated from samples, a confidence interval, usually 95% is computed to assess the reliability of the sample estimates. A *confidence interval* is an interval that is generated from a random sample of an underlying population such that if the sampling was repeated numerous times and the parameter was calculated from each sample according to the same method the samples would contain some interval probability of values in which the true parameter (in this example the RR or the OR) would lie.

The *population attributable risk* (PAR) utilizes ratios to estimate the proportion of the risk of exposure due to a particular factor. This measure has importance as it assesses the theoretically achievable reduction in risk if the risk factor were entirely removed from a population. It is for this reason that the PAR is a useful measure of the efficacy of population-based preventive interventions when multiple risk factors are involved in the emergence of a disease as is the case of chronic conditions. The calculation of the PAR is illustrated in Table 10. The PAR can also be estimated from the OR. Thus, when resources are limited, the PAR can aid in determining which interventions may have the greatest impact upon a population. The PAR of 56% means that 56% of the risk of diabetes is due to weight gain. This indicates that the greatest opportunity for preventing diabetes mellitus is to control weight gain to no more than 7 kg (15.4 lbs) after age 21 years. However, 44% of the risk of diabetes is due to other factor(s). Caution in the interpretation of the PAR is advised as different risk factors may be subject to different levels of bias in the classification of the different exposures evaluated as well as there may be a number of individuals exposed who may not respond to the intervention proposed. Also, the calculated PARs presented in Table 10 are “unadjusted.” Methodologies exist for examining combinations of risk factors and the impact of interpreting a PAR when selectively removing exposure to a risk factor while leaving all other exposures unchanged (Walter, 1998).

**Table 10** Computation of population attributable risk (Data from: Koh-Banerjee et al., 2004)

Subgroup	% of men	RR <sub>DM</sub>	PAR (%)
Weight gain of ≥7 kg since age 21 years	55	3.3	56
Waist gain of ≥2.5 cm over 9 years	56	1.4	20

RR<sub>DM</sub> = relative risk of diabetes mellitus

$$\text{Population Attributable Risk (PAR\%)} = \frac{Pe(RR - 1)}{1 + Pe(RR - 1)} \times 100\%$$

where Pe = proportion controls exposed; RR = relative risk for exposed compared with unexposed where the risk among unexposed is 1

Thus,

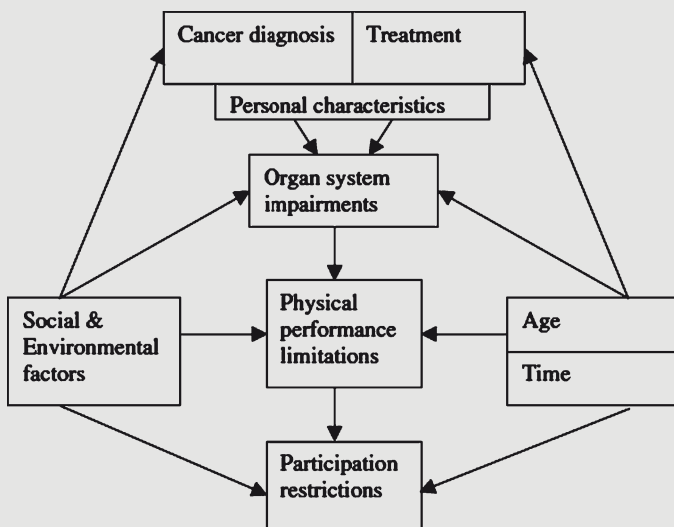
$$PAR_{\text{weight gain} \geq 7 \text{ kg}} = \frac{0.55(3.3 - 1)}{1 + 0.55(3.3 - 1)} \times 100\% = 56\%$$

## Summary

This chapter presented common descriptive epidemiologic measures useful in planning health care services, and for evaluating policies and programs. Rate measures prioritize health problems within a population or across time or geographic areas. They also can be used to evaluate trends within population subgroups. Summary descriptive measures provide a quick assessment of the health of populations.

### **Boxed Example 1 Consequences of Long-Term Survival from Cancer**

**Problem:** There is an increasing prevalence of cancer survivors because of increases in the early detection of cancer and advances in effective treatment. Cancer therapy is often intensive (invasive surgery, high-dose chemotherapy and radiation therapy, multimodality therapy) and potentially results in long-term psychological and physical late effects and even damage to normal tissue affecting physical performance which results in participation restrictions in activities of daily living (see Conceptual Model).



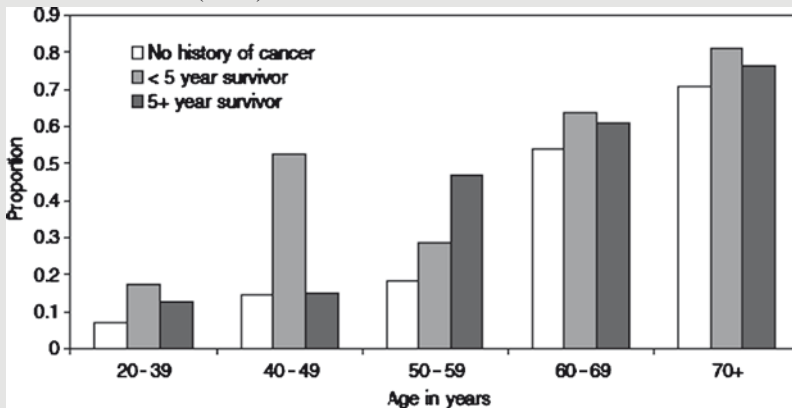
Reprinted from *Ann. Epidemiol.* Vol. 16. Ness, K.K. et al. 197–205. Copyright 2006 with permission from Elsevier.

**Method and Data:** Data from the National Health and Nutrition Examination Survey (NHANES 1999–2002) were used to analyze the information for those who were noninstitutionalized persons age 20+ years at the time of the home interview. Complex multistage samples design was used. Physical performance limitation was ascertained if the responded states “some difficulty,” “much difficulty,” or “unable to do” one or more of 11 physical performance activities

(e.g., getting out of bed, stooping, etc.), without special equipment or one or more of seven routine activities (e.g., dressing, preparing meals, etc.) or a positive response to having “difficulty walking without using any special equipment.”

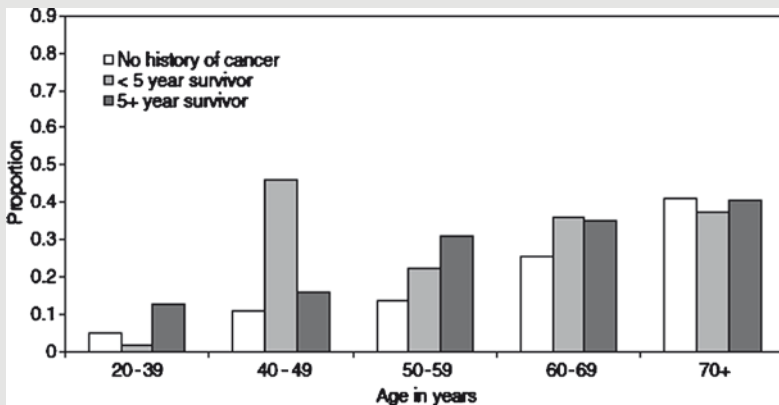
**Results:** The overall prevalence of cancer history was 8.7%. The prevalence of any restriction in routine activities was: 13.0% for those without cancer, 30.5% for those <5 years post cancer diagnosis, 31.3% for those 5+ years cancer survivor. The prevalence of any performance limitation was: 21.2% for those without cancer, 52.7% for those <5 years post cancer diagnosis, 54.4% for those 5+ years cancer survivor.

Source: Ness et al. (2006).



Proportion of adult participants in NHANES 1999–2002 with physical performance limitations.

Reprinted from Ann. Epidemiol., Vol.16, Ness, K.K., et al., 197–205, Copyright 2006, with permission from Elsevier.



Proportion of those with participation restrictions reported among adult participants in NHANES 1999–2002. Reprinted from Ann. Epidemiol., Vol.16, Ness, K.K., et al., 197–205, Copyright 2006, with permission from Elsevier.

**Boxed Example 1** (continued)

Managerial Epidemiology Interpretation: Prevalence measures are estimates of the burden of disease and are the most useful in planning for health services utilization. Both new and long-term cancer survivors experience physical performance limitations and participation restrictions at a greater rate than the general population. Rehabilitation intervention offered soon after active cancer therapy is completed may benefit those particularly for those under 60 years of age as a means of preventing and mitigating long-term disability. The manager may also need to re-formulate the conceptual model and consider other variables which could affect long-term disability. Additional descriptive epidemiologic data may be necessary for the under 60 year age group.

**Boxed Example 2 Are There Black-White Disparities in Breast Cancer Mortality Rates?**

**Problem:** The rate of decline in breast cancer mortality has been declining, but at different rates among the races and geographic areas. Ohio has one of the highest breast cancer mortality rates in the nation.

**Methods:** Breast cancer mortality data (deaths with ICD-9 code of 174) and population data were obtained from the National Center for Health Statistics via the Surveillance, Epidemiology, and End Results (SEER) program software for the period 1970–2001. Data were age-adjusted to the 2000 US standard population. The Breslow-Day statistic was used to evaluate the significance between the rates.

**Data and Results:** Breast Cancer Mortality in Ohio 1970–1974 and 1996–2001 by Race, Women, 30–74 Years

Race	Difference (%) Black			Difference (%) Blacks			Change	
	Rate <sup>a</sup> 1970– 1974	vs. Whites in 1970– 1974	p-value for difference	Rate <sup>a</sup> 1997– 2001	vs. Whites in 1997– 2001	p-value for difference	(%) 1970– 1974 vs. 1997–2001	p-value for difference
Whites	47.7	-0.8	0.18	35.4	41.8	<0.001	-25.8	<0.001
Blacks	47.3			50.2			6.1	0.17

<sup>a</sup>Age-adjusted standardized rates per 100,000 using US 2000 standard population

Source: Tyczynski et al. (2006).

Calculation of percentage difference is as follows:

$$\text{Difference (\%)} = \frac{(\text{Rate comparison group} - \text{rate in referent group})}{\text{Rate referent group}} \times 100\% = \frac{(50.2 - 35.4)}{35.4} \times 100\% = 41.8\%$$

Managerial Epidemiology Interpretation: White women exhibit a favorable rate of decline in breast cancer mortality whereas among Black women the rate increased but was not significantly different from 30 years ago. Reasons for



**Boxed Example 2** (continued)

the high rates in Black females include an increased incidence of breast cancer, a potentially increasing higher tumor grade of the cancer, and a lack of access to health care for screening, initial treatment and follow-up care due to lack of insurance and/or availability of screening services and health care providers. Health care managers should give priority attention to ensuring that follow-up of any abnormal screening mammography test is done particularly quickly, especially for population subgroups at high risk for breast cancer mortality.

Ratios aid in evaluating the degree of the relationship between exposure and outcome. Descriptive measures, in general, are useful in identifying the components of the health care system that could be modified to by identifying populations at risk because of their demographic composition, exposure, or geographic location. The type of epidemiologic measure used depends upon the objective of the assessment, the nature of the health problem being evaluated, and the type of data available for the assessment.

**Discussion Questions**

- Q.1. Compare the rankings of the causes of mortality in Tables 4 and 6. What accounts for the differences in ranking of causes of mortality? Which table of data would you use in determining what should be priority areas for health care services in a community comprised of immigrants and racial or ethnic minorities?
- Q.2. How can mortality data be used in prioritizing which level health care services would be appropriate for a community (see also Tables 3, 5–6, 10, and Fig. 4)? What are the limitations of mortality data in prioritizing health problems for which to develop health care services in a community?
- Q.3. Data from the National Health and Nutrition Examination Survey 1999–2004 found the prevalence of diagnosed diabetes to significantly increase from 5.1% in the period 1988–1994 to 6.5% in the period 1999–2002. During these same time periods undiagnosed diabetes and impaired fasting glucose remained unchanged (Ong et al., 2008). What factors could have accounted for the overall increase in prevalence in diagnosed diabetes?

**References**

- Barker, W.H., Mullooly, J.P., and Getchell, W., 2006, Changing incidence and survival for heart failure in a well-defined older population, 1970–1974 and 1990–1994, *Circulation* 113:799–805.

- Campsmith, M.L., Rhodes, P., Hall, H.I., and Green, T., 2008, HIV prevalence estimates—United States, 2006, *M.M.W.R.* **57**:1073–1076.
- Fleiss, J.L., Levin, B., and Park, M.C., 2003, *Statistical Methods for Rates and Proportions*, 3rd ed., Wiley, Hoboken, NJ.
- Heron, M., Hoyert, D.L., Xu, J., Scott, C., and Tejada, B., 2008, Deaths: Preliminary Data for 2006, *National Vital Statistics Reports* vol. **56**, no. 16. National Center for Health Statistics, Hyattsville, MD.
- Hickman, M., Seaman, S., and de Angelis, D., 2001, Estimating the relative incidence of heroin use: application of a method for adjusting observed reports of first visits to specialized drug treatment agencies, *Am. J. Epidemiol.* **153**:632–41.
- Koh-Banerjee, P., Wang, Y., Hu, F.B., Spiegelman, D., Willett, W.C., and Rimm, E.B., 2004, Changes in body weight and body fat distribution as risk factors for clinical diabetes in US men, *Am. J. Epidemiol.* **159**:1150–1159.
- National Cancer Institute, 2008, *Joinpoint Regression Program, Version 3.3* (August 16, 2008), <http://srab.cancer.gov/joinpoint>.
- National Center for Health Statistics, 2007, Health, United States, 2007, With Chartbook on Trends in the Health of Americans, Hyattsville, MD.
- Ness, K.K., Wall, M.M., Oakes, J.M., Robison, L.L., and Gurney, J.G., 2006, Physical performance limitations and participation restrictions among cancer survivors: a population-based study, *Ann. Epidemiol.* **16**:197–205.
- Nicholas, J.A., Charles, J.M., Carpenter, L.A., King, L.B., Jenner, W., and Spratt, E.G., 2008, Prevalence and characteristics of children with autism-spectrum disorders, *Ann. Epidemiol.* **18**:130–136.
- Ong, K.L., Cheung, B.M.Y., Wong, L.Y.F., Wat, N.M.S., Tan, K.C.B., and Lam, K.S.L. 2008, Prevalence, treatment, and control of diagnosed diabetes in the U.S. National Health and Nutrition Examination Survey 1999–2004, *Ann. Epidemiol.* **18**:222–229.
- Rogers, A.E., Hwang, W., Scott, L.D., Aiken, L.H., and Dinges, D.F., 2004, The working hours of hospital staff nurses and patient safety, *Health Affairs***23**:202–212.
- Sanfilippo, F.M., Hobbs, M.S.T., Knuiman, M.W., and Hung, J., 2008, Impact of new biomarkers of myocardial damage on trends in myocardial infarction hospital admission rates from population-based administrative data, *Am. J. Epidemiol.* **168**:225–233.
- Shapiro-Mendoza, C.K., Tomashek, K.M., Anderson, R.N., and Wingo, J., 2006, Recent national trends in sudden, unexpected infant deaths: more evidence supporting a change in classification or reporting, *Am. J. Epidemiol.* **163**:762–769.
- Tyczynski, J.E., Hill, T.D., and Berkel, H.J., 2006, Why do postmenopausal African-American women not benefit from overall breast cancer mortality decline? *Ann. Epidemiol.* **16**:180–190.
- U.S. Department of Health and Human Services (USDHHS), Centers for Disease Control and Prevention, 2007, *Health, United States, 2007: with Chartbook on Trends in the Health of Americans* (August 16, 2008), <http://www.cdc.gov/nchs/data/abus/abus07.pdf>.
- Ustun, T.B., 1999, The global burden of mental disorders, *Am. J. Pub Health* **89**:1315–1318.
- Walter, S.D., 1998, Attributable risk in practice, *Am. J. Epidemiol.* **148**:411–413.
- Wang, Y., and Beydoun, M.A., 2007, The obesity epidemic in the United States—gender, age, socioeconomic, racial/ethnic, and geographic characteristics: a systematic review and meta-regression analysis, *Epidemiol. Rev.* **29**:6–28.
- Wennberg, J.E., Fisher, E.S., Goodman, D.C., and Skinner, J.S., 2008, *Tracking the Care of Patients with Severe Chronic Illness: The Dartmouth Atlas of Health Care 2008*, The Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth, NH.
- World Health Organization (WHO), 2008, *World Health Statistics*, WHO Press, Geneva, Switzerland.