

1.2 ANALYSIS OF AGE, BIRTH COHORT, AND PERIOD EFFECTS

Health surveys conducted in population samples usually include participants over a broad age range. Age is a strong risk factor for many health outcomes and is also frequently associated with numerous exposures. Thus, even if the effect of age is not among the primary objectives of the study, given its potential confounding effects, it is often important to assess its relationship with exposures and outcomes.

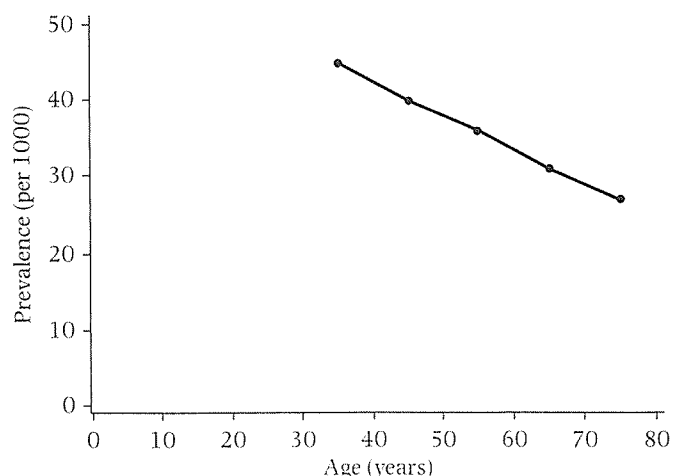
Table 1-1 shows the results of a hypothetical cross-sectional study conducted in 2005 to assess the prevalence rates of a disease Y according to age. (A more strict use of the term "rate" as a measure of the occurrence of incident events is defined in Section 2.2.2. This term is also widely used in a less precise sense to refer to proportions such as prevalence.¹ It is in this more general sense that the term is used here and in other parts of the book.)

In Figure 1-1, these results are plotted at the midpoints of 10-year age groups (e.g., for ages 30–39, at 35 years; for ages 40–49, at 45 years; and so on). These data show that the prevalence of Y in this population decreases with age. Does this mean that the prevalence rates of Y decrease as individuals age? Not necessarily. For many disease processes,

TABLE 1-1 Hypothetical data from a cross-sectional study of prevalence of disease Y in a population, by age, 2005.

Age group (years)	Midpoint (years)	2005 Prevalence (per 1000)
30–39	35	45
40–49	45	40
50–59	55	36
60–69	65	31
70–79	75	27

FIGURE 1-1 Hypothetical data from a cross-sectional study of prevalence of disease Y in a population, by age, 2005 (based on data from Table 1-1).



exposures have cumulative effects that are expressed over long periods of time. Long latency periods and cumulative effects characterize, for example, numerous exposure/disease associations, including smoking-lung cancer, radiation-thyroid cancer, and saturated fat intake-atherosclerotic disease. Thus, the health status of a person who is 50 years old at the time of the survey may be partially dependent on this person's past exposures (e.g., smoking during early adulthood). Variability of past exposures across successive generations (birth cohorts*) can distort the apparent associations between age and health outcomes that are observed at any given point in time. This concept can be illustrated as follows.

Suppose that the same investigator who collected the data shown in Table 1-1 is able to recover data from previous surveys conducted in the same population in 1975, 1985, and 1995. The resulting data, presented in Table 1-2 and Figure 1-2, show consistent trends of decreasing prevalence of Y with age in each of these surveys. Consider now plotting these data using a different approach, as shown in Figure 1-3. The dots in Figure 1-3 are at the same places as in Figure 1-2, except that the lines are connected by *birth cohort* (the 2005 survey data are also plotted in Figure 1-3). Each of the dotted lines represents a birth cohort converging to the 2005 survey. For example, the "youngest" age point in the 2005 cross-sectional curve represents the rate of disease Y for individuals aged 30 to 39 years (average of 35 years) who were born between 1965 and 1974—that is, in 1970 on average (the "1970 birth cohort"). Individuals in this 1970 birth cohort were on average 10 years younger—that is, 25 years of age at the time of the 1995 survey and 15 years of age at the time of the 1985 survey. The line for the 1970 birth cohort thus represents how the prevalence of Y changes with increasing age for individuals born, on average, in 1970. Evidently, the cohort pattern shown in Figure 1-3 is very different from that suggested by the cross-sectional data and is consistent for all birth cohorts shown in Figure 1-3 in that it suggests that the prevalence of Y actually

TABLE 1-2 Hypothetical data from a series of cross-sectional studies of prevalence of disease Y in a population, by age and survey date (calendar time), 1975–2005.

Age group (years)	Midpoint (years)	Survey date			
		1975	1985	1995	2005
<i>Prevalence (per 1000)</i>					
10–19	15	17	28		
20–29	25	14	23	35	
30–39	35	12	19	30	45
40–49	45	10	18	26	40
50–59	55		15	22	36
60–69	65			20	31
70–79	75				27

**Birth cohort*: From Latin *cohors*, warriors, the 10th part of a legion. The component of the population born during a particular period and identified by period of birth so that its characteristics (e.g., causes of death and numbers still living) can be ascertained as it enters successive time and age periods.¹