29

Epidemiology

Mark H. Schiffman, M.D., and Louise A. Brinton, Ph.D.

Most pathologists are part-time epidemiologists as well. The two medical disciplines are more closely allied than most realize. Epidemiologists study the distribution and determinants of diseases in human populations. In current medical practice, diseases often are defined by histopathological diagnoses or by clinical pathological test values. Thus, whenever a pathologist shifts intellectually from the level of the individual slide or specimen to thinking about a group of diagnoses, an informal epidemiologic question is being raised. For example, "How common is this diagnosis?" is a question of prevalence or incidence. "Why am I seeing so many cases of this type of tumor?" is a question of time trends. "How would my colleague interpret these slides compared with me?" is a question of interpathologist agreement. "What causes this disease I am seeing every week?" is a question of etiology that can be addressed by pathologists working as epidemiologists, or with them.

This chapter is meant to introduce the major epidemiologic concepts of greatest use to pathologists who are considering a research project or who wish to think more

R. J. Kurman (ed.), *Blaustein's Pathology of the Female Genital Tract*

formally, at the population level, about their case material or diagnostic criteria. The review is certainly not exhaustive; rather, it is meant to be informal and readable, and to encourage the pathologist to pursue epidemiologic projects and collaborations. The first section, accordingly, is organized around types of possible epidemiologic studies that a pathologist might wish to pursue. The next section outlines nonmathematically a few basics of statistical thinking that pathologists need to know if they wish to do more formal epidemiologic research. The third section discusses a few problem issues that often emerge when epidemiologists and pathologists work together. The final section summarizes what is known about the epidemiology of the major gynecological cancers.

Applications of Epidemiology to Pathology

This section illustrates the types of epidemiologic projects that a pathologist may undertake, either informally or formally. The examples are drawn mainly from etiologic and screening studies of gynecological neoplasia, especially cervical neoplasia.

Throughout this section, epidemiologic terms are introduced and simply defined. There is a useful dictionary of epidemiology for readers interested in learning more terminology.[104] For a more complete understanding of basic epidemiologic concepts, the reader is referred to one of several introductory texts.[109,114,119] For readers with some statistical training, a brief text summarizing the basic methods of epidemiology also is available.[84]

Prevalence, Incidence, and Mortality Rates of Disease

One of the first questions that an expert or novice epidemiologist is likely to ask about a disease under study is "How common is it?" The pathologist at the microscope is interested in how common various conditions are, as one element of differential diagnosis (witness the maxims, "Rare diseases occur rarely," and "If you hear hoofbeats, think of a horse not a zebra.")

For the pathologist considering a research study, the frequency of disease occurrence is crucial for two reasons: On the practical level, very rare conditions are difficult to study epidemiologically. The statistical principles underlying epidemiology require moderately large numbers in order to deal with chance, which is the unavoidable and defining characteristic of observational studies in humans.

More importantly, the amount of disease in a population is the starting point for epidemiologic thought, leading to all of the major epidemiologic comparisons, such as "How much disease occurs in Population A compared with Population B, and what does the difference tell us? Why is the amount of disease changing over time? What risk factors are associated with groups having the most disease?"

Because measuring the occurrence of disease is so important to epidemiologists, they find it important, like skiers discussing snow, to define terms carefully using a resultant epidemiologic jargon. A few key terms related to the frequency of disease occurrence are essential and worth memorizing by anyone interested in epidemiology.

The *prevalence* of a disease is the number of occurrences of the disease in a given population at a given time, for example, "Twenty percent of the patients seen in this clinic have at least reactive changes on their Papanicolaou smears." Often, prevalence is discussed with reference to a single point in time, as in a screening program, yielding a *point prevalence*: "Two percent of the screening smears last month showed changes suggestive of CIN."

The *incidence* of disease is the number of new cases that develop in a given time period. Accordingly, *incident disease* refers to new disease, whereas *prevalent disease* refers to all the cases in the population, whether new or chronic. The connection between prevalence and incidence is the *duration* of the condition (Prevalence = Incidence × Duration); thus, the prevalence of rabies in a given week is close to the incidence because duration is unfortunately short, whereas the prevalence of a long-duration disease such as rheumatoid arthritis much exceeds the incidence for any time period.

Incidence is defined most often as a yearly rate, as in "13,500 incident cases of cervical cancer were diagnosed in the United States in 1993." However, *lifetime cumulative incidence* also is an intuitively useful term, meaning the estimated risk of occurrence of a disease over a woman's life: "About 1% of women in the United States will develop cervical cancer in their lifetime." For chronic diseases such as cancer, endometriosis, or genital herpes, incidence usually is thought of as a one- time phenomenon, for example, second primaries rarely occur. (Recurrences of the same disease imply that it is prevalent, not incident). For acute, self-limited, or curable conditions, such as gonorrhea, incidence must be defined over a fairly narrow range of time if one wishes to avoid counting twice the same patients in a population.

Rates of death from a disease are measured as the *mortality rate*. The connection between incidence and mortality is, of course, survival, measured often by the *case/fatality ratio*.

In sum, the epidemiologist is interested in the prevalence, incidence, and mortality rates of a disease as the fundamental basis of further study. These terms can be applied to any study population, whether that population is a single gynecological practice or hospital, a city, country, or the world.

National incidence and mortality data are cited most often when discussing the scope of a medical problem. Where can national data be obtained? In the United States, most pathologists probably are aware that mortality rates from all causes are compiled and available from a variety of sources, most notably and simply from the National Center for Health Statistics (6525 Belcrest Road, Room 1064, Hyattsville, MD 20784; 301-436-8500). Despite the problems of death certificates as the source of the mortality data, with the obvious uncertainties and errors in ascribing causes of death, mortality rates for often-fatal diseases usually are the most reliable gauge of disease occurrence, when comparing different populations or time periods (except when the case/fatality ratio has been altered sharply by improved treatment).

If a condition is not often fatal, mortality rates may not be useful for disease surveillance. It is often more difficult to obtain reliable incidence data, and the researcher must rely on data from voluntary registries, published surveys, or occasionally government-mandated registries. For cancer, fortunately, careful and complete incidence rates for a (nonrandom but stable) 10% sample of the U.S. population are compiled by the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program. The most accessible source of SEER cancer incidence and survival data (as well as national cancer mortality data) is *CA—A Cancer Journal for Clinicians*, published annually by the American Cancer Society and mailed free to physicians. More detailed cancer data can be obtained from other American Cancer Society publications, such as *Cancer Facts and Figures*, or from the SEER program itself (EPN 343 National Cancer Institute, Bethesda, MD 20892;301-496-8510).

Geographic Differences and Time Trends in Disease Occurrence

Pathologists may wish to go beyond descriptions of disease occurrence at a place and time, to compare rates between geographic areas or over time. The usual hope is that the comparisons may yield clues to etiology and pathogenesis. A cautious approach is critical because of the omnipresent effects of chance on observational data. How can one tell if the amount of disease in one place or time is truly different from the amount found earlier or elsewhere? Disease rates fluctuate over time and place. Many geographic differences and temporal trends do not persist over time, appearing random (to the limit of our understanding!).

Hence, the need for statistics as one of the disciplines underlying epidemiology. Distinguishing chance differences from true differences requires statistical thinking and an appreciation of the types of differences that arise by chance. This point is important because overinterpretation of chance differences is one of most common errors that novice epidemiologists make when comparing disease rates from one place or time with another. For example, many cancer "outbreaks" in which several neighbors get similar tumors turn out to be quite explainable as chance clusterings of events, expected for common malignancies like breast cancer. A good bit of advice might be to treat health statistics like the monthly economic news: it takes a long-term trend or a persistent difference to trust that something important is happening.

When comparing place to place or analyzing time trends, the cardinal rule is to make sure that the comparison is valid. A checklist of common-sense questions should be asked:

1. Are the rates being compared truly comparable (incidence, prevalence, mortality)? In particular, are the sources of data comparable (e.g., a mandatory registry cannot be compared with a voluntary reporting system because of differences in the completeness of reporting).
2. Are the diagnostic criteria the same in both comparison groups? This particular problem has plagued the interpretation of time trend data regarding minor cervical cytological abnormalities because increased recognition by pathologists of subtle koilocytotic changes cannot be distinguished easily from increased incidence of koilocytotic atypia.
3. Are the two populations comparable in age and other factors affecting risk of disease? No one would think of comparing the prevalence of CIN in a gynecological referral practice to the prevalence in a screening clinic because, of course, the prevalence would be higher in the referral clinic. But some researchers make the analogous mistake of comparing populations that differ with regard to age, socioeconomic status, or other more subtle characteristics related to the risk of disease (called *confounding variables* in epidemiologic jargon). Most importantly, almost all diseases vary in incidence and prevalence by age; thus, almost all comparisons should take age into account. The section on error and bias below mentions simple methods of adjustment for age and other confounding variables. The statistical bases of making geographic and temporal comparisons are covered below in the sections on descriptive data and measures of risk.

Validating New (or Old) Histopathologic Diagnostic Distinctions

The creation and refinement of pathological classifications can be aided by epidemiologic corroboration. For example, the Bethesda System of cervical cytology combines koilocytotic atypia and CIN 1 as low grade squamous intraepithelial lesion (LSIL). This combination was supported by epidemiologic data. The two diagnoses, which are not reliably distinguishable on morphologic grounds, share the same epidemiologic profiles of younger average

age and varied human papillomavirus (HPV) types as compared with the older average age and restricted HPV types found in higher grade lesions. As another example, a recent pathological study of squamous vulvar cancer, which proposed new pathological subtypes, was strengthened by a separate epidemiologic analysis showing that the new subtypes had different epidemiologic characteristics. Pathologists and epidemiologists can work iteratively to refine disease classifications, asking each other "Do categories X and Y look the same or different from your point of view?"

Judging Intra- or Interpathologist Agreement

Pathology agreement studies have been motivated by the needs of both disciplines. Pathologists obviously are concerned with the reliability of the diagnoses they make. Epidemiologists are concerned with uniform case definition in their studies. When comparisons of intra- and interpathologist agreement are performed, the epidemiologist can serve the role of scientific organizer, ensuring independence of the reviews by *masking* the reviewers (also called *blinding*) to each others' diagnoses until after the data are complete. It is the widespread opinion of epidemiologists that unmasked comparisons, in which reviewers have access to each others' diagnoses, have limited scientific value. Like all human beings, pathologists tend to agree much more in public than in private, and masking provides a guarantee that a comparison rather than a consensus is being achieved. In the area of cervical pathology, the diagnosis of CIN by either cytology or histology has proved much more variable among experts when masked comparisons were performed than initially expected.

Epidemiologic Studies of Disease Etiology

Epidemiologists attempt to find the determinants of disease by statistically correlating the presence or absence of possible *exposures* (often called *risk factors*) with the presence or absence of disease. Epidemiologic studies attempting to relate exposures and disease are called *analytic studies*, as distinguished from *descriptive studies* that yield rates of disease without explicitly addressing etiology.

A description of the many types of analytic studies is beyond the scope of this chapter. At the simplest level, *prospective* or *cohort studies* start with the measurement of an exposure, then proceed to compare incidence rates or *absolute risk* of disease in the exposed versus the unexposed groups. The ratio of the incidence rate in exposed subjects divided by the incidence rate in the unexposed is called the *incidence rate ratio*. The reader might correctly expect that there are as many types of rate ratios as there are types of rates (e.g., *prevalence rate ratio*, *lifetime cumulative risk ratio*). Many epidemiologists casually refer to the entire group as the *relative risk* of exposed versus nonexposed subjects, and use the abbreviation *RR* as a general shorthand.

Prospective studies are the most appealing type of analytic study because they determine most directly how commonly disease occurs in exposed versus unexposed individuals. The relative risk, measured directly, is an intuitively clear answer to the question: "If a woman has this characteristic (the exposure), how much more likely is she to develop the disease, compared with a similar unexposed woman?" The absolute risk translates as "How likely to get disease is an exposed woman?" (see below, Measures of Risk). The problem with prospective studies is that they are expensive, usually take years to organize and complete, and must be very large to generate enough occurrences of disease for reliable estimates of risk, unless the disease is extraordinarily common.

Other analytic study designs try, in general, to estimate the relative risk that might be observed in the ideal prospective study while saving time and money. Analytic studies that start by collecting a series of *cases* (women diagnosed with a given disease) and appropriate *controls* (women without that disease who are measured for comparison) are called *case-control* studies. The exposures of interest are ascertained for both groups and the relative risk (RR) of disease among the exposed versus the unexposed is estimated by calculating the ratio of the odds of exposure in cases versus controls (for more explanation, see the the statistical section on measures of risk).

The estimation of the prospective relative risk by the case-control *odds ratio* (*OR*) is one of the most important statistical concepts in epidemiology, and one of the most subtle. For this statistical approximation to be valid, incident cases and controls must be chosen to be strictly comparable. The control group must represent the group of subjects at risk of developing disease at the time the incident case was diagnosed, otherwise the estimation of the relative risk can be grossly mistaken because of *bias* (a nonrandom or systematic error in estimation of a statistic, to be distingushed from *random error*).

In practice, it is difficult to define and recruit an unbiased sample of the general population that gave rise to the cases appearing in one hospital or clinic. Thus, all kinds of compromises of convenience and practicality must be made, and it becomes difficult to avoid bias in choosing controls. For example, smoking causes or worsens so many kinds of illness that it is very difficult to use hospitalized controls to estimate the relative risk of a disease in smokers compared with nonsmokers. The exposure to smoking in the hospitalized controls is elevated compared with the general at-risk population; thus, the odds ratio obtained in a naively conducted hospital-based study tends to provide too low an estimate of the relative risk.

Because case-control studies are used so commonly as an analytic design, choosing proper controls is one of the two most important aspects of epidemiology. The other is assuring proper measurements of exposure and disease. The mark of a good epidemiologist is a dedicated attention to control selection, whereas many novices tend to focus more on the cases while relying on a *convenience sample* of whichever controls are most easily available.

Besides prospective and case-control studies, another common analytic study design is the *cross-sectional* study, in which exposure and disease status are ascertained concurrently for a study population. An example would be a screening study of HPV infection and abnormal cervical cytology, in which all women attending a clinic are tested for viral DNA at the same time the cytological smear is taken. The analysis of a cross-sectional study is somewhat similar to that of a case-control study, but the researcher must be careful because the cases are a combination of incident and prevalent disease. The odds ratio computed in a cross-sectional study is a good estimation of the prospective relative risk only if certain conditions are met, including an assumption that the disease under study is rare (an assumption not met for cervical cytological abnormalities in many clinics).

The pathologist collaborator should play a key role in all analytic studies of diseases whose definitions rely on nonroutine pathological expertise. Misclassification of disease status can be damaging to a study because the effect of misclassification on correlative statistics such as the relative risk and odds ratio is, generally, to reduce the apparent strength of the association between disease and exposure. If the disease is poorly defined, no epidemiologic risk factors may be found even if they exist. Moreover, it is often very difficult to measure the risk factors (exposures) without substantial error, whether laboratory testing or interviews are being used. The combination of multiple errors in measuring both exposure and disease can literally make a study worthless. For example, early studies correlating HPV DNA detection and CIN revealed only a moderate association, in that less than 50% of cases were found to be HPV positive.[148] Moreover, HPV infection was not apparently associated with sexual activity, an established strong risk factor for CIN. These weak associations were due to misclassification. Subsequent studies with better HPV tests and expert review of pathology revealed that virtually all cases of CIN contain HPV DNA,[153] and that HPV is the sexually transmitted agent explaining the association of sexual activity and risk of CIN.[148]

As the result of the strong, damaging effects of misclassification on epidemiologic studies, epidemiologists pay careful attention to the pathological classifications that define their study cases and controls, and often establish formal collaborations with reviewing pathologists as part of epidemiologic studies.

Follow-up Studies of Patients with the Same Pathological Diagnosis

Clinicians, pathologists, and epidemiologists are all interested in learning what happens to patients diagnosed with a given disease. For a possibly fatal disease, survival rates are critical, whereas for other chronic diseases, progression rates or recurrence rates often are estimated. It is often of interest to divide the patients into groups to determine whether subtypes of disease follow different courses, or whether different treatments influence outcome. The *randomized clinical trial* is a specialized version of such a follow-up study, in which subjects are randomly assigned to various treatment groups to maximize the comparability of the groups. The hope is that the randomization will minimize differences in both known and unknown confounding variables that could bias the comparison.

Follow-up studies almost invariably involve the concept of *time to an event*. In other words, it is important when death or recurrence or progression occur, not just if they occur. Clearly, all participants in any clinical trial eventually die, the question is when (and why). A good treatment prolongs time to death, a bad type of disease shortens it.

Because of the critical notion of "time to event" in epidemiologic follow-up studies, they depend heavily on actuarial methods, like survival curves and life-table analyses, when comparing exposed to unexposed patients or treated to untreated patients. The central statistical concept in such studies is a kind of rate called a *hazard*, which refers to the risk of disease occurring in a unit of follow-up time. A hazard is computed as the number of *events* (e.g., death, cure, progression) divided by the amount of *person-time* of follow-up. One thousand women followed for a year or 100 women followed for 10 years both yield 1000 person-years of follow-up time. Two cancers arising during that follow-up would yield a hazard of two cases per 1000 person-years.

A hazard is a special kind of rate because it is conceived of as the rate of disease at a single moment in time, as the mathematical "limit" of the rate as time "goes to zero." Accordingly, the hazard of disease can change from moment to moment as conditions change. A woman lights up a cigarette and her hazard for lung cancer increases. She quits smoking the next day and her hazard decreases.

Moreover, the computation of the denominator of hazards, person-time, of follow-up requires some training and thought. For each successive time interval during follow-up, the denominator of women at risk for an event changes. For example, women are lost to follow-up as they drop out of the study, or they die for other reasons, or they experience the event itself (since one can develop a disease for the first time only once). Thus, computing the proper amount of person-time during which the events occurred requires some knowledge of *censoring*, which is the proper

deletion of irrelevant follow-up time during which the subject was not truly at risk of diagnosis of the event.

Usually, researchers are not content to describe the simple survival curve of a disease, which represents the hazard of death over time after diagnosis. They wish to determine which factors affect the hazard, that is, what the relative or *proportional hazard* of death, etc., might be for women in different groups defined by pathological differences or treatment types. The proportional hazard is almost identical to the incident rate ratio discussed above, but the denominator is person-time of follow-up, not just time. Proportional hazard analyses are too complex to be described here, and pathologists performing follow-up studies might consider consulting an epidemiologist or biostatistician early in the design phase of such projects. Data collection must be organized carefully to permit a correct determination of person-time.

Screening for Gynecological Malignancies

Screening is inherently epidemiologic, thus, the pathologist involved in screening programs (e.g., cervical cytological screening or CA125 testing) needs to understand the interrelated concepts of sensitivity, specificity, and predictive value. The basics are outlined below in a statistical section on screening.

A common mistake in evaluating the results of a screening trial is to ignore the clinical setting. The *sensitivity* of a screening technique (percentage of diseased women who test positive) and its *specificity* (percentage of disease-free women who test negative) theoretically do not change when the test is taken out of a high-risk hospital clinic to be applied to the general population. But most clinicians are more interested in the *positive predictive value* and *negative predictive value*, two statistics that are highly dependent on the clinical setting. For example, the positive predictive value is the percentage of women testing positive who truly have disease. Given the same sensitivity and specificity, positive predictive value decreases sharply as the prevalence of the disease decreases. Therefore, the same screening test that looks promising because of high sensitivity in a high-risk clinic often will perform poorly in the general population, producing so many false positives compared with the disease yield that the costs outweigh the benefits. As a general rule, specificity is the prime requirement of a screening test. A screening test such as a tumor marker must be highly specific (negative in virtually all nondiseased women, certainly more than 95%) to be cost effective for general population screening.

Basic Statistical Concepts

Hopefully, the preceding discussion has firmly established the relevance of epidemiology to gynecological pathology research and even daily practice. Epidemiologic work requires an understanding of biostatistics. This section mentions the bare basics of what the author believes pathologists collaborating in epidemiologic research might wish to know about biostatistical methods. Introductory biostatistics texts are available and easy to read for the pathologists wishing to work independently, or for those who want computational formulae for chi-square or other commonly used tests.

Variability as a Fundamental Principle of Pathology

Virtually all measurements that one could make about a human population are variable. Height, weight, fine points of anatomy, metabolic patterns, serum levels of hormones, and nutrients are all commonly recognized to be variable. The same variability is seen by pathologists at the tissue and cellular levels, and by research pathologists at the molecular biological level (e.g., varying tissue levels of DNA adducts given equivalent carcinogenic exposures, genetic polymorphisms in human genes, and varying molecular responses to infection with viral DNA). Even the intricate, multistep molecular pathways to cancer demonstrate substantial variability between individuals who develop the same type of malignancy.[171]

Variability in pathology is mainly described by *categorical* or *discrete* data and statistics, as compared with *continuous* data and statistics (the province of the mean, median, and standard deviation). Similar (but not identical) histological and cytological appearances are categorized and named. More attention is paid to the borderlines and overlaps of the categories, rather than subtler differences within the categories (unless splitting into finer categories is being considered). Categorical data analysis relies on *contingency tables*, which are discussed in a section below. Contingency tables, like the common 2×2 *table*, are frequency counts of categorical data, for example, how many (not what percent of) CIN 2 lesions demonstrated aneuploidy or not, compared with how many CIN 3 lesions demonstrated aneuploidy or not.

The variability in categorical data such as pathology categories shows up in diagnostic error (i.e., the misassignment of a patient to the wrong category). In general, error cannot be avoided. To the epidemiologist, categorization of variable biological continuua virtually dictates that there will be error. If two categories blend into each other with regard to a characteristic (even one as complex and general as microscopic appearance), they cannot be perfectly separated based on that characteristic. Thus, pathologists search for additional characteristics to discriminate difficult-to-distinguish indeterminate cases, such as immunocytochemistry, but these ancillary measurements also have error and overlap. There is a field of statistics called *discriminant analysis*, in which the goal is to determine how many characteristics must be measured to maximize correct assignment to overlapping categories. This compli-

cated set of statistical methods underlies the development of computer-assisted cytology screening.

Error Versus Bias

Error is inevitable, but epidemiologists hope that it is mainly random, not systematically pushing the data in one way or the other. *Random error* reduces the *reliability* of repeated measurements, affecting their *precision*, and reduces the perceived strength of correlations, but the average measured value still becomes increasingly true or *accurate* as the study size increases. Systematic error, called *bias*, impacts directly on the accuracy of the measurement; no matter how large a study based on biased measurements is, the answer will be wrong. Thus, epidemiologists struggle to reduce random measurement error, but they have an even stronger dislike of biased measurements. If the exact direction and magnitude of a fixed bias were known, the data could be adjusted (like a scale that always reads three pounds heavy), but adjustments for bias usually are not possible.

Epidemiologists combat error and bias in a few standard ways. To quantify and reduce random error, reliability is measured by repeating data collection, whether that involves re-asking a question, re-running an assay, or submitting a pathology slide for re-review.

For continuous variables, statistics of reliability include the *variance* (sum of the squared deviations of the measurements from their arithmetic average or *mean*, divided by the number of data points minus one), the *standard deviation* (the square root of the variance) of the raw data or of the mean (called the *standard error* of the mean), and the *coefficient of variation* (the ratio of the standard deviation to the mean).

For categorical variables, statistics of reliability include the simple percentage of agreement and more complicated statistics mentioned below in the section on measures of interpathologist agreement.

To reduce bias, epidemiologists would like to compare the study measurement to a reference standard of truth, but such reference standards virtually never exist. Certainly, there is no source of absolute truth in pathology, only advancing degrees of expertise correlated with decreasing amounts of diagnostic error. Therefore, to reduce bias in pathology, researchers are limited to the comparison of different experts. To the extent that truly independent experts agree (without consideration of each other's opinion), the possibility that either one is biased is reduced.

Of course, experts can share or even teach each other the same biases, destroying the notion of independence and reducing the possibility of identifying the biases. Hence the value to epidemiologists of outspoken pathologists, who point out beliefs in pathology that might be changing.

Also to reduce the possibility of bias, epidemiologists try to ensure that all study measurements are made independent of each other, so that knowledge of one variable cannot bias a decision about another. The difficulties of masking are discussed in a section below.

Descriptive Data

The terms used most often to describe and summarize descriptive data, such as prevalence and incidence, are defined in the section on geographic differences and time trends above, and are not repeated. A few additional statistical concepts critical to the interpretation of descriptive data are mentioned.

First, there is an important choice of scale in the plotting of descriptive data. The scale of the y or vertical axis greatly affects the appearance of the data, and must always be noted when examining plotted data. A log scale flattens curves and reduces the apparent strength of trends and differences, whereas an arithmetic scale does the opposite. On a log scale, an increasing, straight-line trend implies an exponential, not linear, rate of increase.

A common error in inference when interpreting descriptive data is the *ecologic fallacy*, the attribution of causality to an association seen only in descriptive data. For example, the international risk of colon cancer (mortality rates for each country plotted on a graph) correlates with the average dietary intake of those countries for fat, meat, and sugar, and with the average amount of sunlight (the major determinant of vitamin D levels). To assume that all four variables are true risk factors for colon cancer at the level of the individual would be an example of the ecologic fallacy, confusing descriptive data for analytic (individual level) data.

In the interpretation of time trend data, the possibility of a *cohort effect* must be kept in mind. A cohort effect, familiar by analogy to anyone who studies the sociology of baby-boomers, is the variation in disease occurrence that occurs in a population over time, as successive birth cohorts (persons of the same age) experience the unique environment that typifies their life course. For example, based on cross-sectional prevalence data compiled in 1991, the prevalence rates of koilocytotic atypia of the cervix decrease sharply with increasing age, from a peak at about age 20 to 25 years. This age trend might represent a biological phenomenon, the result of immunity, with many women becoming infected with HPV at the time of initiation of sexual intercourse, then becoming increasingly immune and having fewer new sexual partners as they age. Or, the age trend also could reflect a cohort effect, with changing sexual practices and increasing prevalence of HPV infection over the past decades placing younger women today at higher risk for koilocytotic changes compared with their older sisters and mothers.

To distinguish cohort effects from simple age trends requires a *cohort analysis*, a type of descriptive graphing in

Table 29.1. The basic contingency table

	Disease	*No disease*	*Total*
Exposed or test positive	a	b	a + b
Unexposed or test negative	c	d	c + d
Total	a + c	b + d	a + b + c + d = n

which the age- specific prevalence rates are graphed separately for each birth cohort. These analyses usually are difficult enough in interpretation to merit a statistical consultation.

The Basic Contingency Table

The pathology slide of epidemiology is the contingency table, the basic form of which is the 2 × 2 table (Table 29.1).

Most important epidemiologic findings, relating an exposure to risk of a disease, have been derived and can be expressed in this simple form. Extension of the table to more rows or columns does not change the concepts, only the statistical complexity.

The most common statistics computed from a contingency table are simple percentages, which can then be compared: "Ninety percent of the group with disease were smokers [(a/ a + c) = 0.90] compared with 20% of the nondiseased [(b/ b + d) = 0.20]. These proportions could be compared statistically using the well-known *t* test or another test of the difference between independent proportions. More often, the *chi-square statistic* is computed, which gives equivalent interpretations but has a slightly different intent.

The chi-square test is meant to determine whether the disease categories and the exposure categories are associated or independent (i.e., does being exposed affect the probability of having disease?). Chi-square values are derived by comparing the expected counts of a, b, c, and d to the values that would be expected if disease and exposure were totally independent. For example, the expected value of a is the cross-product of (a + b) × (a + c) divided by *n*. The divergence of observed from expected values for all of the *cells* of the table (a, b, c, d) are summed to derive the chi-square statistic. The larger the statistic, summarizing how much observed counts differ from expected, the more likely disease and exposure are associated by more than chance.

The chi-square statistic obtained is compared with the tabled values of the *chi-square distribution* to yield a *p* value, the probability of observing such a chi-square value if disease and exposure are not related. If the *p* value is less than 0.05 or 0.01, then convention dictates that chance is unlikely to explain the degree of association seen in the table, and the association is considered *statistically significant*.

Thanks to many recent published cautions, most clinicians and researchers know that a strict dependence on *p* values is incorrect, because the magnitude of the *p* value depends on the size of the study. Smaller studies require stronger associations to achieve the same level of statistical significance; thus, a *p* value of 0.06 in a small study by no means rules out a true exposure-disease association whereas a highly statistically significant difference from a huge study may be so small as to be clinically irrelevant.

Contingency tables larger than 2 × 2 should be analyzed in a methodical and hierarchical fashion, not restricting the analysis to the most "significant-looking" internal comparisons. First, the evidence for association in the full table should be assessed and, if there is none, then the analysis should stop. A common mistake some novices make is to look at a large contingency table, choosing the most interesting difference seen, then testing the significance of that extracted comparison. Given a large enough contingency table, some subtables will yield statistically significant results by chance alone. Permitting a prescreening of the data before applying a statistical test to the most divergent data points is wrong. If one wishes to define the likely source of the association when the overall contingency table indicates statistical significance, the proper approach is to analyze smaller subtables in a complete and hierarchical manner. A formal description of the proper approach to contingency table analyses can be found in standard biostatistics texts.

When the number of study subjects is very small, such that the expected count in any cell is less than about five, then chi-square analyses are unreliable and should be replaced by a test called *Fisher's exact test*. Of course, if the study is too small, no result will be statistically significant.

One other key point about contingency tables is that the two measurements (disease status and exposure, for example) must be assumed to be independent, as one embarks on statistical testing. Although a significant chi-square statistic indicates that the measurements are not independent, the initial or *null hypothesis* of independence is what the test is designed to reject. Thus, standard chi-square analyses should not be performed to test tables in which the measurements are explicitly correlated, as in interpathologist agreement studies, or comparisons of the efficacy of two cell collection techniques used in the same group of patients. For *paired-sample comparisons*, the *McNemar's test* is easy to use. The test ignores the points of agreement of the two measurements and tests the statistical significance of the amount of divergence.

Measures of Risk (Absolute, Relative, and Attributable Risks)

The chi-square provides limited information regarding the strength of an association (yes/no). Therefore, epidemiologists often prefer instead to compute the more informative

statistic, the relative risk (or odds ratio estimate of the relative risk). These key terms are defined in the section on epidemiologic studies of disease etiology. In this section, the relation of the terms to the contingency table are explained, with a brief discussion of ancillary topics such as statistical adjustment of confounding variables, interaction, and confidence intervals.

Suppose a prospective study started by defining an exposed group and an unexposed group of women, then followed the two groups for disease occurrence. The absolute risk of disease following exposure can be represented as an incidence rate a/(a + b). (See Table 29.1.) The time period for this incidence rate is implicitly the duration of follow-up. The absolute risk of disease in the unexposed group, analogously, would be the incidence rate c/(c + d). The ratio of these absolute risks would be the relative risk (specifically, the incidence rate ratio) in exposed versus nonexposed women, a/(a + b) divided by c/(c + d). A relative risk above 1.0 implies an increased risk. For example, a relative risk of 2.0 means that the risk of disease in exposed women is twice that of unexposed women. In contrast, a relative risk between 0.0 and 1.0 indicates a protective association (a relative risk of 0.5 implies a halving of risk associated with the exposure). A relative risk of approximately 1.0 implies the exposure is not related to risk of the disease.

Prospective studies permit the computational directness and intuitive quality of the relative risk calculation and the ability to decompose the relative risk into the absolute risks among the exposed and unexposed groups.

In contrast, absolute risks usually cannot be calculated in case-control studies because the true numbers of exposed women (a + b) and unexposed women (c + d) are not known. In fact, in 2 × 2 tables from case-control studies the values a + b and c + d are meaningless and should never be computed. The numbers of cases (a + c) and controls (b + d) are chosen first, and not in proportion to the true ratio of cases to controls in the population. Cases are almost always sampled in excess; in fact, oversampling cases to overcome the limitation of rarity is the major reason to perform a case-control analysis.

As mentioned earlier, although case-control data do not permit direct calculation of the relative risk, the odds ratio provides a valid estimate if the cases and controls represent an unbiased sample of all women with and without disease in the population, and if the disease in question is very rare (if the cases are all incident, the rare disease assumption is not as important, unless the disease is so common that a nonnegligible percentage of the population is developing it at any given time).

To understand the concept of the odds ratio, again consider a prospective study. The odds of disease in exposed women is a/b, very close to the risk of disease a/(a + b) if a, the occurrence of disease among the exposed, is very infrequent. Similarly, the odds of disease in nonexposed women is c/d, close to the risk of the disease if uncommon in the nonexposed women, c/(c + d).

With a little algebra, it is easy to see that the relative odds or odds ratio for a rare disease (a/b divided by c/d, often computed as the cross- product ad/bc) is quite close to the relative risk.

The important point is that the cross-product ad/bc can be computed from a case-control study without knowing the total number of exposed and unexposed women. As long as the odds a/c and b/d are unbiased with regard to the entire population, then a/c divided by b/d equals ad/bc equals the prospective odds ratio of a/b divided by c/d. The key is to select an unbiased sample of cases and controls. Because epidemiologists usually try to recruit all cases occurring in a population, bias among cases usually is not an issue unless participation rates are poor. The place where bias is a major concern is among the controls. Epidemiologists spend most of their intellectual energy attempting to ensure that the ratio b/d in controls (also thought of as the percentage of controls exposed to the risk factor) is unbiased compared with the same ratio in the whole population that gave rise to the cases. Without the elimination of bias, the odds ratio does not estimate the relative risk, and the case-control design will yield a false result.

Confounding is the type of bias that concerns epidemiologists the most, particularly when they are conducting case-control studies or nonrandomized prospective studies. *Confounding variables* are factors that influence both the risk of disease and the likelihood of exposure to a risk factor under study. The relationship between exposures, confounding variables, and disease outcome is illustrated in Figure 29.1.

When assessing whether an exposure such as smoking causes cervical cancer, the researcher must consider and adjust for the confounding influence of sexual activity leading to HPV infection, a central cause of cervical cancer. Women who smoke tend to have more sexual partners and, consequently, are more likely to be HPV infected (i.e., the confounding variable is linked to the likelihood of the study exposure). HPV is a cause of cervical cancer. The apparent influence of smoking on risk of cervical cancer is reduced by statistically adjusting for HPV infection status. A residual association between smoking and cervical cancer risk

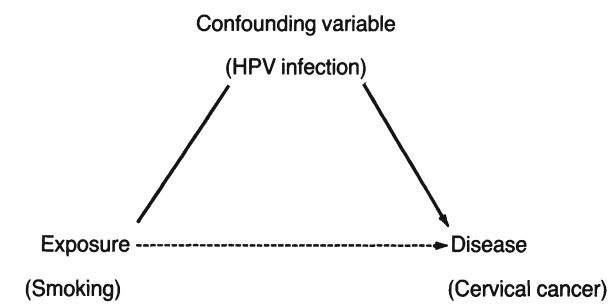

Fig. 29.1. Confounding. Confounding variables are risk factors for diseases correlated with the exposure under study.

may still exist, as mentioned later in the chapter, but the true strength of the association is only estimated, correcting when adjusted for confounding. In summary of this important point, epidemiologic analyses must adjust statistically for the influence of confounding factors to generate unbiased risk estimates.

Adjustment for confounding is commonly undertaken by one of three methods: *exclusion*, *stratification*, or *regression modeling*. Exclusion is exemplified in the example above by restricting the analysis to women known to be infected with HPV, or to virgins (limiting the exposure to HPV). Using stratification, one could analyze either of the two "strata" defined on HPV status, eliminating the possibility of confounding (if HPV infection status could be defined perfectly). Rather than excluding any subjects, the association of smoking with cervical cancer could be examined separately in each of the two strata (HPV −/ HPV +), providing two unconfounded estimates akin to those derived by exclusion. The risk estimates could then be pooled to obtain a global estimate for the risk of smoking adjusted for HPV. This kind of stratified analysis is commonly performed using a group of procedures called a *Mantel–Haenszel analysis* in recognition of its developers.

A more conceptually difficult approach that is widely used is *logistic regression analysis*, a multivariable regression technique available in the major statistical software packages such as SAS and BMDP. This is especially well suited to calculation of the odds ratio as an estimate of the relative risk in case-control studies. This technique permits the simultaneous estimation of the relative risks for multiple risk factors, adjusted for each others' confounding influences. A discussion of this technique, its uses and misuses, is beyond the scope of the chapter. The commercially available statistical packages offer multivariable regression packages in a seductively simple format that might inspire some novice epidemiologists to perform complicated analyses. However, to master the art of multivariable regression analysis takes statistical training and apprenticeship. Moreover, the results cannot be "checked" easily. It is wise both to avoid and distrust complicated analyses, especially because the bulk of what can be learned from most data sets can be expressed using simple tables and intuitively approachable statistics.

Adjustment for confounding often is not perfectly achieved, particularly when the confounding variable cannot be measured well or when variables under study are highly correlated. In fact, it is sometimes virtually impossible using statistical methods to adjust for the confounding influences of correlated variables. For example, the most conceptually difficult areas of chronic disease epidemiology relate to time. In all data analyses involving time, the correlated effects of age at first exposure, duration of exposure, and latency (time since first exposure) are among the most difficult to figure out.

Sometimes the risk of an exposure is modified by levels of another variable. For example, the risk of esophageal cancer associated with smoking is much higher among alcohol drinkers than among nondrinkers. This effect modification, often called *interaction*, is different from confounding in that no global adjustment to arrive at a single correct risk estimate for the exposure is possible. The proper approach is to present the risk estimates for the exposure separately for each level of the effect modifier.

It is common to place *confidence intervals* around relative risk estimates to indicate the likely range of the true risk that we are trying to estimate. Confidence intervals take into account only random error, not bias, and are conceptually somewhat similar to p values although more informative. Thus, a 95% confidence interval and a p value of 0.05 are both commonly chosen as standard, and have analogous interpretations. For example, if the relative risk of an exposure for a disease is 1.8 with a 95% confidence interval of 1.1 to 3.0, this implies that given random error, the true relative risk has a 95% chance of falling within that range. If the confidence interval for a relative risk excludes 1.0, the result is conventionally considered statistically significant. A relative risk with confidence intervals including 1.0 indicates no statistically significant association between exposure and disease. As with p values, confidence intervals should be used as a guide but not followed slavishly in interpreting data.

Most analytic epidemiologic research centers around estimation of relative risks. Another useful concept, especially for public health applications of epidemiologic results, is the *attributable risk*, also known as the *attributable proportion*, or *etiologic fraction*. These terms subsume several computational forms and subtle differences in meaning, but the general meaning is clear: how much of the disease is due to the exposure and would theoretically disappear if the exposure were eliminated (from 0 to 100%). One useful computational formula for the attributable risk, using the notation in Table 29.1, is: Attributable risk = $[(a/a + c) \times (1 - 1/RR)] \times 100\%$. In words, the fraction of disease attributable to the risk factor is equal to the percentage of cases of the disease who are exposed, adjusted for the strength of the estimated relative risk. Although the formula may appear a bit complicated, it is easy to use. The adjustment part of the formula, $(1 - 1/RR)$, goes to zero as the relative risk goes to 1.0, and goes to 1 as the relative risk goes to infinity. Thus, even if all cases are exposed, the attributable risk will be 0% if all controls also are exposed, because the RR is 1.0 and the adjustment term is 0.

Measures of Interpathologist Agreement

Simply put, there is no universally accepted statistical measure of interrater agreement. The problem is adjustment for the influence of chance agreement, which varies with

the numbers of categories and the composition of the study population. All currently available statistical methods have limitations and, therefore, it is best when possible to present the actual data to the reader, in addition to any percentage or statistic.

Consider a study of interpathologist agreement for the categories of the Bethesda System of cervical cytology. A group of 100 smears was given to two pathologists, who were asked to rate them as normal or benign reactive changes, atypical squamous cells of undetermined significance (ASCUS), LSIL, or HSIL. The trouble with simply calculating percent agreement is not only that some agreement is expected by chance, but the results are strongly dependent on how the smears are chosen. If mainly normal smears were submitted, the percentage of agreement would be high. If a wide range of changes were equally represented, then agreement undoubtedly would decrease. In general, the most information is obtained by choosing a wide range of smears, oversampling the rarer grades to achieve a balanced study group.

The most widely used, more sophisticated statistic of agreement of use to pathologists is the *kappa statistic*. Its limitations are discussed in a worthwhile review.[112] The kappa statistic computes the proportion of agreement in excess of the expected chance agreement. Kappa values can range from 1.0 (perfect agreement) to less than 0.0 (zero implies only chance agreement). The interpretation of kappa values is not clear cut, in that researchers disagree as to what good agreement is. According to one group, values greater than 0.75 represent excellent agreement beyond chance, 0.40 to 0.75 is fair to good, and less than 0.40 indicates poor agreement beyond chance.[60]

Screening Terms

Screening is a special area of epidemiology distinct from descriptive or analytic studies. It is rare to find a useful screening test. Finding a strong risk factor for a disease does not imply that we should screen for that risk factor, because the factor often is too common in the general population to permit its use as a trigger for clinical action.

Screening terms have exact meanings, which may vary from other common uses of the same terms. In Table 29.1 above, the women in cell "a" have *true-positive* screening tests, in that they have the disease and tested positive. The women in cell "c" have *false-negative* results because they have the disease but tested negative. The *sensitivity* of a test, also called the true positive rate, is the percentage of diseased women who test positive [a/(a + c) in Table 29.1]. The screening sensitivity must be clearly distinguished from the analytic sensitivity of a laboratory assay, which has a different meaning.

The *true-negative* results are in cell "d"; the *false-negative* results are in cell c. The *specificity*, also called the true negative rate, is the percentage of women without the disease who test negative [d/(b+ d)]. The concept of specificity is more important in screening than most realize. Because the overwhelming majority of women in a population do not have the disease under study, as the specificity percentage falls even slightly, the absolute numbers of false-positive screening tests rise dramatically in comparison to the number of true positives.

Therefore, decreased specificity leads to low *positive predictive value*, the percentage of women with a positive test who truly have the disease [a/(a + b)]. Positive predictive value is, for many diseases, the major screening statistic of interest. Clinicians ask: If a woman tests positive, what is the likelihood that she will have disease confirmed on referral to the next clinical step (e.g., colposcopically directed biopsy, laparoscopy, or more major surgery). Low positive predictive value leads to overreferral and overtreatment.

For grave diseases, in which overtreatment of normal women is less of a concern than not missing any cases, the *negative predictive value* is an important concept. The negative predictive value is the percentage of women who test negative who are truly disease-free [d/(c + d)]. A clinician may ask, accordingly, "If the test is negative, what is the percentage assurance that the disease is not present, and that I can safely stop the diagnostic workup?" The sensitivity of the test usually is the key determinant of negative predictive value.

Some of the current controversy regarding the proper clinical management of inconclusive cervical cytological smears centers around the competing needs for good negative predictive value (assurance that we are not missing any high-grade disease) and good positive predictive value (desire not to overtreat). This problem highlights an inescapable feature of screening (or more fundamentally of trying to categorize overlapping distributions): increased sensitivity virtually always leads to decreased specificity and, as a corollary, increasingly reassuring negative predictive value can be obtained only at the price of decreased positive predictive value.

There is a formal method for choosing the proper screening *cutpoint* (e.g., level of a serologic assay, degree of atypia meriting the diagnosis of LSIL) to achieve an optimal compromise between sensitivity and specificity. The technique is called the *receiver operating characteristic (ROC)*, because the approach was developed to test how well an electronic receiver could distinguish signals from electrical noise. The concepts are easy to understand, useful, and well explained in one key article that is recommended to anyone wishing to evaluate a screening test.[79]

When screening is mentioned, there is always an implicit notion of a *reference standard* or *gold standard* of disease. The performance of screening tests is described statistically in relation to this reference standard and if it is flawed, then the screening statistics will be flawed. For example, colposcopically directed biopsy with pathological

diagnosis often is taken as the reference standard of cervical intraepithelial neoplasia, but the colposcopic biopsy may be misdirected or the histopathological diagnosis may be in error. Thus, the true performance of screening tests such as cytology, cervicography, or HPV testing may be misinterpreted when compared with the results of colposopically directed biopsies.

Screening tests may detect prevalent disease or predict the future diagnosis of disease, and the two time frames may be confused. If some type of HPV test could truly predict incipient cervical neoplasia, even when biopsies were still negative, it would be misleading to compare the HPV screening result only to prevalent (same-day) disease defined by biopsies.

Another mistake is the following: Researchers who wish to compare the sensitivity of two screening tests double-test a research population, referring for a definitive diagnostic procedure those women who are positive for either screening test. If they then compute and report the "sensitivity" of each test, an error of circular reasoning has been made. Because both screening tests could have missed disease (double false-negatives), the true sensitivity of either test cannot be known without referring all women in the study population for the definitive work-up. Sometimes, in large studies it suffices to refer a random sample of the women who screen negative on both tests as a way of correcting (or of verifying, to think optimistically) the estimates of sensitivity.

The point of this discussion is that, when screening terms such sensitivity or specificity are mentioned, then the reference standard must be explicitly stated and, if necessary, questioned.

Problem Areas

The major goal of including an introduction to epidemiology in a textbook on gynecological pathology is to encourage pathologists to do epidemiologic studies, and to work with epidemiologists. Accordingly, it may be worth alerting the pathologist to recurrent problem areas that exist at the juncture of the two disciplines. This section informally catalogs a few practical problems that appear to arise most commonly.

Dividing a Spectrum of Disease into Categories

Unfortunately, some epidemiologists may seek out pathologists to perform a service function of "making sure the cases are right," without understanding much about pathology (just as pathologists might seek out statisticians to do a rote data analysis or to figure "how many cases are needed for statistical significance?"). Providing rote pathology review may prove a difficult collaboration, because epidemiologists are prompted by their statistical methods to seek overly simplistic and discrete categorization of disease outcomes. Because the statistical methods for considering a spectrum of disease are difficult to perform and understand, epidemiologists tend to simplify disease measurements into a few (ideally two) reliably distinguished categories, such as "invasive cervical cancer" versus "normal." But as the example of cervical neoplasia demonstrates, diseases may exist as a spectrum of changes that are impossible to divide perfectly into a few categories.

When an epidemiologist asks a pathologist to state whether a slide shows disease (i.e., defines a case) or not (i.e., rejects the case), an uncertain or heavily qualified diagnosis is difficult to force into the study dichotomy. Often, the epidemiologist must subsequently exclude the uncertain diagnoses from the analyses. It is possible to perform a "malicious analysis" in which the uncertain cases are added to the analysis as cases, then reanalyzed as controls, in order to see whether the uncertainty in pathological definition affects the comparisons being made. But too large a proportion of uncertain diagnoses can make an analytic study unreliable.

The collaborating epidemiologist must be willing to understand diagnostic error as a fact of nature and not a failing of pathologists. The pathologist must be willing to sacrifice absolute truth in order to simplify the statistical data to the point of understanding. The limitations of epidemiology should be recognized. As a great physician–epidemiologist once said: "Epidemiology is a butcher shop, don't try to use a scalpel." In other words, epidemiology can only study strong risk associations, because even strong associations are made to appear weak by unavoidable measurement errors and biases. Truly weak associations probably will be missed by all but the largest and luckiest studies. With this in mind, the routine use of pathological qualifiers such as "consistent with" and "cannot exclude" should be abandoned in epidemiologic studies, with the recognition that diagnostic errors will exist (the extent of which should be measured by reliability studies and reported).

The Need for Masking

Epidemiologists tend to mask all data collection as an automatic part of good research technique, to avoid the influence of possible subtle biases that could distort risk estimates. Thus, they do not routinely tell interviewers the disease status of the subjects to minimize bias in questioning, they do not tell laboratory collaborators the identity of specimens until the results are obtained, and they ask pathologists to make their diagnoses with a minimum of information regarding the patients. Epidemiologists are seeking a completely independent decision from pathologists, without influence from previous diagnoses or clinical tests, which often are being studied as risk factors for the current condition. All common statistical tests assume that

the study measurements are completely independent of each other; thus, using any data to influence a decision on another piece of data is wrong.

Pathologists, however, realize that diagnoses are best made in the context of complete information regarding the patient, and that asking for a microscopic diagnosis out of context, as one would demand a lab result from a machine, risks error. Some pathologists incorrectly view the request for masking as a sign of distrust of their intellectual integrity or ability to make an independent decision. The request actually is a sign of epidemiologists' belief that everyone is biased about every decision unless masked. As a revealing example, an epidemiologist's wine tasting group in Maryland covers all labels from the bottles before tasting, and unmasks the results only after the "data" (opinions) are in. Fortunately, it is usually easy for good collaborators to achieve a balance between automatic demands for complete masking, and the kind of complete disclosure of study information that could lead to serious biases.

Standardization of the Scientific Art of Pathology

A more thorny problem arises when epidemiologists challenge the accuracy and reliability of pathologic diagnoses, either as part of a formal pathology agreement study or as part of a larger epidemiologic project. This challenge takes the form of calculation and publication of rates of (dis)agreement between experts, or between the expert and him/herself on different days. The epidemiologist is trained to believe that all biological phenomena are variable and that all measurements of biological phenomena are prone to random error. The pathologist has the weighty daily task of being the final arbiter of disease definition, a responsibility that does not mesh well with error.

The epidemiologist authors have learned something about the world of gynecological pathology only because of the intellectual humility of expert gynecological pathologists (responsible for several of the chapters in this text) whose curiosity outweighed their urges to preserve their national reputations for infallibility. Most of the comparisons performed have related to the cytopathology and histopathology of cervical intrepithelial neoplasia and benign "look-alikes." Agreement rates between expert pathologists have been only fair at best, but have led to a greatly increased understanding of the diagnoses.

If a pathologist ever feels irritated at the demands for reliability studies from new epidemiologist colleagues, it should suffice to ask them when the last time was that they performed a masked logistic regression data analysis in parallel with other epidemiologists, then published the agreement obtained for the relative risk estimates. Since such painful comparative exercises are almost never perpetrated by epidemiologists on themselves, mutual humility and curiosity should reign.

Specimen Adequacy Versus the Bias of Convenience Samples

Epidemiologists seeking to minimize bias are loathe to permit exclusions from a complete series. They suspect that the excluded members of the set will differ from those included in a systematic (biased), rather than random, way. Thus, epidemiologists working with pathologists wish to start their analyses by considering the entire collection of pathological specimens available, winnowing down as needed to usable specimens but always with an eye to possible biases of exclusion that could affect the general applicability of the results. Epidemiologists distrust *convenience samples*, groups of specimens that happen to be available for testing or for review. Pathologists may view the task of defining and retrieving all relevant specimens from their center to be unnecessary. It may be difficult to decide in advance when a convenience sample is sufficient and when a more definitive collection is required. In general, convenience samples are useful for preliminary methodologic work, as in checking if genomic DNA can be amplified from the paraffin blocks available, but such studies cannot be used to reach definitive, generalizable conclusions.

Deciding How Large a Study to Do: Statistical Significance Versus Practicality

Bigger is better for the epidemiologist. It is not much more difficult to do a statistical analysis of 1000 patients than 100; in fact, it is methodologically easier because the numbers are clearly sufficient. But the pathologist collaborator may view it differently. The question of study size is almost always negotiable, in that bigger studies permit the detection of smaller differences, but the critical difference that needs to be detected usually is open to discussion.

There are minimum numbers of subjects that permit epidemiologic analyses. It is impossible to generate a statistically significant result with fewer than five subjects, regardless of how strong an association is. Thirty subjects is another breakpoint. One M.D. epidemiologist's wife reports that when she asked her husband about her rash, he replied: "Give me 30 of them, with the right controls, and I'll venture a guess at what it is." Thirty subjects is a common minimum number in that common statistics such as means start to "behave" more reliably when there are about 30 or more data points. About 200 cases and 200 controls are needed to find reliably a relative risk of about 2.0 (a doubling of risk), given typical prevalences of common exposures. Case-control studies of more than 1000 subjects are relatively rare. Cohort studies, however, often require thousands or even tens of thousands of subjects to generate enough disease end points for analysis. Clinical trials range from small (20 subjects) to large (thousands of subjects) based on the size of the difference being sought.

In general, small studies miss weak associations, do not permit adequate adjustment for confounding, and generate less reliable estimates of risk. Still, many landmark studies of new topics have been small.

The key to defining the proper size of the study is to agree on the hypothesis and the range of expected results. Sample size calculations are very assumption-dependent, and usually demand information not available until the study is completed. Most epidemiologists choose a reasonable number based on cost and time available, then compute the *statistical power* of such a study to detect associations of various strengths. It is standard to require the study to have an 80% or greater chance of finding (as statistically significant) the key disease-exposure association under study, assuming the association truly exists.

Epidemiologic Features of the Major Gynecological Cancers

There have been a number of investigations focusing on the etiology of the major gynecological cancers, with a number of risk factors identified. Both cervical and endometrial cancers have been intensively studied, and for both there have been important recent insights regarding possible mechanisms of action of many of the previously identified risk factors. Fewer advances have been made with respect to ovarian cancer, although several etiologic clues have emerged from recent studies. Less is known about the etiology of cancers of the vulva and vagina and of trophoblastic disease, mainly because their rarity renders epidemiologic investigations difficult.

Cervical Cancer

An estimated 13,500 new cases of cervical cancer were diagnosed in the United States in 1993.[14] The average annual age-adjusted incidence in all SEER areas was 8.8 per 100,000 women for 1990, with a corresponding age-adjusted mortality rate of 3.0. The 5-year survival rate for cervical cancer is 67%, with the rate rising to 88% for cancers diagnosed at early stages.[1]

Substantial decreases in the incidence of invasive cervical cancer have occurred over time in the United States. Among whites, the incidence per 100,000 women declined 75% from 32.6 in the late 1940s to 8.3 in the early 1980s.[53] The decline in incidence in blacks started later than that for whites.

There is about a twofold difference in age-adjusted incidence for invasive cervical cancer for blacks as compared with whites. This differential, although previously observed for all ages, now appears restricted to older women. The incidence also is approximately two times higher for Hispanics and even higher for American Indians. Racial differences also exist in survival experience, with blacks having a 59% 5-year survival rate compared with a 67% 5-year survival rate among whites.

At least some of the racial differences can be explained by strong inverse associations observed between cervical cancer rates and socioeconomic indicators, such as education and income. These relationships prevail among both whites and blacks. When adjustment is made for socioeconomic differences, the excess risk of cervical cancer among blacks is substantially reduced from more than 70% to less than 30%.[51]

There is considerable geographic diversity in cervical cancer rates, with the highest rates having been reported from certain Latin American countries, where rates exceed those of the United States by approximately sixfold.[124] Although rates in the United States are among the lowest in the world, mortality rates are higher in certain Southern areas, particularly in Appalachia.

Recent upturns in incidence and mortality rates among young women have been observed in a number of countries, including Canada, Great Britain, New Zealand, and Australia. Similar increases may be slower to appear in the United States because of the effectiveness of cytology screening programs. The greatest evidence of increased incidence in the United States is for cervical adenocarcinomas among white women 35 to 54 years of age.[54]

Cervical cancer is believed to result from the progression of milder epithelial abnormalities, that is, dysplasia or cervical intraepithelial neoplasia. Support for a continuum of disease is provided by the observations that cervical dysplasia is diagnosed most often among women in their 20s, carcinoma in situ in women in their 30s, and invasive cancer after age 40 years. Because of this presumed continuum, there is little doubt that exfoliative cytology or the Pap smear can have profound effects on morbidity and mortality. The eradication of precursor lesions has preceded significant declines in cervical cancer incidence and mortality in areas where screening has been widespread, such as Kentucky and British Columbia.[16,40] The rates for cervical cancer have not declined in regions with limited screening programs.[88] Case-control studies that have evaluated the role of screening in preventing invasive cervical cancer have found relative risks ranging from 0.2 to 0.4 associated with screening within the last 3 to 5 years.[41,100] In a Finnish study, even patients who had been screened more than 5 years previously had a relative risk of 0.7 compared with those who had never been screened.[128]

Risk factors for cervical cancer are shown in Table 29.2. It is well established that women with either early ages at first sexual intercourse or multiple sexual partners are at an elevated risk.[15,22,77,91,143,166] Although several studies have attempted to determine the independence of these two risk factors (since they usually are highly correlated), both appear to be important predictors. Early age at first intercourse is thought to support a notion of increased suscepti-

Table 29.2. Risk factors for cervical cancers

Factors influencing risk	*Estimated relative risk*[a]
Older ages	2
Residence in certain parts of Latin America, Asia, or Africa	2–6
Lower levels of education or income	2–3
Black, Hispanic, or American Indian	2
Multiparity	2–4
Early ages at first sexual intercourse	2–4
Multiple sexual partners	2–5
Previous episodes of sexually transmitted disease, especially genital herpes and warts	2–10
Long-term smoking	2–4
Long-term oral contraceptive use	1.5–2
No prior regular Pap smear screening	2–6
Diets low in carotene, vitamin C	2–3
Human papillomavirus infection	>20

[a] Relative risks depend on the study and referent group employed.

bility of the cervix during adolescence, and the effect of multiple partners is postulated to operate through an infectious mechanism, in particular HPV infection. In contrast, most investigations have failed to find any influence on risk of frequency of intercourse after adjusting for the effects of these other sexual factors.[15,22,143,166]

Although most attention has focused on the role of female sexual behavior, there has been recent interest in the contribution of the male factor in the etiology of cervical cancer, especially given geographic clusters of cervical and penile cancer rates[62] and elevated rates of cervical cancer among the wives of men with penile cancer.[64,118,160] Several case-control studies have examined the sexual behavior of the husbands of women with cervical cancer, finding that the husbands of these women report more sexual partners than the husbands of controls.[30,33,92,190] The husbands of affected women also were more likely to report prior venereal diseases, early sexual experiences, affairs during marriage, and visits to prostitutes.[33] In contrast, there appears to be no effect on risk of circumcision status of the partner.[15,30,143]

Although there is little evidence that the character of menses affects cervical cancer risk,[15,22,143] there is increasing support for the role of reproductive factors. A number of studies indicate that women with multiple births are at elevated risk.[29,91,131] Possible explanations for the association include cervical trauma during parturition, hormonal or nutritional influences of pregnancy, or an immunologic effect on HPV infection.

The epidemiologic association of cervical neoplasia with sexual risk factors motivated the search for a venereally transmitted agent. Among the agents most widely investigated have been herpes simplex virus type 2 (HSV-2) and HPV. An etiologic role for HSV-2 was suggested on the basis of laboratory findings that HSV-2 can transform cells in culture, that HSV-2 proteins and integrated DNA can be found in some cervical cancers, and that the prevalence of antibody to HSV-2 was generally greater among cervical cancer patients than controls.[120] However, several follow-up studies[106,172] have cast doubt on the association because of the failure to demonstrate significantly increased risks of cervical neoplasia related to HSV-2 serology. Thus, more recent studies have focused on an etiologic role for HPV.

There is now extensive laboratory evidence for the oncogenic potential of HPV.[137] Initial studies demonstrated the presence of viral particles in cervical tissues by electron microscopy and of structural proteins by immunocytochemistry. However, these techniques are relatively insensitive and, because serologic techniques were not available to permit assessment of past exposure to HPV, it was not until the application of DNA hybridization methods that the frequent presence of venereal HPV infection among cervical neoplasia patients was fully recognized. Approximately 85–90% of squamous cervical carcinomas have been found to contain HPV, mainly types 16 and 18.

Only recently have analytic epidemiologic investigations been undertaken to estimate the relative risk of cervical neoplasia associated with HPV infection after adjustment for known risk factors. One of the first studies, which used filter in situ hybridization to test for HPV (a method now generally considered fairly insensitive and nonspecific) found a ninefold increased risk associated with high levels of HPV types 16 and 18 after controlling for other risk factors.[136] In more recent studies, which have used the more valid polymerase chain reaction (PCR) test to assess HPV infection, the estimates have been found to be considerably higher—more than 20.[126,145] The more recent studies, in contradiction to early studies, also support the notion that HPV detection correlates with most suspected risk factors for cervical cancer, including number of sexual partners, use of oral contraceptives, and race.[108] In one recent study, these factors failed to persist as risk predictors after adjustment for HPV, supporting the view that HPV is a central causal factor.[145]

Despite the central etiologic role of HPV in cervical cancer, an effect of other infections as independent or supporting factors has not been dismissed. Most recently, infection with the human immunodeficiency virus (HIV) has been correlated with the detection of HPV-related cytological changes.[116]

In recent years, there have been an increasing number of reports linking cigarette smoking to an elevated risk of both preinvasive and invasive cervical cancer.[184] Although initially the association was thought to reflect the influence of correlated factors, such as increased sexual activity and greater use of oral contraceptives, studies now seem to support an independent role for smoking.[7,27,42,98,157] It is notable that the recent Surgeons' General Report noted cervical cancer as a likely consequence of smoking. In most studies, a history of smoking is associated with about a twofold elevation in risk, with stronger relationships ob-

served among women who report either high intensity or long durations of smoking. The association with smoking is most apparent for squamous cell tumors.[27] Further supporting a biologic mechanism are studies that have demonstrated high levels of smoke-derived nicotine and cotinine in the cervical mucus of smokers.[146] However, the immunosuppressive effects of smoking also should be considered,[8] particularly as an enhancement to the effects of infectious agents (e.g., HPV).

Oral contraceptive use also has emerged recently as a predictor of cervical cancer, particularly invasive disease. Issues of study design and analysis have been complex, generating questions about confounding factors, especially sexual behavior.[18] Even after considering correlated effects, most studies show some evidence of an increased risk, rising to approximately twofold for users of 5 or more years.[24,28] In several studies, higher risks have been observed for adenocarcinomas, in line with descriptive surveys showing increasing rates of this cell type among young women.[39,151]

In a number of studies, users of barrier methods of contraception have been shown to be at a low risk of developing cervical cancer.[15,185] This association usually is attributed to the prevention of spread of infectious agents.

A number of recent studies have suggested a role for dietary factors in the etiology of cervical cancer. Micronutrients associated with protection include vitamin A, beta-carotene, vitamin C, and vitamin E.[9,32,95,117,156,170,174] In addition, it has been suggested that folate deficiency may increase risk, possibly by acting as a cocarcinogen with HPV.[35] Although case-control studies do not generally support an effect of folates on risk,[32,189] the hypothesis deserves further attention, particularly as an explanation to the high risks associated with parity.

Vulvar Cancer

Carcinoma of the vulva is a rare genital neoplasm, with an average annual age-adjusted incidence in all SEER areas during 1985 to 1989 of 1.6 per 100,000 women. The disease occurs primarily in older women.

Cancers of the vulva occur significantly more frequently among women with primary cancers of the cervix, and the two diseases often occur simultaneously.[82,139] Approximately 15–20% of women with vulvar cancer have a second primary cancer occurring simultaneously or nonsimultaneously in the cervix, vagina, or anogenital area. Many patients with vulvar cancer have multifocal genital lesions, commonly including a mixture of condyloma acuminatum and intraepithelial neoplasia, with evidence that these multifocal lesions are HPV-related.[94,124]

Given synchronous occurrences of vulvar and cervical cancers, there has been interest in determining whether cervical cancer risk factors are predictive of vulvar cancer risk. In one study, there appeared to be a strong relationship between the number of reported sexual partners and risk, with women reporting five or more partners being at a two- to threefold excess risk compared with sexually monogamous women.[25] A sexual etiology for vulvar carcinoma derives further support from serologic findings of antibodies against HSV-2 in vulvar cancer patients and identification of HSV-2 in vulvar tumor tissue.[87] More recently, interest has focused on the role of HPV, with a number of studies showing that certain types of vulvar cancers contain HPV DNA whereas other types do not.[2,94,167] This may explain recent time trends in the occurrence of the disease, with rates of in situ vulvar cancer having nearly doubled while rates of invasive disease having remained relatively stable.[164] Several studies suggest that a history of vulvar warts is associated with an elevated risk of vulvar cancer, with the relative risks ranging from 15 to 23.[25,49] In one study, a particularly high risk was associated with multiple episodes of genital warts,[25] possibly reflecting poor immunologic response among these women.

Other suggested, although unconfirmed, risk factors for vulvar cancer include low socioeconomic status, nulliparity and/or late ages at first birth, diabetes, obesity, and hypertension. A more consistently observed risk factor is cigarette smoking,[25,111,127] with one study showing particularly high risks for current smokers.[25] Smokers with a history of genital warts were at especially high risk, possibly supporting the suggestion that the effects of HPV depend on the presence of co-factors.[191] An excess risk of vulvar cancer among users of oral contraceptives was found in one study,[25] but not in another.[127]

Vaginal Cancer

Cancer of the vagina is rare, with an average annual age-adjusted incidence of 0.7 per 100,000 women in the SEER areas for the period from 1985 to 1989. The incidence is approximately three times higher for blacks than for whites, but the reasons for the discrepancy are unknown. About 1000 new cases and 350 deaths from vaginal cancer occur each year in the United States.[1] The average 5-year survival rate is 46% for whites.

Vaginal cancer is primarily a disease of older women, with almost 60% occurring among women 60 years or older. In the past, carcinoma of the vagina was only rarely reported in young girls, but starting in the late 1960s cases of clear cell adenocarcinoma of the vagina, an uncommon cancer at any age, began to be observed with much greater frequency than expected among women between 15 and 22 years of age. Most of these cases have been related to prenatal exposure to diethylstilbestrol (DES).[75]

There are few etiologic clues regarding vaginal cancer, and most clues for this cancer derive from clinical studies. Among factors that have been suggested, trauma to the vagina has received the most attention. Injury to the vagina from wearing ring pessaries (i.e., to support the uterus or

rectum or as a contraceptive device) has been mentioned as a possible carcinogen.[150] The one case-control study of vaginal cancer, based on relatively few cases (n=41), found associations of risk with low socioeconomic status, histories of genital warts or other genital irritation, and previous abnormal Pap smears.[26] Women who had a previous hysterectomy were at high risk, consistent with several clinical observations,[10,163] but in contrast with one analytic study, in which vaginal cancer cases were matched to controls on history of previous dysplasia or neoplasia of the cervix.[76]

Vaginal cancer is found frequently as a synchronous or a metachronous neoplasm with cervical cancer.[139] This has led to the suggestion that there may be shared etiologic features between vaginal and cervical cancers. Recent reports of the co-existence of condylomatous lesions with vaginal cancer and the existence of HPV antigens and DNA in preinvasive and invasive vaginal tumors provide further support to this notion.

Table 29.3. Risk factors for endometrial cancers

Factors influencing risk	*Estimated relative risk*[a]
Older ages	2–3
Residence in North America, Northern Europe	3–18
Higher levels of education or income	1.5–2
White race	2
Nulliparity	3
History of infertility	2–3
Menstrual irregularities	2
Early ages at menarche	1.5–2
Late ages at natural menopause	2–3
Long-term use or use of high doses of menopausal estrogens	10–20
Use of oral contraceptives	0.3–0.5
Stein–Leventhal disease or estrogen-producing tumors	>5
Histories of diabetes, hypertension, gallbladder disease, or thyroid disease	1.5–3
Cigarette smoking	0.5

[a] Relative risks depend on the study and referent group employed.

Endometrial Cancer

Cancer of the endometrium is the most common invasive gynecological cancer and the fourth most frequently diagnosed cancer among American women today. An estimated 31,000 new cases were diagnosed in the United States in 1993.[14] The average annual age-adjusted incidence for corpus and not otherwise specified uterine cancers from the SEER Program, was 21.5 per 100,000 women for 1990; the age-adjusted mortality rate for corpus cancers was 3.5 per 100,000, reflecting the relatively good prognosis for this cancer.

Endometrial cancer rates are highest in North America and Northern Europe; intermediate in Israel, Southern Europe, and Latin America; and lowest in Asia and Africa.[124] The disease is rare before the age of 45 years, but the risk rises sharply among women in their late 40s to middle 60s. The age-adjusted incidence for whites is approximately twice as high as for nonwhites, with reasons for the discrepancy largely undefined. Within the last several decades in the United States, a dramatic change in the incidence pattern for endometrial cancer has occurred, characterized by a marked increase that peaked about 1975.[173] This rise and fall has been linked with the widespread use of estrogen replacement therapy in the late 1960s and early 1970s. Recent reports show a continued decline in endometrial cancer incidence rates since 1979, despite an upswing in the use of menopausal estrogens to prevent osteoporosis and cardiovascular disease.[134] This may reflect the increased use since 1980 of progestins with estrogen replacement therapy to offset the adverse effects of unopposed estrogens.

Table 29.3 shows the variety of the risk factors and the range of magnitude of relative risks that have been identified for endometrial cancer. Many of these risk factors support a central role for estrogenic factors in the etiology of the disease. Apart from age and residency in North America or Europe, the most consistent and strongest risk factors include nulliparity, menstrual irregularities (including late ages at menopause), and obesity. Most studies demonstrate approximately a threefold or greater excess risk for nulliparous than parous women, and decreases in risk with increasing parity.[57,74,101] There is some indication that the association with nulliparity may reflect prolonged periods of infertility,[59,74,169] especially if anovulation was involved.[59] Although early ages at menarche have been inconsistently related to endometrial cancer risk,[58,74,88,169] several studies have noted associations with various measures of menstrual irregularity, including amenorrhea and heavy bleeding.[20,74,187] Most endometrial cancers are diagnosed postmenopausally, and among these women late ages at natural menopause are associated with increased risk, on the order of a twofold greater risk for those with menopause after the age of 52 as compared with before age 49.[115] Obesity also is a well-recognized risk factor for endometrial cancer, with as much as 25% of the disease possibly explained by this factor.[57,81,89,96,178] Obesity (as measured either by total weight or a measure of weight adjusted for height) appears to affect both premenopausal as well as postmenopausal onset disease, and there is some indication that very heavy women appear to have a disproportionately high risk.[20,187] Although one study has suggested that adolescent obesity might be the most important risk factor,[12] other studies indicate that weight gain later in life is the stronger predictor.[74,154,187] In addition, recent interest has focused on the distribution of body fat, with several studies showing that those whose fat distributes on either the trunk or upper body are at highest risk.[5,55,154]

Despite the fact that obesity has been consistently re-

lated to endometrial cancer risk, few studies have focused on the etiologic role of dietary factors. Geographic correlations between per capita fat intake and endometrial cancer incidence[4] have raised interest in this dietary effect, with there being some support from analytic studies for the association.[96] Studies in vegetarians suggest that dietary factors may affect endometrial cancer risk through modifications in hormone metabolism.[3] Alterations in endogenous hormones also have been offered as the explanation for the reduced risk of endometrial cancer observed among regular consumers of alcoholic beverages,[175] although this association has not been consistently observed.[96]

Endometrial cancer also has been noted to be elevated in women with histories of certain diseases (e.g., diabetes, hypertension, thyroid disease),[57,115] but until recently the extent to which the association might reflect correlations with obesity was unclear. Of interest therefore was a recent study[20] that showed that associations with most diseases did not persist after adjustment for body mass; the relationship with diabetes, however, did appear independent, possibly reflecting the role of associated hormonal alterations.

As previously noted, use of menopausal estrogens has been found associated with a 2- to 12-fold elevation in risk.[23,65,81,113,144,152,179] In most investigations, the increased risk was not observed until the drugs were used for at least 2 to 3 years, and longer use of estrogens was generally associated with higher risk.[23,65,81,144,179] The highest relative risks have been observed after 10 years of use, reaching risks of approximately 10 to 20. In most studies, cessation of use appears associated with a relatively rapid decrease in risk, continuous use is associated with higher risks than cyclic administration, and use of preparations with higher estrogen doses imparts the highest risks, although these relationships have not been consistently observed.[23,65,113,144,152,179]

Studies also have attempted to identify whether certain subgroups of estrogen users may be more adversely affected, with there being some evidence that effects are strongest among women who are thin, nondiabetic, or normotensive.[81,113,159] In addition, a recent study showed that cigarette smokers may be more adversely affected by estrogen use than nonsmokers.[23] These findings suggest that estrogen metabolism differs in these groups of women. Alternatively, risk already may be high enough in obese, hypertensive, or diabetic women that exposure to exogenous estrogens has only a small additional effect. Furthermore, it has been shown that estrogen use predisposes toward tumors that demonstrate favorable characteristics, including earlier stages at diagnosis, lower grade, and fewer instances of myometrial invasion.[65,81,89,113,144] Estrogen users also tend to be younger at diagnosis than patients who have not used estrogens, and the tumors are accompanied more frequently by hyperplasia or adenomyosis.[57,159]

Further evidence for the role of exogenous hormones in the pathogenesis of endometrial cancer derives from studies that have demonstrated significantly high risks in users of sequential oral contraceptives (i.e., containing a high dose of estrogen and a weak progestin)[74,178] and significantly low risks of endometrial cancer in women using estrogen-progestin combination pills.[37,80,86,178] Users of combination oral contraceptives experience about half the risk of nonusers, and long-term users have even further reductions in risk. Recent evidence suggests that the reduced risk associated with use of oral contraceptives relates more to the dose of the progestin rather than to the dose of the estrogen in the oral contraceptives.[142] The extent to which reduced risk persists after discontinuation remains an area of controversy.[86,178] In several studies, the protective effect of the pill appears greatest among nulliparous women, nonobese subjects or those who have not used menopausal estrogens.[37,74,178]

Several recent studies suggest that smokers are at a reduced risk of endometrial cancer.[7,19,105,107,169] In most of these studies, a gradient of decreasing risk with increasing amounts smoked has not been observed, but current smokers are at the lowest risk, leading to the suggestion that smoking may affect risk through an alteration in the absorption, distribution or metabolism of hormones.[19]

Women of upper socioeconomic status also have been reported to be a higher risk of endometrial cancer.[58,89] Findings related to socioeconomic status may be partially explained by other endometrial cancer risk factors correlated with affluence (e.g., overnutrition or use of replacement estrogens).

Genetic factors also may explain a small proportion of endometrial cancer, since some women present with Lynch syndrome II, a familial aggregation of endometrial, ovarian, and colon cancers.[110]

Ovarian Cancer

Ovarian cancer accounts for 4% of all cancers in women, with approximately 1 in 70 American women developing the disease in their lifetimes.[1] The average annual age-adjusted incidence for all SEER areas during 1990 was 14.9 per 100,000 women, with an estimated 22,000 new cases having been diagnosed in 1993. A relative survival of 85% can be achieved if ovarian cancer is diagnosed early, but usually the disease is not detected until it has reached an advanced stage, which imposes a high fatality rate (38% 5-year survival rate).

Ovarian cancer rates are high in North America and Northern Europe and low in Japan.[124] White women had considerably higher rates of ovarian cancer than blacks, but there is evidence that this difference may be narrowing.[52]

Table 29.4 shows the identified risk factors for ovarian cancer. Although the incidence of ovarian cancer increases with age, there is a marked flattening in the age-specific

Table 29.4. Risk factors for ovarian cancers

Factors influencing risk	*Estimated relative risk*[a]
Older ages	3
Residence in North America, Northern Europe	2–5
Higher levels of education or income	1.5–2
White race	1.5
Nulligravidity	2–3
History of infertility or use of infertility drugs	2–5
Early ages at menarche	1.5
Late ages at natural menopause	1.5–2
History of a hysterectomy	0.5–0.7
Use of oral contraceptives	0.3–0.5
Perineal talc exposure	1.5–2
Female relative with ovarian cancer	3–4

[a] Relative risks depend on the study and referent group employed.

incidence curves shortly after menopause. Fewer risk factors have been identified for ovarian cancer than for endometrial cancer, although nulligravidity and infertility have been fairly consistent predictors.[13,73,83,180,186] Compared with nulligravidous women, women with a single pregnancy have a relative risk of 0.6 to 0.8, with each additional pregnancy lowering risk by about 10–15%. This derives primarily from associations with number of full-term births, although in several studies, risk also has been found to decrease with increasing number of incomplete pregnancies. In most studies that have adjusted effects of age at first pregnancy by number of pregnancies, no residual effect of age at first pregnancy persisted.[73,78,83,186] There is, however, support from several recent studies for a reduced risk of ovarian cancer for women who breast feed for extended periods.[13,66,73,180]

The extent to which the relationship of risk to pregnancy history reflects a hazardous role for infertility or a protective role for pregnancy remains unresolved. Studies appear to support a role for infertility in the etiology of ovarian cancer.[13,73,83,181] In a recent combined analysis of data from multiple studies, the relationship with infertility was linked to use of infertility drugs.[180] Although provocative because of the fact that many of the drugs used in the treatment for infertility stimulate ovulation, the relationship requires confirmation from additional studies.

Several studies have reported early ages at menarche and late ages at menopause as ovarian cancer risk factors, although these patterns have not always been observed.[13,61,78,130,186] Numerous studies have noted a reduced risk of ovarian cancer associated with hysterectomy, with the apparent protective effect ranging from 30% to 40%.[13,45,61,71,83,182] This may reflect the opportunity for visualization and removal of abnormal ovaries during surgery, although it is also possible that hysterectomy compromises ovarian function through reduced blood supply to the ovaries.[56]

Similar to endometrial cancer, oral contraceptive use has been linked with a reduced risk of ovarian cancer. A reduction in risk is apparent after only a few months' use, but the apparent protection is greatest among long-term users.[38,141,177,183,186] The reduction in risk appears to persist for a number of years after discontinuation and applies to all histological types of ovarian cancer.

Most studies that have examined the effect of menopausal estrogens on ovarian cancer risk have not found an association.[13,45,61,73,78,85,186] However, in one study, a threefold excess risk of endometrioid ovarian cancers was linked with estrogen use.[176] This relationship was not confirmed in another study.[85]

The role of dietary factors in the etiology of ovarian cancer only recently has been given attention. International data correlating ovarian cancer incidence and per capital fat availability[3] and the increased incidence of ovarian cancer among Japanese migrants to the United States[68] has stimulated interest in the role of dietary fat. Several follow-up studies support an association,[135,138,161] although the relationship has not been consistently observed.[90] Similarly, several case-control studies show a greater consumption of a variety of indicators of fat consumption among cases than controls.[46,97,155] However, other studies do not support the association, but rather show that risk is more dependent on low consumption of fruits and vegetables and foods containing either beta-carotene or vitamin A.[36,158] Cramer and others[43,44] found a high risk of ovarian cancer associated with consumption of lactose-rich dairy products. This association was restricted to women with low levels of galactose-1-phosphate uridyl transferase activity, an enzyme linked with hypergonadotropic hypogonadism.

Familial clusters of ovarian cancer suggest a genetic component. Case- control studies have attempted to estimate the magnitude of the genetic contribution,[93,133,149] with the largest of the studies showing estimated relative risks of 3.6 and 2.9 associated with having a first- and second-degree relative with ovarian cancer, respectively.[149]

Other suggested factors afffecting ovarian cancer risk include talc exposure, a history of mumps infection, and alcohol consumption. Talc exposure, which has been related to an excess risk of ovarian cancer in a number of case-control studies,[47,69,70,182] is of interest biologically in that ovarian cancer is thought to arise from the mesothelium that lines the peritoneal cavity. Ovarian cancer may be analogous, therefore, to pleural mesothelioma, which has been shown to be caused by asbestos, a chemical similar to talc. Mumps infection, which usually has been assessed by history, has been shown to correlate poorly with mumps serology, raising questions about the biological reality of infection history as a true predictor.[121,147] Smoking has not been related to ovarian cancer in most investigations. Some studies have shown slightly lower risks of ovarian cancer among alcohol drinkers.[36,63,67] Coffee drinking, linked to an excess risk of ovarian cancer in several

studies, has not been confirmed as a risk factor in other investigations.[72,99,123,168,182]

Gestational Trophoblastic Disease

Choriocarcinoma is a rare malignancy in the United States, with a recent incidence in all SEER areas of 0.2 per 100,000 women, or approximately 1 per 22,623 live-births.[21] Hydatidiform mole occurs about once in every 1000 pregnancies, and approximately one of six occurrences results in invasion. Trophoblastic disease has been reported to be more common in certain parts of the world, although part of the differences may be due to a variety of selection biases.[17] The epidemiologic study of choriocarcinoma has been complicated by its relative infrequency. Most studies have focused on defining risk factors for hydatidiform or invasive mole, and it is uncertain the extent to which these findings can be extrapolated to choriocarcinoma.

Apart from a history of hydatidiform mole, the most clearly established risk factor for choriocarcinoma and hydatidiform mole is late maternal age. In one study, a 24-fold increased incidence of choriocarcinoma was found for women with a pregnancy after 45 years as compared with those with a pregnancy between 20 and 39 years.[165] Rates of trophoblastic disease also appear to be considerably higher in Asian and African countries, but the true extent of difference from Western rates is difficult to decipher because of variations in reporting practices. One incidence survey in the United States showed that, even after adjustment for birth distribution effects, blacks and other nonwhite races had 2.1- and 1.8-fold greater risks, respectively, than whites.[21]

An association between blood group A and choriocarcinoma has been found in two studies, and the combination of mother's group A and father's group O was considerably higher than expected (10.4-fold risk).[6,50] Blood groups A and AB were associated with elevated risks of hydatidiform mole in one study, although blood group was not predictive of risk in two others.[122,129,165]

In several studies that have adjusted for effects of later maternal age, parous women have remained at substantially reduced risk of hydatidiform mole compared with nulliparous women, with some evidence of further reductions in risk with multiple births.[31,122,132] Several studies found an increased risk associated with the occurrence of a prior spontaneous abortion, although this has not been consistently observed.[31,102] An increased risk of hydatidiform mole was associated with induced abortions, although information was not available on reasons for the terminations.[31] A history of infertility has been suggested in one study, although other studies do not confirm the association.[31,102,132] In one study, Chinese patients reporting use of herbal medicines during the first trimester of a previous pregnancy were at elevated risk.[31]

Low body mass, unrelated to dieting or exercise, has been reported as a risk factor for choriocarcinoma in one study.[34] Patients also had later onset of menarche and lighter menstrual periods than controls, possibly reflecting lower estrogen levels.

Although several studies have found an increased risk of trophoblastic diseases associated with long-term use of oral contraceptives,[11,31,140] others have found no such influence.[102,122] Others have suggested that oral contraceptives may increase the risk of malignant sequelae after mole evacuation through a tumor-stimulating effect.[162,188] In one study, this effect was restricted to users of high-dose estrogens, although in others, there were no such effects.[11,48,188]

Late paternal age has been suggested in one study as a risk factor for trophoblastic disease,[103] although other studies have failed to confirm this.[31,122] Cigarette smoking also has been linked with the occurrence of trophoblastic disease.[102] One study suggested that low carotene intake affected the risk of hydatidiform mole,[11] but no specific dietary associations were observed in another study.[31]

References

1. American Cancer Society (1991): Cancer Facts and Figures
2. Anderson WA, Franquemont DW, Williams J, Taylor PT, Crum CP (1991): Vulvar squamous cell carcinoma and papillomaviruses: Two separate entities. Am J Obstet Gynecol 16S: 329–336
3. Armstrong BK, Brown JB, Clarke HT, et al (1981): Diet and reproductive hormones: A study of vegetarian and non-vegetarian postmenopausal women. JNCI 67: 761–767
4. Armstrong B, Doll R (1975): Environmental factors and cancer incidence and mortality in different countries, with special reference to dietary practices. Int J Cancer 15: 617–631
5. Austin H, Austin MJ Jr, Partridge EE, et al (1991): Endometrial cancer, obesity, and body fat distribution. Cancer Res 51: 568–572
6. Bagshawe KD, Rawlins G, Pike MC, et al (1971): ABO blood groups in trophoblastic neoplasia. Lancet 1: 553–557
7. Baron JA, Byers T, Greenberg ER, et al (1986): Cigarette smoking in women with cancers of the breast and reproductive organs. JNCI 77: 677–680
8. Barton SE, Jenkins D, Cuzick J, et al (1988): Effect of cigarette smoking on cervical epithelial immunity: A mechanism for neoplastic change? Lancet 2: 652–654
9. Basu J, Palan P, Vermund S, et al (1991): Plasma ascorbic acid and beta-carotene levels in women evaluated for HPV infection, smoking, and cervix dysplasia. Cancer Detection Prev 15: 165–170
10. Bell J, Sevin B, Averette H, et al (1984) Vaginal cancer after hysterectomy for benign disease: Value of cytologic screening. Obstet Gynecol 54: 699–701
11. Berkowitz RS, Cramer DW, Bernstein MR, et al (1985) Risk factors for complete molar pregnancy from a case-control study. Am J Obstet Gynecol 152: 1016–1020
12. Blitzer PH, Blitzer EC, Rimm AA (1976): Association be-

tween teen-age obesity and cancer in 56,111 women: All cancers and endometrial carcinoma. Prev Med 5: 20–31
13. Booth M, Beral V, Smith P (1989) Risk factors for ovarian cancer: A case-control study. Br J Cancer 60: 592–598
14. Boring CC, Squires TS, Tong T (1993): Cancer statistics, 1993. CA-A Ca J for Clin 43: 7–26
15. Boyd JT, Doll R (1964): A study of the aetiology of carcinoma of the cervix uteri. Br J Cancer 18: 419–434
16. Boyes DA (1981) The value of a pap smear program and suggestions for its implementation. Cancer 48: 613–621
17. Bracken MB, Brinton LA, Hayashi K (1984): Epidemiology of hydatidiform mole and choriocarcinoma. Epidemiol Rev 6: 52–75
18. Brinton LA (1991): Oral contraceptives and cervical neoplasia. Contraception 43: 581–595
19. Brinton LA, Barrett RJ, Berman ML, et al (1993): Cigarette smoking and the risk of endometrial cancer. Am J Epidemiol 137: 281–291
20. Brinton LA, Berman ML, Mortel R, et al (1992): Reproductive, menstrual, and medical risk factors for endometrial cancer: Results from a case-control study. Am J Obstet Gynecol 167: 1317–1325
21. Brinton LA, Bracken MB, Connelly RR (1986): Choriocarcinoma incidence in the United States. Am J Epidemiol 123: 1094–1100
22. Brinton LA, Hamman RF, Huggins GR, et al (1987): Sexual and reproductive risk factors for invasive squamous cell cervical cancer. JNCI 79: 23–30
23. Brinton LA, Hoover RN, and the Endometrial Cancer Collaborative Group (1993) Estrogen replacement therapy and endometrial cancer risk: Unresolved issues. Obstet Gynecol 81: 265–271
24. Brinton LA, Huggins GR, Lehman HF, et al (1986): Long-term use of oral contraceptives and risk of invasive cervical cancer. Int J Cancer 38: 339–344
25. Brinton LA, Nasca PC, Mallin K, et al (1990): Case-control study of cancer of the vulva. Obstet Gynecol 75: 859–866
26. Brinton LA, Nasca PC, Mallin K, et al (1990): Case-control study of in situ and invasive carcinoma of the vagina. Gynecol Oncol 38: 49–54
27. Brinton LA, Schairer C, Haenszel W, et al (1986): Cigarette smoking and invasive cervical cancer. JAMA 255: 3265–3269
28. Brinton LA, Reeves WC, Brenes MM, et al (1990): Oral contraceptive use and risk of invasive cervical cancer. Int J Epidemiol 19: 4–11
29. Brinton LA, Reeves WC, Brenes MM, et al (1989): Parity as a risk factor for cervical cancer. Am J Epidemiol 130: 486–496
30. Brinton LA, Reeves WC, Brenes MM, et al (1989) The male factor in the etiology of cervical cancer among sexually monogamous women. Int J Cancer 44: 199–203
31. Brinton LA, Wu B, Wang W, et al (1989) Gestational trophoblastic disease: A case-control study from the People's Republic of China. Am J Obstet Gynecol 161: 121–127
32. Brock KE, Berry G, Mock PA, et al (1988) Nutrients in diet and plasma and risk of in situ cervical cancer. JNCI 80: 580–585
33. Buckley JD, Harris RWC, Doll R, et al (1981) Case-control study of the husbands of women with dysplasia or carcinoma of the cervix uteri. Lancet 2: 1010–1015
34. Buckley JD, Henderson BE, Morrow CP, et al (1988): Case- control study of gestational choriocarcinoma. Cancer Res 48: 1004–1010
35. Butterworth CE Jr, Hatch KD, Soong S-J, et al (1992): Oral folic acid supplementation for cervical dysplasia: A clinical intervention trial. Am J Obstet Gynecol 166: 803–809
36. Byers T, Marshall J, Graham S, et al (1983): A case-control study of dietary and nondietary factors in ovarian cancer. JNCI 71: 681–686
37. Cancer and Steroid Hormone Study of the Centers for Disease Control and the National Institute of Child Health and Human Development (1987): Combination oral contraceptive use and the risk of endometrial cancer. JAMA 257: 796–800
38. Cancer and Steroid Hormone Study of the Centers for Disease Control and the National Institute of Child Health and Human Development (1987) The reduction in risk of ovarian cancer associated with oral contraceptive use. N Engl J Med 316: 650–655
39. Chilvers C, Mant D, Pike MC (1987): Cervical adenocarcinoma and oral contraceptives. Br Med J 295: 1446–1447
40. Christopherson WM, Parker JE, Mendez WM, et al (1970): Cervix cancer death rates and mass cytologic screening. Cancer 26: 808–811
41. Clarke EA, Anderson TW (1979): Does screening by "pap" smears help prevent cervical cancer? A case-control study. Lancet 2: 1–4
42. Clarke EA, Morgan RW, Newman AM (1982): Smoking as a risk factor in cancer of the cervix: Additional evidence from a case-control study. Am J Epidemiol 115: 59–66
43. Cramer DW (1989): Lactase persistence and milk consumption as determinants of ovarian cancer risk. Am J Epidemiol 130: 904–910
44. Cramer DW, Harlow BL, Willett WC, et al (1989): Galactose consumption and metabolism in relation to the risk of ovarian cancer. Lancet 2: 66–71
45. Cramer DW, Hutchison GB, Welch WR, et al (1983): Determinants of ovarian cancer risk. I. Reproductive experiences and family history. JNCI 71: 711–716
46. Cramer DW, Welch WR, Hutchison BG, et al (1984) Dietary animal fat in relation to ovarian cancer risk. Obstet Gynecol 63: 833–838
47. Cramer DW, Welch WR, Scully RE, et al (1982) Ovarian cancer and talc. A case-control study. Cancer 50: 372–376
48. Curry SL, Schlaerth JB, Kohorn EI, et al (1989) Hormonal contraception and trophoblastic sequelae after hydatidiform mole (a Gynecologic Oncology Group study). Am J Obstet Gynecol 160: 805–811
49. Daling JR, Chu J, Weiss NS, et al (1984): The association of condylomata acuminata and squamous carcinoma of the vulva. Br J Cancer 50: 533–535
50. Dawood MY, Teoh ES, Ratnam SS (1971): ABO blood group in trophoblastic diseases. J Obstet Gynaecol Br Commonw 78: 918–923
51. Devesa SS, Diamond EL (1980): Association of breast cancer and cervical cancer incidences with income and education among whites and blacks. JNCI 65: 515–528
52. Devesa SS, Silverman DT (1978) Cancer incidence and mortality trends in the United States: 1935–74. JNCI 60: 545–571

53. Devesa SS, Silverman DT, Young JL Jr, et al (1987) Cancer incidence and mortality trends among whites in the Unites States, 1947–84. JNCI 79: 701–770
54. Devesa SS, Young JL Jr, Brinton LA, et al (1989) Recent trends in cervix uteri cancer. Cancer 64: 2184–2190
55. Elliott EA, Matanowski GM Rosenshein NB, et al (1990) Body fat patterning in women with endometrial cancer. Gynecol Oncol 39: 253–258
56. Ellsworth LR, Allen HH, Nisker JA (1983): Ovarian function after radical hysterectomy for stage IB carcinoma of the cervix. Am J Obstet Gynecol 145: 185–188
57. Elwood JM, Boyes DA (1980) Clinical and pathological features and survival of endometrial cancer patients in relation to prior use of estrogens. Gynecol Oncol 10: 173–187
58. Elwood JM, Cole P, Rothman KJ, et al (1977): Epidemiology of endometrial cancer. JNCI 59: 1055–1060
59. Escobedo LG, Lee NC, Peterson HB, et al (1991) Infertility-associated endometrial cancer risk may be limited to specific subgroups of infertile women. Obstet Gynecol 77: 124–128
60. Fleiss JL (1981): *Statistical Methods for Rates and Proportions, 2nd ed.* New York, John Wiley and Sons
61. Franceschi S, La Vecchia C, Helmrich SP, et al (1982): Risk factors for epithelial ovarian cancer in Italy. Am J Epidemiol 115: 715–719
62. Franco EL, Filho NC, Villa LL, et al (1988): Correlation patterns of cancer relative frequencies with some socieconomic and demographic indicators in Brazil: An ecologic study. Int J Cancer 41: 24–29
63. Franks AL, Lee NC, Kendrick JS, et al (1987) Cigarette smoking and the risk of epithelial ovarian cancer. Am J Epidemiol 126: 112–117
64. Graham S, Priore R, Graham M, et al (1979): Genital cancer in wives of penile cancer patients. Cancer 44: 1870–1874
65. Gray LA, Christopherson WM, Hoover RN (1977): Estrogens and endometrial carcinoma. Obstet Gynecol 49: 385–389
66. Gwinn ML, Lee NC, Rhodes PH, et al (1990) Pregnancy, breast feeding, and oral contraceptives and the risk of epithelial ovarian cancer. J Clin Epidemiol 43: 559–568
67. Gwinn ML, Webster LA, Lee NC, et al (1986): Alcohol consumption and ovarian cancer risk. Am J Epidemiol 123: 759–766
68. Haenszel W, Kurihara M (1968) Studies of Japanese migrants. I. Mortality from cancer and other diseases among Japanese in the United States. JNCI 40: 43–68
69. Harlow BL, Cramer DW, Bell DA, et al (1992): Perineal exposure to talc and ovarian cancer risk. Obstet Gynecol 80: 19–26
70. Hartge P, Hoover R, Lesher LP, et al (1983): Talc and ovarian cancer (Letter). JAMA 250: 1844
71. Hartge P, Hoover R, McGowan L, et al (1988): Menopause and ovarian cancer. Am J Epidemiol 127: 990–998
72. Hartge P, Lesher LP, McGowan L, et al (1982): Coffee and ovarian cancer (Letter). Int J Cancer 30: 531
73. Hartge P, Schiffman MH, Hoover R, et al (1989): A case-control study of epithelial ovarian cancer. Am J Obstet Gynecol 161: 10–16
74. Henderson BE, Casagrande JT, Pike MC, et al (1983): The epidemiology of endometrial cancer in young women. Br J Cancer 47: 749–756
75. Herbst AL, Anderson S, Hubby MM, et al (1986): Risk factors for the development of diethylstilbestrol-associated clear cell adenocarcinoma: A case-control study. Am J Obstet Gynecol 154: 814–822
76. Herman JM, Homelsey HD, Dignan MB (1986): Is hysterectomy a risk factor for vaginal cancer? JAMA 256: 601–603
77. Herrero R, Brinton LA, Reeves WC, et al (1990) Sexual behavior, venereal diseases, hygiene practices, and invasive cervical cancer in a high- risk population. Cancer 65: 380–386
78. Hildreth NG, Kelsey JL, LiVolsi VA, et al (1981): An epidemiologic study of epithelial carcinoma of the ovary. Am J Epidemiol 114: 398–405
79. Hsiao JK, Bartko JJ, and Potter WZ (1989): Diagnosing diagnoses: Receiver operating characteristic methods and psychiatry. Arch Gen Psychiatry 46:664–667
80. Hulka BS, Chambless LE, Kaufman DG, et al (1982) Protection against endometrial carcinoma by combination-product oral contraceptives. JAMA 247: 475–477
81. Hulka BS, Fowler WC, Kaufam DG, et al (1980): Estrogen and endometrial cancer: Cases and two control groups from North Carolina. Am J Obstet Gynecol 137: 92–101
82. Jimerson GK, Merrill JA (1970): Multicentric squamous malignancy involving both cervix and vulva. Cancer 26: 150–153
83. Joly DJ, Lilienfeld AM, Diamond EL, et al (1974): An epidemiologic study of the relationship of reproductive experience to cancer of the ovary. Am J Epidemiol 99: 190–209
84. Kahn HA (1983): *An Introduction to Epidemiologic Methods.* New York, Oxford University Press
85. Kaufman DW, Kelly JP, Welch WR, et al (1989): Noncontraceptive estrogen use and epithelial ovarian cancer. Am J Epidemiol 130: 1142–1151
86. Kaufman DW, Shapiro S, Slone D, et al (1980): Decreased risk of endometrial cancer among oral-contraceptive users. N Engl J Med 303: 1045–1047
87. Kaufman RH, Bornsteing J, Adam E, et al (1988): Human papillomavirus and herpes simplex virus in vulvar squamous cell carcinoma in situ. Am J Obstet Gynecol 158: 862–871
88. Kelsey JL, Hildreth NG (eds) (1983): Breast and Gynecologic Cancer Epidemiology. Boca Raton, FL, CRC Press
89. Kelsey JL, LiVolsi VA, Holford TR, et al (1982): A case-control study of cancer of the endometrium. Am J Epidemiol 116: 333–342
90. Kinlen LJ (1982): Meat and fat consumption and cancer mortality: A study of strict religious orders in Britain. Lancet 1: 946–949
91. Kjaer SK, Dahl C, Engholm G, et al (1992) Case-control study of risk factors for cervical neoplasia in Denmark. II. Role of sexual activity, reproductive factors, and venereal infections. Cancer Causes Control 3: 339–348
92. Kjaer SK, de Villiers E-M, Dahl C, et al (1991) Case-control study of risk factors for cervical neoplasia in Denmark: I. Role of the 'male factor' in women with one lifetime sexual partner. Int J Cancer 48: 39–44
93. Koch M, Gaedke H, Jenkins H (1989) Family history of ovarian cancer patients: A case-control study. Int J Epidemiol 18: 782–785
94. Kurman RJ, Toki T, Schiffman MH (1993) Basaloid and

warty carcinomas of the vulva. Am J Surg Pathol 17: 133–145
95. La Vecchia C, Decarli A, Fasoli M, et al (1988) Dietary vitamin A and the risk of intraepithelial and invasive neoplasia. Gynecol Oncol 30: 187–195
96. La Vecchia C, Decarli A, Fasoli M, et al (1986) Nutrition and diet in the etiology of endometrial cancer. Cancer 57: 1248–1253
97. La Vecchia C, Decarli A, Negri E, et al (1987): Dietary factors and the risk of epithelial ovarian cancer. JNCI 79: 663–669
98. La Vecchia C, Franceschi S, Decarli A, et al (1986) Cigarette smoking and the risk of cervical neoplasia. Am J Epidemiol 123: 22–29
99. La Vecchia C Franceschi S, Decarli A, et al (1984): Coffee drinking and the risk of epithelial ovarian cancer. Int J Cancer 33: 559–562
100. La Vecchia C, Franceschi S, Decarli A, et al (1984) "Pap" smear and the risk of cervical neoplasia. Quantitative estimates from a case- control study. Lancet 2: 779–782
101. La Vecchia C, Franceschi S, Decarli A, et al (1984) Risk factors for endometrial cancer at different ages. JNCI 73: 667–671
102. La Vecchia C, Franceschi S, Parazzini F, et al (1985) Risk factors for gestational trophoblastic disease in Italy. Am J Epidemiol 121: 457–464
103. La Vecchia C, Parazzini F, Decarli A, et al (1984) Age of parents and risk of gestational trophoblastic disease. JNCI 73: 639–642
104. Last JM, Abramson JH, Greenland S, et al (1983): *A Dictionary of Epidemiology*. New York, Oxford University Press
105. Lawrence C, Tessaro I, Durgerian S, et al (1987): Smoking, body weight, and early-stage endometrial cancer. Cancer 49: 1665–1669
106. Lehtinen M, Hakama M, Aaran R-K, et al (1992) Herpes simplex virus type 2 infection and cervical cancer: A prospective study of 12 years of follow-up in Finland. Cancer Causes Control 3: 333–338
107. Lesko SM, Rosenberg L, Kaufman DW, et al (1985) Cigarette smoking and the risk of endometrial cancer. N Engl J Med 313: 593–596
108. Ley C, Bauer HM, Reingold R, et al (1991): Determinants of genital human papillomavirus infection in young women. J Natl Cancer Inst 83: 997–1003
109. Lilienfeld AM, Lilienfeld DE (1980): *Foundations of Epidemiology, 2 ed.* New York, Oxford University Press
110. Lynch HT, Cavalieri RJ, Lynch JF, et al (1992): Gynecologic cancer clues to Lynch syndrome II diagnosis: A family report. Gynecol Oncol 44: 198–203
111. Mabuchi K, Bross DS, Kessler II (1985): Epidemiology of cancer of the vulva. A case-control study. Cancer 55: 1843–1848
112. Maclure M and Willett WC (1987): Misinterpretation and misuse of the kappa statistic. Am J Epidemiol 126: 161–169
113. Mack TM, Pike MC, Henderson BE, et al (1976): Estrogens and endometrial cancer in a retirement community. N Engl J Med 294: 1262–1267
114. MacMahon B and Pugh TF (1970): *Epidemiology: Principles and Methods*. Boston, Little, Brown, and Co
115. MacMahon B (1974) Risk factors for endometrial cancer: Gynecol Oncol 2: 122–129
116. Mandelblatt JS, Fahs M, Garibaldi K, et al (1992) Association between HIV infection and cervical neoplasia: Implications for clinical care of women at risk of both conditions. AIDS 6: 173–178
117. Marshall JR, Graham S, Byers T, et al (1983) Diet and smoking in the epidemiology of cancer of the cervix. JNCI 70: 847–851
118. Martinez I (1969) Relationship of squamous cell carcinoma of the cervix uteri to squamous cell carcinoma of the penis among Puerto Rican women married to men with penile carcinoma. Cancer 24: 777–780
119. Mausner AK, Bahn JS (1974): *Epidemiology: An Introductory Text*. Philadelphia, W. B. Saunders
120. McDougall JK, Beckmann AM, Galloway DA (1986): The enigma of viral nucleic acids in genital neoplasia. In Peto R, zur Hausen H (eds): Viral Etiology of Cervical Cancer. Banbury Report 21, pp 199–210. New York, Cold Spring Harbor Laboratory
121. Menczer J, Modan M, Ranon L, et al (1979) Possible role of mumps virus in the etiology of ovarian cancer. Cancer 43: 1375–1379
122. Messerli ML, Lilienfeld AM, Parmley T, et al (1985) Risk factors for gestational trophoblastic neoplasia. Am J Obstet Gynecol 153: 294–300
123. Miller DR, Rosenberg L, Kaufman DW, et al (1987) Epithelial ovarian cancer and coffee drinking. Int J Epidemiol 16: 13–17
124. Mitchell MF, Prasad CJ, Silva EG, et al (1993) Second genital primary squamous neoplasms in vulvar carcinoma: Viral and histopathologic correlates. Obstet Gynecol 81: 13–18
125. Muir C, Waterhouse J, Mack T, et al (eds) (1987) Cancer Incidence in Five Continents, Vol V. Lyon, France, IARC Scientific Publication 88
126. Munoz N, Bosch FX, Sanjose S, et al (1992) The causal link between human papillomavirus and invasive cervical cancer: A population-based case-control study in Colombia and Spain. Int J Cancer 52: 743–749
127. Newcomb PA, Weiss NS, Daling J (1984): Incidence of vulvar carcinoma in relation to menstrual, reproductive, and medical factors. JNCI 73: 391–396
128. Olesen F (1988) A case-control study of cervical cytology before diagnosis of cervical cancer in Denmark. Int J Epidemiol 17: 501–508
129. Parazzini F, La Vecchia C, Franceschi S, et al (1985) ABO blood groups and the risk of gestational trophoblastic disease. Tumori 71: 123–126
130. Parazzini F, La Vecchia C, Negri E, et al (1989) Menstrual factors and the risk of epithelial ovarian cancer. J Clin Epidemiol 42: 443–448
131. Parazzini F, La Vecchia C, Negri E, et al (1989) Reproductive factors and the risk of invasive and intraepithelial cervical neoplasia. Br J Cancer 59: 805–809
132. Parazzini F, La Vecchia C, Pampallona S, et al (1985) Reproductive patterns and the risk of gestational trophoblastic disease. Am J Obstet Gynecol 152; 866–870
133. Parazzini F, Negri E, La Vecchia C, et al (1992) Family history of reproductive cancers and ovarian cancer risk: An Italian case-control study. Am J Epidemiol 135: 35–40

134. Persky V, Davis F, Barrett R, et al (1990) Recent time trends in uterine cancer. Am J Public Health 80: 935–939
135. Phillips RL, Garfinkel L, Kuzma JW, et al (1980) Mortality among California Seventh-Day Adventists for selected cancer sites. JNCI 65: 1097–1107
136. Reeves WC, Brinton LA, Garcia M, et al (1989) Human papillomavirus (HPV) infection and cervical cancer in Latin America. N Engl J Med 320: 1436–1441
137. Reeves WC, Rawls WE, Brinton LA (1989): Epidemiology of genital papillomaviruses and cervical cancer. Rev Infect Dis 11: 426–439
138. Rose DP, Boyar AP (1985) Diet and ovarian cancer. JAMA 254: 2553
139. Rose PG, Herterick EE, Boutselis JG, et al (1987) Multiple primary gynecologic neoplasms. Am J Obstet Gynecol 157: 261–267
140. Rosenberg L, Shapiro S, Slone D, et al (1989) Gestational trophoblastic disease and use of oral contraceptives (Letter). Am J Obstet Gynecol 161: 1087–1088
141. Rosenberg L, Shapiro S, Slone D, et al (1982) Epithelial ovarian cancer and combination oral contraceptives. JAMA 247: 3210–3212
142. Rosenblatt KA, Thomas DB, and the WHO Collaborative Study of Neoplasia and Steroid Contraceptives (1991) Hormonal content of combined oral contraceptives in relation to the reduced risk of endometrial carcinoma. The WHO Collaborative Study of Neoplasia and Steroid Contraceptives. Int J Cancer 49: 870–874
143. Rotkin ID (1967) Adolescent coitus and cervical cancer: Associations of related events with increased risk. Cancer Res 27: 603–617
144. Rubin GL, Peterson HB, Lee NC, et al (1990) Estrogen replacement therapy and the risk of endometrial cancer: Remaining controversies. Am J Obstet Gynecol 162: 148–154
145. Schiffman MH, Bauer HM, Hoover RN, et al (1993) Epidemiologic evidence showing that human papillomavirus infection causes most cervical intraepithelial neoplasia. JNCI 85: 958–964
146. Schiffman MH, Haley NJ, Felton JS, et al (1987) Biochemical epidemiology of cervical neoplasia: Measuring cigarette smoke constitutents in the cervix. Cancer Res 47: 3886–3888
147. Schiffman MH, Hartge P, Lesher LP, et al (1985) Mumps and postmenopausal ovarian cancer (Letter). Am J Obstet Gynecol 152: 116–117
148. Schiffman MH, Schatzkin A (1994) Test reliability is critically important to molecular epidemiology: An example from studies of HPV infection and cervical neoplasia. Cancer Res
149. Schildkraut JM, Thompson WD (1988) Familial ovarian cancer: A population-based case-control study. Am J Epidemiol 128: 456–466
150. Schraub S, Sun XS, Maingon PH, et al (1992) Cervical and vaginal cancer associated with pessary use. Cancer 69: 2505–2509
151. Schwartz SM, Weiss NS (1986) Increased incidence of adenocarcinoma of the cervix in young women in the United States. Am J Epidemiol 124: 1045–1047
152. Shapiro S, Kaufman DW, Slone D, et al (1980) Recent and past use of conjugated estrogens in relation to adenocarcinoma of the endometrium. N Engl J Med 303: 485–489
153. Sherman ME, Schiffman MH, Lorincz AT, et al (1994) Towards objective quality assurance in cervical cytopathology: correlation of cytopathologic diagnoses with detection of high-risk HPV types. J Clin Pathol, Aug
154. Shu XO, Brinton LA, Zheng W, et al (1992) Relation of obesity and body fat distribution to endometrial cancer in Shanghai, China. Cancer Res 52: 3865–3870
155. Shu XO, Gao GT, Yuan JM, et al (1989) Dietary factors and epithelial ovarian cancer. Br J Cancer 59: 92–96
156. Slattery ML, Abbott TM, Overall JC Jr, et al (1990) Dietary vitamins A, C, and E and selenium as risk factors for cervical cancer. Epidemiology 1: 8–15
157. Slattery ML, Robison LM, Schuman KL, et al (1989) Cigarette smoking and exposure to passive smoke are risk factors for cervical cancer. JAMA 261: 1593–1598
158. Slattery ML, Schuman KL, West DW, et al (1989) Nutrient intake and ovarian cancer. Am J Epidemiol 130: 497–501
159. Smith DC, Prentice R, Thompson DJ, et al (1975) Association of exogenous estrogen and endometrial carcinoma. N Engl J Med 293: 1164–1167
160. Smith PG, Kinlen LJ, White GC, et al (1980) Mortality of wives of men dying with cancer of the penis. Br J Cancer 41: 442–448
161. Snowdon DA (1985) Diet and ovarian cancer (Letter). JAMA 254: 356–357
162. Stone M, Dent J, Karana A, et al (1976) Relationship of oral contraception to development of trophoblastic tumour after evacuation of a hydatidiform mole. Br J Obstet Gynaecol 83: 913–916
163. Stuart GC, Allen HH, Anderson RJ (1981) Squamous cell carcinoma of the vagina following hysterectomy. Am J Obstet Gynecol 139: 311–315
164. Sturgeon SR, Brinton LA, Devesa SS (1992): In situ and invasive vulvar cancer incidence trends (1973 to 1987). Am J Obstet Gynecol 166: 1482–1485
165. Teoh ES, Dawood My, Ratnam SS (1972) Observations on choriocarcinoma in Singapore. Obstet Gynecol 40: 519–524
166. Terris M, Wilson F, Smith H, et al (1967) The relationship of coitus to carcinoma of the cervix. Am J Public Health 57: 840–847
167. Toki T, Kurman RJ, Park JP, Kessis T, Daniel RW, Shah K (1991) Probable nonpapillomavirus etiology of squamous cell carcinoma of the vulva in older women: A clinicopathologic study using in situ hybridization and polymerase chain reaction. Int J Gynecol Pathol 10: 107–125
168. Trichopoulos D, Papapostolou M, Polychronopoulou A (1981): Coffee and ovarian cancer. Int J Cancer 28: 691–693
169. Tyler CW Jr, Webster LA, Ory HW, et al (1985): Endometrial cancer: How does cigarette smoking influence the risk of women under age 55 years having this tumor? Am J Obstet Gynecol 151: 899–905
170. Verreault R, Chu J, Mandelson M, et al (1989): A case-control study of diet and invasive cervical cancer. Int J Cancer 43: 1050–1054
171. Vogelstein B, Fearon ER, Hamilton SR, et al (1988): Genetic alterations during colorectal-tumor development. N Engl J Med 319: 525–532

172. Vonka V, Kanka J, Hirsch I, et al (1984) Prospective study on the relationship between cervical neoplasia and herpes simplex type–2 virus. II. Herpes simplex type–2 antibody presence in sera taken at enrollment. Int J Cancer 33: 61–66
173. Walker AM, Jick H (1980): Declining rates of endometrial cancer. Obstet Gynecol 56: 733–736
174. Wassertheil-Smoller S, Romney SL, Wylie-Rosett J, et al (1981) Dietary vitamin C and uterine cervical dysplasia. Am J Epidemiol 114: 714–724
175. Webster LA, Weiss NS, the Cancer and Steroid Hormone Study Group (1989) Alcoholic beverage consumption and the risk of endometrial cancer. Int J Epidemiol 18: 786–791
176. Weiss NS, Lyon JL, Krishnamurthy S, et al (1982) Noncontraceptive estrogen use and the occurrence of ovarian cancer. JNCI 68: 95–98
177. Weiss NS, Lyon JL, Liff JM, et al (1981) Incidence of ovarian cancer in relation to the use of oral contraceptives. Int J Cancer 28: 669–671
178. Weiss NS, Sayvetz TA (1980) Incidence of endometrial cancer in relation to the use of oral contraceptives. N Engl J Med 302: 551–554
179. Weiss NS, Szekely DR, English DR, et al (1979): Endometrial cancer in relation to patterns of menopausal estrogen use. JAMA 242: 261–264
180. Whittemore AS, Harris R, Itnyre J, et al (1992) Characteristics relating to ovarian cancer risk: Collaborative analysis of 12 US case-control studies. Am J Epidemiol 136: 1184–1203
181. Whittemore AS, Wu ML, Paffenbarger RS, et al (1989) Epithelial ovarian cancer and the ability to conceive. Cancer Res 49: 4047–4052
182. Whittemore AS, Wu ML, Paffenbarger RS, et al (1988): Personal and environmental characteristics related to epithelial ovarian cancer. II. Exposures to talcum powder, tobacco, alcohol and coffee. Am J Epidemiol 128: 1228–1240
183. WHO Collaborative Study of Neoplasia and Steroid Contraceptives (1989) Epithelial ovarian cancer and combined oral contraceptives. Int J Epidemiol 18: 538–545
184. Winkelstein W Jr (1990) Smoking and cervical cancer—current status: A review. Am J Epidemiol 131: 945–957
185. Wright NH, Vessey MP, Kenward B, et al (1978) Neoplasia and dysplasia of the cervix uteri and contraception: A possible protective effect of the diaphragm. Br J Cancer 38: 273–279
186. Wu ML, Whittemore AS, Paffenbarger RS Jr, et al (1988) Personal and environmental characterics related to epithelial ovarian cancer. I. Reproductive and menstrual events and oral contraceptive use. Am J Epidemiol 128: 1216–1227
187. Wynder EL, Escher GC, Mantel N (1966) An epidemiologic investigation of cancer of the endometrium. Cancer 19: 489–520
188. Yuen BH, Burch P (1983) Relationship of oral contraceptives and the intrauterine contraceptive devices to the regression of concentrations of the beta subunit of human chorionic gonadotropin and invasive complications after molar pregnancy. Am J Obstet Gynecol 145: 214–217
189. Ziegler RG, Brinton, Hamman RF, et al (1990) Diet and the risk of invasive cervical cancer among white women in the United States. Am J Epidemiol 132: 432–435
190. Zunzunegui MV, King MC, Coria CF, et al (1986) Male influences on cervical cancer risk. Am J Epidemiol 123: 302–307
191. zur Hausen H (1982) Human genital cancer: Synergism between two virus infections or synergism between a virus infection and initiating event? Lancet 2: 1370–1372