Comparative Medicine
Copyright 2001
by the American Association for Laboratory Animal Science

Vol 51, No 3
June 2001
Pages 208-217

Overview

A Primer in Epidemiologic Methodology

Benjamin J. Weigler, DVM, PhD

Epidemiology is defined as the study of the distribution and determinants of disease within populations. In addition to the requirements for disease surveillance, epidemiologic methods have numerous applications in laboratory animal science and can reveal important insights into the multifactoral mechanisms of disease, thereby aiding in the design of optimized intervention strategies. Observational approaches to data collection can be used to quantify the role of causal factors under natural circumstances, complementing the value of experimental studies in this field. The meaning and appropriate use of standard measures of disease frequency and exposure-disease relationships are reviewed, along with explanations of bias and confounding. Recommendations for reporting the methods and findings from this type of work in comparative medicine literature are presented. Aspects of model-based approaches to data analysis are introduced, offering further opportunities for gaining needed information from epidemiologic study of problems in laboratory animal medicine and management.

The practice of laboratory animal medicine necessarily involves many sub-disciplines, including clinical medicine and management, pathology, surgery, physiology, nutrition, and epidemiology. The benefits of epidemiology for laboratory animal science come from their ability to provide unbiased estimates of disease frequency and in identifying and quantifying the relative importance of risk factors, thereby guiding programs of disease surveillance, outbreak investigation, and prevention and elimination strategies. Insights into the biology and transmission pathway of communicable agents are also gained through these methods. Risk assessment and risk management present other potential uses of epidemiology in laboratory animal science, particularly as they relate to the initial design and ongoing refinement of occupational health and safety programs in the research animal workplace (1).

Epidemiologists obtain data through population-based observations of natural events, often to shed light on possible risk factors that form the causal pathway to disease (2-6). All diseases are considered multifactoral, where various factors can be viewed to function as necessary, and/or sufficient components of this pathway. Rarely in causal relationships is there a direct one-to-one (necessary and sufficient) correlation between these factors and the disease. More often, disease is the result of a multistage process, wherein a series of factors act in concert (e.g., infective agent, immune status, host genotype, and host behaviors) to determine which individuals will be affected. Quantifying the elements of that biologically meaningful framework provides the strongest foundation on which intervention strategies are developed and aids in the refinement of experimental-mechanistic investigations. This is the power of analytical epidemiology in the medical, behavioral, and veterinary sciences. Experimental work, including randomized controlled trials, often help add validity to associations drawn by epidemiologic study, but both types are susceptible to problems of bias and confounding.

Animal Health Resources, Fred Hutchinson Cancer Research Center, Mailstop: BE-114, 1100 Fairview Avenue North, Seattle, Washington 98109.

Epidemiologic approaches have been useful in the study of laboratory animal infectious diseases (including zoonotic pathogens), cancer, reproductive problems, metabolic disease, trauma, and occupational health concerns, among others. A literature search of articles published in *Comparative Medicine* (formerly *Laboratory Animal Science)*, using the key word "epidemiology" indexed 267 articles published in this one journal through December 31, 2000, documenting its importance to the field. The article published here provides an introductory overview of the principles and methods used in epidemiologic research in hopes of stimulating its many possible applications within the laboratory animal science community. Readers are encouraged to peruse the referenced sources of information in this vast discipline for elaboration on specific topics, quantitative methodology, mathematical proofs and theories, and computer software-based approaches to data analysis.

Fundamental Issues

Epidemiology is defined as the study of the distribution (person, place, and time) and determinants (risk factors or exposure variables) of disease in populations, with the aim being to reduce morbidity and mortality through development of intervention strategies (7). Traditionally, this term has been limited to applications in human populations, with epizootiology reserved for animal species. For simplicity and due to biologically common principals, epidemiology is now preferred in either instance (8). The study of disease distribution typically involves calculation of rates and proportions expressing disease frequency by standard measures. This simplifies interpretation, allows comparisons between studies, and provides clearly defined mathematical parameters for statistical tests of possible differences or trends. Stating the case definition for the disease in clear and consistent terms is the most important initial step, even if it needs to be revised in future work. The case definition in epidemiologic studies may differ from its clinical use, and should include as many or few criteria as necessary for its functionality while maximizing sensitivity and specificity. Disease can

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-28

be any discrete and definable outcome, typically involving a condition or state of bodily dysfunction, and is sometimes identified by a battery of tests. Exposure variables may include etiologic agents, genetic factors, chemical agents, diets, behaviors, workplace functions, age, breed (or race, stock, or strain), micro- and macroenvironments, and all other factors that could represent components in the causal pathway to disease.

Epidemiologists traditionally convey this relationship as a triad (Fig. 1), where host characteristics have a key role and the possibility of arthropod vectors (as in ticks or mosquitoes) may be involved in some conditions. The use of epidemiologic methods allows all components of this triad to be quantitatively assessed for their role as risk factors (determinants) of disease development. Estimates of the risk, or probability that the event will occur in association with these disease determinants is done using methods described in part through this overview. Common measures of disease occurrence are its incidence, defined as the number of new cases during a specified period within a population at risk, and its prevalence, defined as the total number of cases (pre-existing and new) within the population during a specific point or span of time (such as a month or year).

Incidence is calculated as a rate measure, whereas prevalence is a proportion that varies according to disease duration. In steady-state conditions where these measures and the study population size are unchanging, prevalence is equivalent to the product of incidence and disease duration. Many examples of prevalence surveys exist in laboratory animal literature, with most applications having been to document that a specific disease or condition exists in a facility, either to confirm suspicions or to establish a baseline for future actions. However, incidence-based measures are required if the objective is to study causal pathways since they allow the exposure-disease relationship to be quantified with greater conviction. Incidence-based studies of communicable agents typically first require assessing which individuals are to be considered at risk through serologic or other tests, often allowing both measures to be determined within the same investigation.

Epidemics are situations in which a rapid increase in disease frequency has occurred in a population beyond some arbitrary background level, or for the first time in an area previously thought to be disease free. The term "outbreak" is typically used to describe local or regional events over a limited time span. Investigation of disease outbreaks follows standard epidemiologic approaches where the period of interest is often restricted to a narrower interval, particularly for acute diseases. In this instance, the principal goal is to prevent further spread by identifying the source and other components of the causal pathway. Obtaining sufficient data to confirm that the observed disease frequency rates are consistent with an outbreak is an important initial step in this process.

The study of natural events or conditions through observational epidemiologic methods avoids the potential pitfall of misrepresenting important aspects of real-world conditions (e.g., dose, route, intercurrent factors, and antigenic variability) that can occur during experimental studies. However, the study of large populations can be costly in time and resources, especially for rare conditions. Both approaches require a well-reasoned approach to study design and analysis, and neither should be initiated before the biology of the disease and the availability and limitations of the data are understood through literature-based ground work. Like all good science, the best epidemiologic work is predicated on a well-formulated, refutable hypothesis and deductive reasoning. Ranking the issues and outcomes considered to be most important will aid in the design of the study. Disproving the null hypothesis through appropriately chosen analytical methods is as important in epidemiologic work as in other medical sciences that use statistics as the basis for establishing proof (2, 5). Results of these investigations can be used to reveal aspects of the causal mechanisms, which often lead to further epidemiologic and experimental studies to narrow the focus areas. Each study will have possible constraints that drive the methodology and add caveats to interpretation of the conclusions. This is not problematic when recognizing that strength in conclusions drawn from epidemiologic work comes from similarly repeated findings in other contexts. Many studies of the same disease condition lead to inconclusive results because of limitations in the study design, population differences, and weaknesses in the magnitude of the exposure-disease relationship (6).

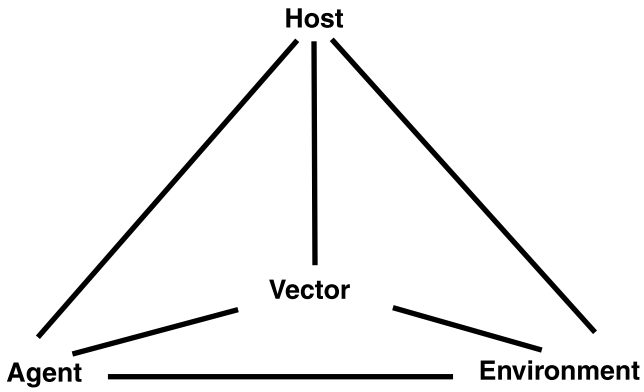

Figure 1. The epidemiologic triad.

Data Sources and Sampling Considerations

In all epidemiologic studies, it is critical to assess the types of data that can be made available, both in terms of the population to which the results are to be extrapolated and with regard to the measures of disease and exposure. Compiled findings from reference diagnostic laboratories provide a valuable overview of the etiologic agents circulating in various institutions but without information regarding the size and nature of the source population, only limited conclusions can be drawn regarding differences, causative factors, or trends. Certainly, key information required of veterinary studies in the species and breed (or stock/strain), type of housing for primary and secondary enclosures, age, sex, diet, management, and features of the program of veterinary care that communicate the likelihood that cases of disease would be recognized (8). For communicable disease investigations, the premise that all animals in the population have a reasonable probability of exposure cannot be assumed, especially given modern arrangements of caging types and housing locations. Unlike human communities or free-ranging species, biomedical research facilities and agricultural production methods force any potential for random mixing to the herd, room, pen, or cage level. This has clear implications to the appropriate methods of data analysis, as will be discussed. An initial step in the design of epidemiologic studies comes in defining the population

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-28

to which we wish to generalize (i.e., the target population), then deciding on the best methods to sample that population for the set of risk factors and outcomes of interest. Specifying the basis for the approach to sampling makes clear essential aspects regarding the likelihood that the findings could be meaningful beyond simply the set of animals that contributed to the data (i.e., its external validity [2, 6]).

Simple random sampling, stratified sampling (e.g., by age, breed or strain, sex, location, or housing type), systematic sampling, and cluster sampling techniques represent some options for consideration (9). In all instances, each element (animal) from the source population must have had a definable probability for inclusion. This allows unbiased comparisons of the resulting disease rates and proportions via statistical methods. Some epidemiologic methods (e.g., case-control studies and cohort studies) do not have assessments of disease frequency as an objective, but focus instead on quantifying the role of possible disease determinants to aid in prevention and intervention strategies. The possibility for bias due to errors in the design or conduct can plague interpretation of the results in either event. For estimation of means and proportions, determination of the appropriate sample size requires specification of the source population size, the desired bounds on the error of estimation, and an estimate of the variance (or population proportion). Sample size requirements for studies in which a specified statistical power in measuring the importance of risk factors on disease is the objective are reviewed elsewhere (10- 12).

Principles of disease surveillance in the context of various laboratory animal species have been published in comparative medicine literature (13-16). Differences in systems of animal housing, assemblages of species, stocks, strains, and other aspects of laboratory animal medicine and management greatly impact the underlying assumptions of the process and should influence the choice of a sampling strategy. Wherever possible, the approach should be well-defended through detailed description of the prior information and statistical methods used for the sampling algorithm.

Bias, Confounding, and Interaction

Bias is defined as a non-random (systematic) error in a study that leads to a result distorted from the truth. Rothman and Greenland (2) differentiate two principal types of bias that can reduce the ability of epidemiologic studies to draw meaningful conclusions: selection bias and information bias. Information bias results from problems in ascertainment of the individuals selected for study in regards to the exposures and/or outcomes of interest. Selection bias reflects problems in generating the sample of individuals included for study in a matter that distorts the outcome and the exposure-disease relationship. Information bias can occur when animals are misclassified with respect to the diseases of interest on the basis of diagnostic test results that are limited in sensitivity and specificity. It can also occur with respect to exposure variables when methods used to ascertain risk factors among members of the sample are inappropriate, as in the case of significant management or research related differences within the population. Both types of bias can be differential (the values for one group are erroneously weighted more than another) or non-differential (the errors are spread randomly among groups) with respect to either the exposures or the disease. Differential misclassification can bias the exposure-disease relationship against the truth in either direction (i.e., either toward or away from a significant association). Non-differential misclassification of the exposure or the disease tends to bias the relative risk measures toward insignificance, even in the presence of a true relationship (2-6).

Established methods exist to design epidemiologic studies in a manner that can minimize the potential for large bias from these sources and sometimes to partly correct for them during data analysis when they are suspected or known to exist (2, 5, 10). Testing the possible extent of the suspected bias against the association measure is important to determine the likelihood that the conclusions could have been influenced by those effects.

Confounding in epidemiologic studies can also bias the results by introducing misleading evidence for the role of one or more risk factors in predicting the outcome. This is known as a "mixing effect," where some extraneous variable(s) are associated with an important exposure (but not causally following exposure) and are associated with disease in the absence of such exposure. Since the goal of much epidemiologic work is to elucidate causal pathways, thoughtful attention to reducing or eliminating confounding and other types of bias through design and analysis of the study is paramount. One example of confounding comes from a study of risk factors for B-virus infection in three corral-housed groups of rhesus monkeys (17). In this instance, the housing group confounded the monkey age-antibody relationship and failure to include it in the analysis would have introduced bias to estimates of that exposure-disease measure.

The criteria used by epidemiologists when critically evaluating causation were presented by Hill in 1965 (18) and differ from the traditional Henle-Koch postulates in part since they allow for conditions where infective agents play no or only partial roles, and also where lack of animal models makes difficult establishment of experiment-based proofs. The criteria described by Hill are: strength of the exposure-disease association; consistency of the association in different populations under different circumstances; specificity of the association for the outcome; temporal relationship of the exposure preceding the disease; dose-response relationship of increasing risk with increasing levels of exposure; biological plausibility of the exposure-disease relationship in light of the current body of knowledge; coherence of the relationship without conflict, given known aspects of the disease's natural history; experimental evidence for the association; and analogy with other relationships of similar exposure types and outcomes. Some of these criteria were admittedly weak, and Hill noted that establishing causality requires sound scientific reasoning without a simple check-list approach. More recent monographs have indicated a ruling-out of alternate plausible hypotheses and a documented decrease in risk following cessation of the exposure (6) in lieu of the last two criteria, which are particularly tenuous.

Data on factors that are known or suspected to confound the exposure-disease relationship of interest must be collected during animal studies or no effective arguments against the potential for spurious findings can be made. Evaluation of the results may shed light on additional confounders which should be incorporated into future investigations. Matching on confounding variables (individually or at the group-level) accentuates their importance in the study design phase, as does a stratified analysis or forcing confounders into the model for statistical regression methods used in the data analysis phase (5, 10, 11, 19).

Also known as effect modification, epidemiologic interaction

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-28

occurs when differences in the magnitude of the exposure-disease association exist according to the level of another factor. While the interaction between one or more variables in modifying the risk-relationship of interest obviates the possibility of corrections for confounding by that variable, it raises important evidence for biologically meaningful positive or negative synergism in the data. This can have profound implications to some groups, perhaps due to other exposures, environmental, or genetic factors. Interaction terms are always included in epidemiologic models when they are detected. Epidemiologic models that build upon these features have incorporated assumptions about additivity or multiplicity between cofactors (2, 19). Unlike confounding, statistical tests exist to evaluate the hypothesis of significant interaction among the variables at any specified α-level.

Measuring Disease and Exposure

Differentiating the population of animals with and without the disease of interest relies on the use of one or more diagnostic tests. Every diagnostic test has inherent technical features that determine its sensitivity and specificity for the condition, which together provide its validity. This is true for clinical methods and devices (e.g., abdominal palpation or stethoscope auscultation) as well as laboratory assays (e.g., an enzyme-linked immunosorbent assay [ELISA] or a polymerase chain reaction [PCR] test), and these features do not vary with the frequency of disease in the population. Sensitivity is a test's ability to correctly identify those with the disease, whereas specificity is a test's ability to correctly identify those without it. Absolute certainty in differentiating these populations is typically not possible due to the overlapping continuum of characteristics used to define the disease condition (e.g., the frequency of loose stools as used to define chronic colitis in nonhuman primates) and lack of perfectly valid diagnostic tests that are still feasible to implement. Estimates of the efficiency of a diagnostic test, representing the total proportion of correctly classified results (i.e., true positives plus true negatives over the total number tested) are reported by some authors, but are less informative than are the separate values.

Known biological features of the disease and evaluation through additional high-validity diagnostic tests (i.e., "gold-standards") are used as the measure against which the sensitivity and specificity of an assay can be determined on a population-wide basis. Only relative sensitivity and specificity, in relation to another test's findings, can be determined in the absence of certainty regarding true disease status. Realistic assessment of test performance can only be gained by evaluations in populations of diseased and non-diseased animals that reflect the actual characteristics (e.g., age, breed or strain, sex, breeding and other attributes of use) and disease spectrum for the condition of interest when it is used in other contexts. It is important that the non-diseased groups without the condition include other salient characteristics that are likely to be confused with the disease of interest (e.g., potentially cross-reacting microbes or antibodies) to gain the best measure of specificity under field conditions (8, 20). Recent work has documented use of maximum likelihood methods and Bayesian inference theory to provide estimates of test sensitivity and specificity even under conditions where no perfect gold standard exists (21).

Using one or more diagnostic tests in series increases specificity and the expense of sensitivity, and may be appropriate for conditions in which the penalty of false-positive results is high, as in the case of declaring an individual to be infected with human immunodeficiency virus. For this approach, all results found to be positive in one assay (e.g., ELISA) are submitted to a second assay (e.g., Western blot immunoprecipitation), and only those specimens yielding positive results of both tests are considered to be from infected hosts. Using tests in parallel offers another approach wherein the greatest penalty would come from missing a true case, and specificity is sacrificed at the expense of sensitivity. Specimens are then submitted to two or more assays simultaneously, and those returning positive results on any test are considered infected. It is important to recognize that the performance features of test sensitivity and specificity are generally knowable for many (but not all) infective agents, in that experiments or observational studies could be designed to quantify those parameters with relative confidence. If the basis for making such assessments is reasonable, estimates of disease incidence or prevalence in a population should be numerically adjusted to account for this source of error (20, 22). Failing to do so can bias programs of disease surveillance and waste valuable time and resources. The statistical uncertainty around all of these measures, as indicated for example through 95% confidence intervals, should also be reported so that objective judgments can be made regarding the cause for concern.

Increasing use of the tools from molecular biology in epidemiologic research has allowed increased precision in resolving important stages and events along the pathway to disease, including markers of exposure, biologically relevant doses, altered cellular functions, and variants of outcomes within the disease definition (23). This has the potential to significantly reduce the extent of misclassification bias and thereby improve on the resulting risk estimates. Molecular epidemiology can include use of biomarkers that can represent an exposure event or its correlate, or be a predictor for the extent of disease among those exposed. Many studies of infectious disease, including those in laboratory animal science (24-27), have involved use of molecular diagnostic techniques to differentiate among genotypes of etiologic agents in epidemiologic investigations.

For clinical decision making (including test-and-slaughter programs in livestock and in laboratory animals), the positive predictive value and negative predictive value of a diagnostic assay are critical matters. Given the use of a diagnostic test in a defined population, the positive predictive value is the probability that each of the individuals with positive test results actually has the disease, whereas the negative predictive value is the probability that they don't (3, 8, 20). The former is most greatly influenced by test specificity, whereas the latter is dictated more by test sensitivity. Prevalence of the disease in the population markedly affects both measures at fixed sensitivity and specificity, especially when comparing scenarios where the condition is nearly absent to the situation where it is more common (say, > 5% prevalence).

Re-testing of samples from individuals with negative or indeterminate results on an initial test battery can be done using the same assays over time to help document absence of a condition, particularly to increase test sensitivity and thereby improve quality assurance at the aggregate (e.g., herd) level for test-and-removal programs in disease eradication (8). The use of two or more assays in this regard can, however, fail to produce significant gains in identification of infected animals under circumstances where there is conditional dependence between diagnostic tests (28).

In laboratory animal science, limited data regarding the sensitivity and specificity performance attributes of available diag-

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-28

nostic assays for infective agents, especially with regard to different species and different states of immune competency, complicate some investigations where the expectation is for unbiased estimates of disease frequency. This is particularly true when test systems developed for use in one species are extrapolated to another without controlled validation for factors likely to produce false-positive (as in cross-reacting agents) or false-negative (as in species-specific components of the assay) reactions. Another concern is the expectation for having uniformly comprehensive programs of disease surveillance (quality assurance) for many infective agents while using the same set of sampled individuals without consideration of possible differences in prevalence, diagnostic test limitations, housing configurations, immune response capabilities, and basic biology for the organisms potentially circulating in the target population.

Since the sample size, test methods, and optimal sampling strategy often vary from one situation to another, especially in light of differences among institutions in their disease experience and listed priorities, failure to justify the chosen strategy through deductive reasoning can result in pitfalls. A program that maximizes the sensitivity of disease detection for one agent may not be appropriate for another, thereby raising potential challenges for cost-constrained institutions. Sentinel rodents are commonly housed in rooms with research animals in a manner that increases their likelihood of exposure to agents transmitted via the fecal-oral route (e.g., on dirty bedding), but some agents are not readily detectable that way (29, 30). Justifying the minimum sample size for these programs solely on the commonly cited formula that derives from the binomial distribution (14) may not be sufficient, since the assumptions of large population size, equal likelihood of exposure among individual (random mixing), random sampling, and minimum detectable prevalence are not usually met, in addition to the multiple-agent concern where again the conditional dependence between performance attributes of each test type should be considered (28). One good example of a model-based approach to the design of a disease surveillance program in laboratory rodents on the basis of institution-specific experience documents some of the opportunities for improvement (31). In this instance, a two-parameter exponential distribution was found to best represent the mathematical process behind 17 episodes of viral disease contamination in 81 barrier rooms for rats and mice over three years. Maximum likelihood methods were used to estimate the model parameters and statistically based confidence intervals were constructed to suggest appropriate sample sizes and test intervals for rodent colonies in their facilities with high assurance.

Common Epidemiologic Study Designs

The five designs most frequently encountered in epidemiologic investigations are: case studies (or case series); cross-sectional studies; cohort studies; case-control studies; and ecologic studies. Case studies involve descriptive findings from case-patients without reference to a comparison or control group, and therefore have no ability to discern causal relationships with respect to different exposures or clinical treatments. They are useful in providing descriptive summaries regarding characteristics of diseased individuals as a prelude to establishing a robust case definition, typically in advance of pursuing one of the other study types or for disease surveillance purposes. Cross-sectional studies (also known as prevalence studies) are done by obtaining a snapshot of the source population through sampling at a particular point in time, then cross-classifying each individual in the sample by their exposure and disease status. Although the temporal relationship between exposure and disease onset cannot be made using cross-sectional designs, the sample inherently contains a control group and the risk relationship can, therefore, still be quantified and used to generate research questions. Risk factors identified in cross-sectional studies are typically pursued in other observational and experimental studies.

In cohort studies, the source population is sampled on the basis of exposure for the factor(s) of interest and the disease incidence rates are compared prospectively among exposed and unexposed groups to quantify the risks of exposure on disease development. Retrospective cohort studies are essentially the same, except that information on exposure is drawn from historical data pertinent to the same source population. Prospective cohort studies can be designed to minimize the likelihood of bias and allow the proper temporal assessment of exposure prior to disease development, but they can be quite costly and limited in statistical power, especially for rare conditions (5).

There are many variations of case-control studies, depending on the type of study subjects, the disease, exposure information, and availability of appropriate reference groups. The general principle is that sampling of the source population is based on disease status (i.e., those with and without disease), and information regarding the exposure histories among cases and controls is then ascertained. The exposure-disease relationship is then quantified by cross-classifying the data based on each to provide estimates of risk. The well-reasoned choice of an appropriate control group, clarity in the criteria of study inclusion for case-patients, collection of sufficient exposure data (including likely confounders), and the discussion of possible caveats to interpretation can minimize some of the controversies surrounding this design option (2, 10, 11). Case-control studies can either be retrospective or prospective, and matched samples (individual or group) can be incorporated into the design for key variables likely to confound the exposure-disease relationship of interest. Notice, however, that the variables used for matching cannot then be examined as risk factors, since they have been forced into the sampling method. The efficiency of case-control studies for studying risk factors of rare diseases has made them popular and important tools in epidemiologic work, especially when done with matched samples. Hybrid designs, for example where case-control studies are nested within cohort studies, also are possible (5).

Ecologic studies represent another type of epidemiologic study design, in which entire populations or groups comprise the observational unit, and only averaged exposure information is available for the entire group instead of for each individual. In this event, measures of disease incidence or mortality are frequently compared among groups differing in their exposure status. The exposure-disease relationship from ecologic studies can be suggestive of an important underlying biological process, but they are prone to confounding and the "ecological fallacy" that results from inappropriately extending the values of group-wise associations to individuals (2, 4, 6).

Estimating the Magnitude of the Exposure-Disease Relationship

Quantitative assessment of the determinants of disease in a population is done through various methods in epidemiology,

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-28

contingent upon certain statistical properties and opportunities dictated by the study design. The goal is to estimate the risk of disease associated with one or more exposures, where risk is defined as the probability that the event will occur in an individual within a specified time period (7). As described in the context of measurements for disease frequency, the time-basis is fundamental to this concept. The cumulative incidence of a disease is a direct statement of average risk from time t_1 to t_2, contingent upon the individuals not dying from any cause during that period. Analogous to all probabilities in statistics, it is bounded by zero and one and has no terms of dimension. In contrast, incidence rates reflect the frequency of disease occurrence as a function of population-time of observation, are bounded by zero and infinity, and have time (e.g., "per year") as the dimensional units. Either of these measures can be calculated as two separate components, stratifying the rates or risks for those with and without the exposure of interest. A ratio expression of the exposed group over the non-exposed group for the cumulative incidence (risk ratio) or incidence rate (rate ratio) provides another statement of risk (relative risk) as the strength of association between the exposure and the disease under investigation (Fig. 2). Relative risks equal to one imply no evidence for association between the exposure and disease. Relative risks greater than one support the hypothesis of an association between exposure and disease, in that individuals with the exposure are more likely to acquire the disease. Any relative risks that are less than one again imply an association, but as a protective relationship against disease development. Methods to calculate *P*-values and confidence intervals against the study-wise α-level for these ratio measures exist and should be used along with the point estimates to interpret their significance (2, 5).

The odds ratio offers another measure of exposure-disease association, and is the only option available for case-control studies and cross-sectional studies, since neither affords the opportunity to measure incidence. The odds ratio is determined by dividing the odds of exposure among cases by the odds of exposure among controls (Fig. 2). This provides a mathematical estimate of the incidence rate ratio (and therefore the relative risk) in case-control studies, especially under conditions where the disease is rare (say, < 10% prevalence). Some case-control sampling methods allow the odds ratio to provide such estimates even for more common conditions (11). Interpretation of odds ratio values as evidence for an increased or protective effect on disease development given exposure, relative to the baseline reference group, is the same as that for relative risks. Odds ratios are also used when analyzing the results of cohort studies where the population sampling is based on exposure status, although many reviewers would prefer the reporting of incidence-based ratios whenever possible. For cross-sectional studies, the duration of the risk period and the relationship between exposure and survival affect interpretation and potential bias from the use of odds ratios as risk estimators. There are statistical properties associated with odds ratios that promote their use in multivariable models, such as logistic regression, where multiple exposure factors and confounding variables can be included and thereby allow assessment of each variable while controlling for the others.

While the relative risk and odds ratio quantify the magnitude of the likelihood that a particular exposure is associated with the disease of interest, the attributable fraction among the exposed is the proportion by which their incidence rate would decrease if such variable were eliminated or prevented. Likewise, the attributable fraction for the population (etiologic fraction) estimates the extent to which its source population-wide incidence rate would be reduced if the risk factor were to be eliminated, and can be used to document the benefits of entire disease prevention programs. Large relative risks do not necessarily translate into large etiologic fractions if exposure to the variable in the population is rare. Both of these values can be expressed either as a proportion or percentage. A related measure is the risk difference, reflecting the balance of the risk (or cumulative incidence) of disease in the exposed group after that for the unexposed group has been subtracted (2, 4, 5). The rate difference, obtained by substituting incidence rate values for the aforementioned, offers a conceptually similar approach to determining the "excess" disease associated with the exposure, and all have been used as arguments in support of public policy and disease control actions.

Statistical Approaches to Data Analysis

So far, this theoretical framework has been presented in the context of the simplest case where a dichotomous exposure variable (i.e., exposed versus unexposed) is ascertained for a dichotomous disease outcome (i.e., diseased versus non-diseased). In most situations, exposures are not dichotomous but continuous, so categories must be established to represent the levels (doses) of interest, given the type of data available and existing scientific knowledge about the pathway to disease. For diseases with long latent periods it is important to ascertain exposure at a time more likely to have had influence, if the relationship is causal. One of Hill's criteria (18) for establishing causality includes assessments for a biological gradient (dose-response curve) in the exposure-disease relationship, although such apparent trends are also subject to confounding (2). The increasing availability and use of biomarkers such as DNA adducts (23) in epidemiology can circumvent some of these problems and allow population-based investigations at specific stages of a disease process within the causal pathway of exposed individuals. This does not eliminate the potential for confounding but can sometimes increase the precision of the measures. In all instances, it is essential to carefully define the exposure variables to allow for critical interpretation of the results and provide the possibility of repeating the study in other populations.

Likewise, the case definition of disease must be explicit and allow for unambiguous differentiation of affected and unaffected individuals during the specified period of observation. Typically, disease outcomes are dichotomous but procedures also exist for analysis of epidemiologic data through polytomous approaches, whereby multiple levels of outcome can be specified in relation to a set of exposure factors (19). In all instances, it is important to consider not only the point estimates of the association measure and any corresponding *P*-values, but also their variability as provided through interval estimation (e.g., 95% confidence intervals). The width of those intervals conveys valuable information about the variability among the point estimates, as well as their magnitude and likelihood of significance relative to the established α-level (2). They should always be included when reporting epidemiologic findings.

Methods of contingency table analysis, typically involving variations of χ^2-tests, are used to evaluate the statistical significance of point estimates for the odds ratio and relative risk (32). This includes assessments for the overall degree of association and for

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-28

	Disease (Cases)	No Disease (Controls)	Row Totals
Exposed to Risk Factor	a	b	a+b
Not Exposed to Risk Factor	c	d	c+d

For Cohort Studies and Cross-Sectional Studies:

$$\text{Odds Ratio} = \frac{\text{Odds of becoming a case among exposed}}{\text{Odds of becoming a case among non-exposed}} = \frac{a/b}{c/d} = \frac{ad}{bc}$$

$$\text{Relative Risk} = \frac{\text{Incidence rate (risk) among the exposed}}{\text{Incidence rate (risk) among the non-exposed}} = \frac{a/(a+b)}{c/(c+d)}$$

For Unmatched Case-Control Studies:

$$\text{Odds Ratio} = \frac{\text{Odds of exposure among cases}}{\text{Odds of exposure among controls}} = \frac{a/c}{b/d} = \frac{ad}{bc}$$

Figure 2. Calculation of the exposure-disease relationship.

each individual strata representing different categories of the exposure variable upon disease development. Most epidemiologic study designs involve this approach in the data analysis phase to explore the univariate-significance of each putative risk factor as a prelude to development of more complex statistical models. These univariate summary measures and statistical tests can be calculated by hand or generated via most computer software packages. Candidate variables for inclusion in the chosen model are based on those initial findings, in conjunction with known biological aspects of the disease and any associated confounding variables (33).

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-28

Model-based approaches have advantages over stratified analysis, in that they avoid imprecision from having too few observations per stratum and allow simultaneous evaluation of multiple covariates, thereby reducing one potential for confounding. However, the statistical assumptions used to predicate the chosen model should be validated and well reasoned, and clear understanding of the software package output is essential to prevent erroneous conclusions.

Logistic regression models have widespread applications in the statistical analysis of epidemiologic data from cross-sectional, cohort, and case-control studies, allowing quantitative exploration of one or more risk factors on a dichotomous disease outcome (19). This contrasts with linear regression (34), in which case the dependent variable is typically continuous. Statistical assumptions between the two vary slightly, but the strategies for building the model and testing significance of individual predictor variables are analogous. Most statistical software packages accommodate these models, allowing estimation of the model parameters through maximum likelihood iterative procedures. The coefficients that emerge from logistic regression models (logits) are the log odds of the outcome for each covariate (or categories within covariates), which are directly convertible to odds ratios by exponentiation. Simultaneous inclusion of multiple covariate risk factors in a single model reduces the potential for bias and improves precision of the estimates. Exploration of the data for evidence of confounding and interaction are readily achieved through logistic regression, and model-based approaches are, in fact, a good strategy for statistically adjusting the exposure effects of interest for confounding variables. These models can be readily adapted to case-control studies involving pair-matched samples and can include parameters for extra-binomial variation, as in situations where biologically related factors force the outcome variables into natural groupings, such as litters of offspring or herds of animals (35).

Statistical models of survival analysis exist to accommodate epidemiologic investigations that involve time-to-event data, especially where there is loss to follow-up for some individuals (censoring) prior to conclusion of the defined study period (12, 36). The events of interest in this case can be any well-defined outcome, not just death. A variety of statistical distributions exist for evaluation against these data sets, and Cox's proportional hazards model is especially valuable for this since it can be used for testing a set of covariates on the disease incidence (or mortality) rates. In this instance, the values from estimated coefficients are interpreted as relative risks for exposed groups relative to those unexposed to the factors of interest. Parameter estimates are most often produced through maximum likelihood estimation methods, as for logistic regression, and graphic display of the resulting survivorship curves and hazard functions helps to validate inherent statistical assumptions. Cox models also allow for time-dependent covariates, whereby exposures may come and go for intermittent periods among each individual prior to the study endpoint (12, 36). Lifetable approaches to data analysis, such as the product-limit method of Kaplan-Meier plots, are also useful for epidemiologic work, but are generally restricted to evaluation of a single (univariate) risk factor on the outcome of interest.

Since the final selected set of risk factors and their corresponding magnitude of importance that emerge from epidemiologic studies can vary with the choice of model, investigators are advised to evaluate the appropriateness of different models for their particular problem on statistical and biological grounds. The findings should be conveyed in a way that delineates why a model was used as the basis for analysis and how the final set of covariates were decided on for inclusion, as well as the model-building strategy itself. This will help convey some of the potential limitations to the readership and allow possible explanations of any differences in results between studies (33).

Investigators should not base the interpretation of their findings simply on a *P*-value, especially in light of potential problems of low statistical power due to small sample size and the relatively arbitrary nature of establishing the study-wise α-level at any nominal percentile (e.g., 5%). Even factors with documented statistical significance may not be biologically or clinically relevant, and conclusions regarding causal relationships should be made by use of deductive reasoning that extends beyond simple *P*-values. As noted, the inclusion of confidence intervals for point estimates of risk conveys statistical meaning in addition to aspects of the variability that can strengthen or refute a causal hypothesis. Suggested steps that can be used for the critical review of the epidemiologic literature in light of the many issues presented are shown in Table 1.

Table 1. Stepwise approach to critical appraisal of published medical research[a]

Step 1.	Consider the research hypothesis. • Is there a clear statement of the research hypothesis? • Does the study address a question that has clinical relevance?
Step 2.	Consider the study design. • Is the study design appropriate for the hypothesis? • Does the design represent an advance over prior approaches? • Does the study use an experimental or an observational design?
Step 3.	Consider the outcome variable. • Is the outcome being studied relevant to clinical practice? • What criteria are used to define the presence of disease? • Is the determination of the presence or absence of disease accurate?
Step 4.	Consider the predictor variable(s). • How many exposures or risk factors are being studied? • How is the presence or absence of exposure determined? • Is the assessment of exposure likely to be precise and accurate? • Is there an attempt to quantify the amount or duration of exposure? • Are biologic markers of exposure used in the study?
Step 5.	Consider the methods of analysis. • Are the statistical methods employed suitable for the types of variables (nominal versus ordinal versus continuous) in the study? • Have the levels of type I and type II errors been discussed appropriately? • Is the sample size adequate to answer the research question? • Have the assumptions underlying the statistical tests been met? • Has chance been evaluated as a potential explanation of the results?
Step 6.	Consider possible sources of bias (systematic errors). • Is the method of selection of subjects likely to have biased results? • Is the measurement of either the exposure or the disease likely to be biased? • Have the investigators considered whether confounders could account for the observed results? • In what direction would each potential bias influence the results?
Step 7.	Consider the interpretation of results. • How large is the observed effect? • Is there evidence of a dose-response relationship? • Are the findings consistent with laboratory models? • Are the effects biologically plausible? • If the findings are negative, was there sufficient statistical power to detect an effect?
Step 8.	Consider how the results of the study can be used in practice. • Are the findings consistent with other studies of the same questions? • Can the findings be generalized to other populations? • Do the findings warrant a change in current clinical practice?

[a] From (3), page 168. *Reprinted with permission from McGraw-Hill Co., N.Y.*

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-28

Summary

The goal of this report was to provide an overview of methodologic issues and considerations when designing and interpreting epidemiologic studies. Most aspects are to be considered in the realm of general principles for this discipline, with the caveat that veterinary epidemiology and laboratory animal populations in particular can present unique challenges. Perusal of the literature can be used to document instances where inappropriate or novel methods for describing disease rates and proportions have been used, sometimes in conjunction with haphazard sampling without regard to probability distributions that complicate efforts to interpret findings or to provide the basis for valid comparisons between populations. Failure to describe the target population and defend the approach to sampling, along with uncertainty in the attributes of diagnostic tests, has limited the internal and external validity on some occasions. Lack of an appropriate comparison group and potential sampling bias raises suspicion when drawing meaning from diagnostic test findings from laboratory animal populations, even when considering one institution's experience over time. Failure to take advantage of opportunities to analyze epidemiologic data through appropriate statistical methods can result in misleading conclusions and represents potential wastage of valued resources. Use of less powerful statistical models is a related concern, as is too much reliance on *P*-values for interpretation. The comments contained in this report are intended to stimulate interest in epidemiologic study of current problems in laboratory animal science, and to provide a broad outline with references for further advances of this field. The benefits of use of the epidemiologic method along with the population-based nature of medicine and management techniques used for laboratory animal species lead credence to the likelihood that it will continue to be an important discipline in the years ahead.

Acknowledgments

The author thanks Ronald DiGiacomo of the University of Washington and Nicholas Lerche of the University of California, Davis for their useful comments and suggestions. Portions of this work were presented at the American College of Laboratory Animal Medicine 2001 Annual Forum, April 1-4, Point Clear, Ala.

References

1. **National Research Council, Institute of Laboratory Animal Resources.** 1997. Occupational health and safety in the care and use of research animals. National Academy Press, National Academy of Sciences, Washington, D.C.
2. **Rothman, K. J., and S. Greenland.** 1998. Modern epidemiology, 2nd ed. Lippincott-Raven, Philadelphia.
3. **Greenberg, R. S., S. R. Daniels, W. D. Flanders, J. W. Eley, and J. R. Boring.** 1996. Medical epidemiology, 2nd edition. Appleton & Lange, Stamford, Conn.
4. **Kelsey, J. L., A. S. Whittemore, A. S. Evans, and W. D. Thompson.** 1996. Methods in observational epidemiology, 2nd ed. Oxford University Press, N.Y.
5. **Kleinbaum, D. G, L. L. Kupper, and H. Morgenstern.** 1982. Epidemiologic research: principles and quantitative methods. Van Nostrand Reinhold Co., N.Y.
6. **Gordis, L**. 2000. Epidemiology, 2nd ed. W. B. Saunders Co., Philadelphia.
7. **Last, J. M.** 1995. A dictionary of epidemiology, 3rd ed. Oxford University Press, N.Y.
8. **Thrusfield, M.** 1995. Veterinary epidemiology, 2nd ed. Blackwell Science Ltd., Oxford.
9. **Scheaffer, R. L., W. Mendenhall, and L. Ott.** 1986. Elementary survey sampling. Duxbury Press, Boston, Mass.
10. **Breslow, N. E., and N. E. Day.** 1980. Statistical methods in cancer research, vol. 1. The analysis of case-control studies. International Agency for Research on Cancer. World Health Organization, Lyon.
11. **Schlesselman, J. J**. 1982. Case-control studies: design, conduct, analysis. Oxford University Press, N.Y.
12. **Lee, E. T.** 1992. Statistical methods for survival data analysis, 2nd ed. John Wiley & Sons, N.Y.
13. **Weisbroth, S. H., R. Peters, L. K. Riley, and W. Shek.** 1998. Microbiological assessment of laboratory rats and mice. ILAR J. **39:** 272-290.
14. **Hsu, C.-K., A. E. New, and J. G. Mayo.** 1980. Quality assurance of rodent models, p.17-28. *In* A. Spiegel, S. Erichsen, and H. A. Solleveld (ed.), Animal quality and models in biomedical research. Gustav Fischer Verlag, Stuttgart.
15. **Lerche, N. W., J. L. Yee, and M. B. Jennings.** 1994. Establishing specific retrovirus-free breeding colonies of macaques: an approach to primary screening and surveillance. Lab. Anim. Sci. **44:**217-221.
16. **Ward, J. A., J. K. Hilliard, and S. Pearson.** 2000. Herpes B-virus specific-pathogen-free breeding colonies of macaques (*Macaca mulatta*): diagnostic testing before and after elimination of the infection. Comp. Med. **50:**317-322.
17. **Weigler, B. J., J. A. Roberts, D. W. Hird, N. W. Lerche, and J. K. Hilliard.** 1990. A cross-sectional survey for B virus antibody in a colony of group housed rhesus macaques. Lab. Anim. Sci. **40:**257-261.
18. **Hill, A. B.** 1965. The environment and disease: association or causation? Proc. R. Soc. Med. **58:**295-300.
19. **Hosmer, D. W. and S. Lemeshow.** 1989. Applied logistic regression. John Wiley & Sons, N.Y.
20. **Greiner, M., and I. A. Gardner.** 2000. Epidemiologic issues in the validation of veterinary diagnostic tests. Prev. Vet. Med. **45:**3-22.
21. **Enøe, C., M. P. Georgiadis, and W. O. Johnson.** 2000. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. Prev. Vet. Med. **45:**61-81.
22. **Roger, W. J., and B. Gladen.** 1978. Estimating prevalence from the results of a screening test. Am. J. Epidemiol. **107:**71-76.
23. **Schulte, P. A., and F. P. Perera**. 1993. Molecular epidemiology: principles and practices. Academic Press, San Diego.
24. **Whary, M. T., J. H. Cline, A. E. King, K. M. Hewes, D. Chojnacky, A. Salvarrey, and J. G. Fox.** 2000. Monitoring sentinel mice for *Helicobacter hepaticus*, *H. rodentium*, and *H. bilis* infection by use of polymerase chain reaction analysis and serologic testing. Comp. Med. **50:**436-443.
25. **Weigler, B. J., J. E. Thigpen, M. F. Goelz, C. A. Babineau, and D. B. Forsythe.** 1996. Randomly amplified polymorphic DNA polymerase chain reaction assay for molecular epidemiologic investigation of *Pasteurella pneumotropica* in laboratory rodent colonies. Lab. Anim. Sci. **46:**386-392.
26. **Kodjo, A., L. Villard, F. Veillet, F. Escande, E. Borges, F. Maurin, J. Bonnod, and Y. Richard.** 1999. Identification by 16S rDNA fragment amplification and determination of genetic diversity by random amplified polymorphic DNA analysis of *Pasteurella pneumotropica* isolated from laboratory rodents. Lab. Anim. Sci. **49:** 49-53.
27. **Lerche, N. W., R. F. Cotterman, M. D. Dobson, J. L. Yee, A. N. Rosenthal, and W. M. Heneine.** 1997. Screening for simian type-D retrovirus infection in macaques, using nested polymerase chain reaction. Lab. Anim. Sci. **47:**263-268.
28. **Gardner, I. A., H. Stryhn, P. Lind, M. T. Collins.** 2000. Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. Prev. Vet. Med. **45:**107-122.
29. **Cundiff, D. D., L. K. Riley, C. L. Franklin, R. R. Hook, Jr., and C. Besch-Williford.** 1995. Failure of a soiled bedding sentinel system to detect cilia-associated respiratory bacillus infection in rats. Lab Anim Sci. **45:**219-221.
30. **Artwohl, J. E., L. M. Cera, M. F. Wright, L. V. Medina, and L. J. Kim.** 1994. The efficacy of a dirty bedding sentinel system for detecting Sendai virus infection in mice: a comparison of clinical signs and seroconversion. Lab. Anim. Sci. **44:**73-75.

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-28

31. **Selwyn, M. R., and W. R. Shek.** 1994. Sample sizes and frequency of testing for health monitoring in barrier rooms and isolators. Contemp. Top. Lab. Anim. Sci. **33(3):**55-60.
32. **Fleiss, J. L.** 1981. Statistical methods for rates and proportions, 2nd ed. John Wiley & Sons, N.Y.
33. **Greenberg, R. S., and D. G. Kleinbaum.** 1985. Mathematical modeling strategies for the analysis of epidemiologic research. Ann. Rev. Public Health **6:**223-245.
34. **Kleinbaum, D. G., L. L. Kupper, K. E. Muller, and A. Nizam.** 1998. Applied regression analysis and other multivariable methods. Duxbury Press, Pacific Grove, Calif.
35. **Curtis, C. R., R. H. Mauritsen, M. D. Salman, and H. N. Erb.** 1988. The enigma of herd: a comparison of different models to account for group effects in multiple logistic regression analysis. Acta Vet. Scand. Suppl. **84:**462-465.
36. **Kalbfleisch, J. D., and R. L. Prentice.** 1980. The statistical analysis of failure time data. John Wiley & Sons, N.Y.

http://prime-pdf-watermark.prime-prod.pubfactory.com/ | 2025-02-28