In: F. Griffin and G. de Lisle (eds) Tuberculosis in Wildlife and Domestic Animals. Proceedings of the Second International Conference on *Mycobacterium bovis*, University of Otago, 28 August - 1 September 1995. Otago Conference Series 3, University of Otago Press, Dunedin, New Zealand, 247-251.

Epidemiological Methods in Disease Investigation

Dirk U. Pfeiffer

Department of Veterinary Clinical Sciences, Massey University, New Zealand

Introduction

Among the major aims of epidemiology are to *describe* the health status of populations, to *explain* the etiology of diseases, to *predict* disease occurrence and to *control* the distribution of disease (Kleinbaum *et al* 1982). An understanding of causal relationships is the basis of the last three objectives. Such associations between causes and disease occurrence can be discovered by individual *case studies*, by experimental *laboratory studies* and by *field studies*. *Case studies* focusing on individual sick animals have long been at the center of clinical knowledge. They are based on direct personal observations relating to anatomical structure and physiological function which can be quantified and systematic but are still largely qualitative. While these observations can be extremely intensive and detailed their disadvantage is their subjectivity and the possibly extreme variation between cases (Susser 1973). In the *laboratory experiment* - the classic experiment- great precision in measurements and optimal control of influencing variables can be achieved resulting in sound inferences. The disadvantage is that it is usually not possible to represent the myriad of factors affecting disease occurrence in the natural environment of the animal and it may be difficult to work with sufficient animals to represent true variation between animals in the natural population. A *field study* is conducted in the natural environment of the animals and measurements are made on sick as well as healthy animals. The differences between sick and healthy animals can be described with respect to presence or absence of potential risk factors. With this type of study animals are exposed to all the known and unknown environmental factors present in their natural environment.

Causality

A primary concern of veterinary research is the search for etiologic causes of disease. A approach which historically has been particularly successful was based on Koch's postulates. This *deterministic* methodology required that if an agent was to be the cause of a disease (1) that it be present in every case of the disease under appropriate circumstances (necessary condition), (2) that the agent should occur in no other disease as a non-pathogenic event (specificity of effect) and (3) that it must be possible to isolate the agent from the body and induce disease in susceptible animals (sufficient condition). Among the disadvantages of this approach are that it cannot account for diseases with multifactorial etiology, factors which have multiple effects or dose-response relationships and imperfect knowledge about the underlying mechanisms in a biological system. As a consequence of investigations into the relationship between smoking and lung cancer a more general approach to causal interpretation has been formalized by Hill (1971) and later on by Evans (1978). These epidemiological criteria for investigating cause-effect relationships include the following factors:

- Proportion of disease in exposed animals higher than in non-exposed animals.
- Exposure more common in diseased than in non-diseased animals.
- Exposure precedes the disease.
- Presence of spectrum of measurable host responses.
- Elimination of exposure reduces proportion of disease.
- Prevention or modification of host response decreases/eliminates disease.
- Disease can be experimentally reproduced.

Field Studies

Field research is *empirical* and involves measurement of variables, estimation of population parameters and statistical testing of hypotheses. It is of a *probabilistic nature* in that as a result of a population study it will not be possible to predict with certainty which animal will develop a particular disease given the presence of certain risk factors. But it will be possible to predict how many cases of the disease will occur in the population in the future. Field research involves comparisons among groups in order to estimate the magnitude of an *association* between a putatively causal factor and a disease. The objective is to assess if there is a cause-effect relationship between a single or multiple risk factors and the disease.

Scientific Method

Approaches for connecting observations and theories have been studied by philosophers since the beginning of the last century (Kuhn 1970, Pearson 1937, Popper 1960). There have been many attempts to characterize a generic methodology applicable to all scientific disciplines called the *Scientific Method*. Kleinbaum *et al* (1982) describe an idealized version of the *Scientific Method* which is particularly suited to epidemiological research (see Figure 1). It is based on a continuous cycle evolving around the development of hypotheses and subsequent empirical studies resulting in adjusted, redefined or new hypotheses. Study design is a key stage in epidemiological research as it requires the researcher to generate testable hypotheses from conceptual hypotheses. It has to be accepted that every study design has its limitations. During the inferential process it is of extreme importance that the researcher is aware of these limitations.

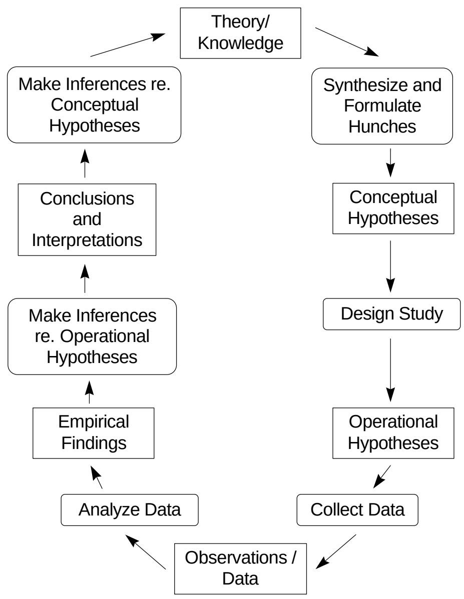

Figure 1: Idealized Conceptualization of Scientific Method (from Kleinbaum *et al* 1982)

Study Design

The interrelatedness of phenomena within a biological system complicates the situation for an investigator who must select a segment of the system for the investigation (Susser 1973). Attempting to isolate the segment from the rest of the system can result in an outcome which does not represent the real situation in the system anymore.

Epidemiological studies can be broadly divided into *experimental* and *observational studies*. The first category, *experimental studies* (clinical trials, intervention or non-observational studies), involves manipulation of study factors by the investigators. An attempt is made to simplify observation by *creating* suitable conditions for the study. The second group, *observational studies*, does not allow any artificial manipulation of the study factors. Here, the investigator is only allowed to *select* suitable conditions for the study.

Having to choose between these two major approaches the well-conducted *experiment* will allow more precise testing of the hypothesis than the *observational study*. But a properly conducted *observational study* is more likely to represent the biological system under study. It also allows testing of a wider range of hypotheses at the same time than the *experimental study*. It is also more difficult to enroll large populations for an *experimental* than for an *observational study*. Maintaining an experimental intervention over time can be difficult and costly.

Experimental studies

These types of studies typically involve dividing a group of animals into a subgroup which is being *treated* and another subgroup which is being left *untreated* and acts as a control. The decision to treat an animal or leave it untreated can be based on random allocation - *randomization*. After a period of time the status with regard to a response variable (e.g. disease status) is assessed for each animal. Summary measures of the response are then compared between both subgroups. Differences in the summary values suggest the presence of an effect of the *treatment* on the response variable. Such experimental studies can be conducted as *laboratory experiments* or as *field studies* such as *clinical trials*. The latter are usually used to evaluate therapeutic or preventive effects of particular interventions, but are also useful to investigate etiologic relationships.

The *experimental study* provides the researcher with effective control over the study situation. If the sample size is large enough a well-designed experiment will limit the effect of unwanted factors even if they are not measurable. Control of factors other than treatment which are likely to have an effect on disease can be achieved by using them to define homogeneous subgroups - *blocks* - within which treatment is allocated randomly. The possibility to have excessive control over the study situation can become a weakness of the experimental approach as it may not be representative of the situation in the biological system any more.

Observational studies

In epidemiological research the *observational study* is one of the most frequently used techniques. There are three major study designs. The *cohort study* is based on selecting two groups of non-diseased animals, one exposed to a factor postulated to cause a disease and another one unexposed to the factor. They are followed over time and their change in disease status is recorded during the study period. In a *case-control study* sufficient animals with the disease (*cases*) and without the disease (*controls*) are selected. Their status with regard to potential risk factors is then examined. In a *cross-sectional study* a sample of a population is taken at one point in time and individual animals included in the sample are examined for the presence of disease and their status with regard to other risk factors.

The *cohort study* is the *observational study* design which provides the best evidence for the presence of *cause-effect relationships* because the putative cause has to be present before disease occurs. But as it is based on pure observation within a largely uncontrolled environment it is possible that there are still other unmeasured factors which have produced the apparent cause-effect relationship. The *cohort study* is inefficient for studying rare diseases, which is a particular strength of the *case-control study*. A carefully designed *cross-sectional study* is more likely to be representative of the population under study than a *case-control study*. With regard to the development of new etiologic hypotheses this can be done efficiently with a *cross-sectional*, but less with *cohort studies*.

Conclusion

Awareness of the limitations and advantages of particular study designs is essential during the planning, analysis and interpretation phases of epidemiological studies. Experimentation and determination of biological mechanisms provide the most direct evidence of a causal relationship between a factor and a disease. Epidemiological field studies can provide strong support for causal hypotheses. Combined epidemiological and other evidence can lead to the conclusion that a causal hypothesis becomes highly probable.

References

Evans,A.S. 1978. Causation and disease: A chronological journey. *Am. J. Epidemiol.* 108: 249-258.

Hill,A.B. 1971. *Principles of medical statistics.* 9th edition. Oxford University Press, New York.

Kleinbaum,D.G., Kupper,L.L. and Morgenstern,H. 1982. *Epidemiologic research - principles and quantitative methods.* Van Nostrand Reinhold, New York, 529pp.

Kuhn,T.S. 1970: *The structure of scientific revolutions.* 2nd edition. University of Chicago Press, Chicago, 210pp.

Pearson,K. 1937: *The grammar of science.* 3rd edition. Dent, London, 357pp.

Popper,K.R. 1960: *The logic of scientific discovery*. Hutchinson, London, 479pp.

Susser,M. 1973: *Causal thinking in the health sciences - concepts and strategies of epidemiology*. Oxford University Press, New York, 181pp.