MEMO

TO: All Faculty Department of Epidemiology

FROM: John Cassel

DATE: June 4, 1975

RE: Outline of Methods Section of Epid 160

With the increasing autonomy of the sections in Epid 160 and looking forward to the day when there will be a series of independent courses on the Principles of Epidemiology, I thought it might be helpful for faculty, teaching assistants and possibly students to have a more explicit outline of the methodological principles that we have attempted to teach in the past.

The enclosed is an attempt at this. It is meant to be a guide not a set of instructions. While the order in which the material is presented appears to me to be reasonably logical, there is no good reason why this order need be maintained in each section/course.

The guide and accompanying reading material and exercises have been designed to be handed out to students and to provide the framework for the structuring of the section/course. However, certain faculty members may wish to modify this (if they choose to use the guide) by either handing out the guide but using their own reading material and exercises or alternatively restrict the guide for their own purposes in organizing their course and not hand it out.

I think if the articles enclosed here are read and digested, the only other reading I would recommend for nonepidemiology majors would be Friedman's Primer of Epidemiology. For epidemiology majors however I think that in addition they should read McMahon and if possible Jerry Morris.

I welcome your comments.

JC:jf

GUIDE TO PRINCIPLES OF EPIDEMIOLOGY
(EPID 160)

Contents:

Course objectives

Principles of method

1. Associations
2. Causal relations
3. Strategies and Research Design

Examples, exercies and reading material

Recommended text

Non-epidemiology majors - Friedman, Primer of Epidemiology

Epidemiology majors - McMahon, Epidemiologic Methods
Friedman, Primer of Epidemiology
Morris, Uses of Epidemiology

OBJECTIVES AND GENERAL COURSE OUTLINE

Epidemiology may be viewed both as a specific body of knowledge concerning various states of health and as a method of study. Thus it is appropriate to talk of 'the epidemiology of" typhoid fever or lung cancer, for example (i.e., the specific body of epidemiological knowledge concerning those two diseases) and also to talk of "epidemiological investigation" to determine the factors responsible for any disease or disorder. This course is concerned mainly with the principles underlying epidemiology as a method of study and the scope, potentialities and limitations of this approach.

In the minds of many, the objectives of epidemiological investigation are restricted to discovering the factors responsible for an outbreak or epidemic of some infectious disease. Modern epidemiologists regard this as only one contribution of epidemiology. The scope and uses of epidemiological study have been considerably broadened. This point will be amply documented in this course.

Stated formally the objectives of this course are:

1. To develop a conceptual model of epidemiological enquiry as the basis for scientific public health practice.
2. To illustrate the scope and uses of epidemiological enquiry.
3. To familiarize students with the basic principles of the observational sciences (of which epidemiological enquiry is one).
4. To teach a number of the more important aspects of epidemiological method.

To accomplish these objectives the course will be divided into a lecture and a laboratory/seminar series. The lecture series will be con-

cerned with the philosophy, principles and methods of epidemiology. The laboratory/seminar series will review and illustrate these principles using various areas of application.

PRINCIPLES OF EPIDEMIOLOGY

The principles underlying epidemiologic method can be stated quite succintly. All epidemiological inquiry is based upon finding answers to two sets of questions.

1. Is there an association between a set of characteristics (of the populations under study and/or their environment) and the health conditions under study?

2. How strong is the evidence indicating that such an association is likely to be causal?

To answer these questions the following rules of evidence need to be taught and illustrated

1. Is there an association?
   Answers to this question require determining whether the correct comparisons have been made.

   The two most common sources of error in this regard are due to:
   a) Assuming that an association exists based upon comparison of cases (with and without the condition) only, i.e. no controls - so called numerator data only (see Example 1)
   b) Assuming that an association exists at the individual level when the characteristic is measured solely at the ecological level. (so called ecological fallacy) (see Example 2)

2. If there is an association is it likely to be causal?
   2.1 Categories of non-causal association
      2.1.1. Association due to chance (i.e. sampling variation)
      Assessed by statistical test of significance

      2.1.2. Association produced by artifact (i.e. artificially produced by the way in which the study was done)
      Two major sources of artificial (or spurious) associations

         2.1.2.1. BIAS (i.e. false labelling of either characteristic or condition)
         Bias can be produced by lack of reliability and/or validity

            Sources of lack of reliability
            - laboratory error
            - inter and intra observer error
            - selective recall (e.g. cases remembering certain characteristics differently from controls)
            - changes in diagnostic custom over time
            - diagnoses made only after knowledge of presence of presumed causal characteristic

Sources of lack of validity:
- criteria and/or techniques used to assess conditions or characteristics do not truly measure condition or characteristic.

Assessment of Bias

Measures of reliability
where possible measures of inter and intra observer concordance.
Measures of validity
requires external validating criterion measures
Measures of sensitivity and specificity
(See 1) Comparibility in International Epidemiology, Roy M. Acheson, ed. Milbank Memorial Fund, 1965 for useful examples
2. Xerox from Medical surveys and Clinical Trials L.J. Witts, ed. Oxford University Press, 1954 (enclosed) for a good discussion)

In many instances no direct test available to assess possibility of bias and indirect methods have to be used based upon knowledge of phenomenon and application of logical thinking.
e.g., To determine whether a rise in mortality for a given disease over time could be due entirely to changing diagnostic criteria it would be useful to examine the rates in both males and females. If the rise has been mainly in males for example - as in lung cancer with a much less rise in females, it is unlikely that the increase is solely due to such a bias.

(See Example 3 especially pages 171-174 in which such an indirect test for bias was used in a study)

2.1.2.2. Selection

By some fault in the research design it has been made easier for people in whom there is an association between the characteristic and the condition to be included (or excluded) in the study population

Perhaps the best example of the effects of selection is the so called "Healthy worker effect" described by McMichael and Tyroler.

2.1.3. Secondary Associations

Secondary associations are produced by confounding factors. These are factors associated with both the characteristic of interest and the condition.

The term secondary association refers to the association between the characteristic of interest and the condition which has been produced by the association of both these with the confounding factor.

e.g. Coffee drinking found to be associated with myocardial infarction.
But cigarette smoking (a possible confounding factor) found to be associated with both coffee drinking and myocardial infarction. (Coffee drinkers smoke more than do nondrinkers People with myocardial infarction smoke more than do people without myocardial infarction).
Controlling for cigarette smoking eliminates the original association between coffee drinking and myocardial infarction - indicating this was a secondary association.

Assessment of Confounding Factors

The determination as to whether an association is secondary (i.e. produced by confounding factors) or not can obviously be determined only if data on the potentially confounding factors has been gathered. (This incidentally is one of the areas requiring the greatest skill in designing an epidemiological study or in critically reviewing the results of one - the determination upon the basis of knowledge of the phenomenon whether the appropriate potentially confounding factors have been taken into consideration.

If the data are available there are numerous techniques for determining whether confounding factors can account for any association found. (e.g. matching, contigency tables, adjustment, co-variance partial correlations, etc.)

For purpose of this course it is suggested that two approaches be taught.

1) The logic of contingency tables (of which we have many examples from previous laboratory exercises)
2) One form of adjustment the Standardized Mortality Ratio (SMR) - See Example 4 and xerox from Mausner and Bahn, Epidemiology: An Introductory Text. (enclosed)

2.2 Evidence stregthening the possibility of causal associations
Having excluded (as far as possible) non-causal reasons for an association, the following evidence can be used to stregthen the possibility that an association is causal.

2.2.1 Knowledge of time relationship (i.e. which is antecedent and which consequent)

2.2.2. Strength of the association. The stronger the association the less chance that this could have been produced by unknown confounding factors.

2.2.3. Consistency of association. Repeatedly observed by different persons in different places, circumstances and times and by different research designs.

2.2.4. Dose response relationship

2.2.5. Coherence. Does not contradict known phenomena about distribution in population.

2.2.6. Experiment
(See enclosed xerox Bradford Hill, "The Environment and Disease: Association or Causation?" Proc. Roy. Soc. Med. 58: 1965, 295-300.

3. Strategies and Logic of Different Forms of Research Design
For cohort, case-control and cross-sectional studies:

3.1. Strategy: i.e. Basis of classification
Time sequence
Nature of sample (cohort -those free of condition
case-control -cases and controls(nonrepresentative usually)
cross-sectional-survivors with and without condition)

Comparisons to be made

3.2. Rates that can be assessed
Cohort: Incidence: Definition and interpretation possible
Case Control: Proportion of cases with characteric (i.e. no rates)
Cross-sectional: Relative and attributable risks.

3.3. Risk estimates:
Cohort: Relative and attributable risks
Definitions and interpretations
Case-Control: Relative odds
Cross-sectional: Relative and attributable risks

3.4. Advantages and Limitations:

3.4.1. In inferring causation

	Cohort	Case-Control	Cross-Sectional
Antecedent-consequent	+	-	-
Selective recall (bias)	+	-	0
Selective survival	+		
Attrition	-	$\overline{+}$	$\overline{+}$
Direct quantitation of strength of association(risk)	+	-	0

Key: + strength
- weaknesses
0 variable

3.4.2. Administrative and Logistic

	Cohort	Case-Control	Cross-sectional
Time and money	-	+	0
Rare Diseases	-	+	-

See example 5 (from Laboratory Manual) and xeroxes

Richard Doll: Retrospective and Prospective Studies: In Medical Surveys and Clinical Trials. L.J. Witts, ed. Oxford University Press. 1964.

Jerome Cornfield and William Haenszel "Some Aspects of Retrospective Studies." J. Chronic Dis. 1960, v.11, 523-534.

Phillip Sartwell: "Retrospective Studies; A Review for the Clinician." Annals of Internal Medicine . 81: 1974, 381-386.

Bayes Theorem and Its Application in Determining Relative Risk in Case Control Studies . Mineographed by David Kleinbaum.