MEDICAL EDUCATION

The Clark-Omran System of Research Design in Epidemiology

By Abdel R. Omran*

PROJECT IDENTIFICATION

Topic
Principal investigator (name, degrees, position, address)
Institutional affiliations
Funds required for first year of research
Total funding required for project
Research period (dates of proposed research)

STEP I SELECTION AND FORMULATION OF RESEARCH PROBLEM

Step I requires the investigator to choose an appropriate research topic, to identify research questions and the possible value of seeking answers to these questions, to make a clear statement of research objectives, and to define key terms.

1.1 Choice of a research topic
1.2 Definition of nature, extent and significance of the problem
1.3 Framing of specific research questions
1.4 Statement of objectives, immediate and ultimate
1.5 Provision of workable definitions of key terms

STEP II APPRAISAL OF EXISTING INFORMATION AND RECONSIDERATION OF PROBLEM AND RESOURCES

The second step is for the investigator to familiarize himself with the existing knowledge about his research problem and to discover whether or not others have investigated the same or similar problems before. The investigator is thus able to compare his results with those of others, to learn from their experience, their approaches and any new techniques they have developed. This step is accomplished by a thorough and critical review of literature and personal communication with experts. The investigator should then decide whether or not he needs statistical, methodological and/or technological consultations. It is at this early stage in research design, not after collection of data, that such methodological and statistical collaboration should be sought and obtained. The investigator should also reconsider his research interests, the resources available to him, and the requirements of the project in selecting specific questions to be tested.

2.1 Appraisal of existing information
- 2.1.1 Literature review and search for other sources on the subject
- 2.1.2 Classification of existing pertinent information on the subject
- 2.1.3 Critical appraisal of existing information

2.2 Consideration of statistical and technical collaboration

2.3 Reconsideration of interests and resources
- 2.3.1 Consideration of research needs and investigator's interests
- 2.3.2 Consideration of funds and personnel

2.4 Selection of specific research questions to be tested

STEP III FORMULATION AND STATEMENT OF RESEARCH HYPOTHESES

The value of scientific work depends heavily on the originality and logic with which hypotheses are formulated. A hypothesis is a shrewd supposition, or

* The University of North Carolina at Chapel Hill

inference, adopted to explain observations and/or to guide further investigation. It may be simply an educated answer to a specific question. The hypotheses may be derived from the body of knowledge on the subject, from the experience of the investigator or of others, or from previous research endeavors. It must therefore be emphasized that hypotheses are not meant to be haphazard guesses but should reflect the depth of knowledge, imagination and experience of the investigator.

STEP IV RESEARCH PLANS FOR TESTING HYPOTHESES

Plans for testing the hypotheses that have been formulated require selection of the project's research strategy and setting, definition of the unit of observation, the methods and particulars of observation, and an outline of the sampling procedures and the choice of controls. Plans for analysis of the data to be collected should also be made. Administrative details are also to be specified.

4.1 Selection of the research strategy

The selection of a research strategy is the core of research design and is probably the single most important decision the investigator has to make. The choice of strategy, whether descriptive, analytic, experimental, operational, or a combination of these, depends on a number of considerations. Because of the importance of the research strategy, the specific types of studies are described briefly in the Appendix, along with a discussion of the advantages and disadvantages of each.

4.1.1 Descriptive, validating and surveillance strategies

4.1.2 Observational or analytic strategies
- a. Cohort (or prospective) study
- b. Historical (or reconstructed) cohort study
- c. Case history (case-control or retrospective) study
- d. Cross-sectional study
- e. Follow-up study

4.1.3 Experimental strategies
- a. Animal studies
- b. Therapeutic clinical trials
- c. Prophylactic clinical trials

4.1.4 Operational strategies

4.2 Selection of the research setting

The research setting includes all the pertinent facets of the study, such as the population to be studied, the place and time of the study, the type of observation, and the collaborating institutions, if any. The investigator is entitled to choose a convenient setting, but he should be careful not to sacrifice appropriateness of methodology to convenience.

4.2.1 Selection of the population, place and time

4.2.2 Definition of the unit of observation and data to be collected

4.2.3 Choice of methods and particulars of observation

4.2.4 Consideration of ethical problems

4.3 Sampling

A sample is a part of the whole population — alternatively called the universe or reference population. Sampling is the process or technique of selecting a sample of appropriate and manageable size for study. In epidemiologic investigations and in health work in general, it is almost always possible to deal with a sample drawn from a reference population or universe. This universe may be a population of people (healthy and sick), a population of cases of certain diseases, the clientele of a family planning clinic, or recipients of a certain treatment, for example. The universe may not be people at all, as in the case of a universe of birth or death certificates or of medical records; or the universe may consist of health centers, village units or hospital units.

4.3.1 Selection of the sampling procedure
- a. Probability sampling methods : random, systematic, and stratified sampling; use of panels; area, cluster sampling; other
- b. Non-probability sampling : quota and convenience sampling; other

4.3.2 Determination of sample size

The sample should be of sufficient size to be dependable and to allow tests of statistical significance to be applied. Statistical methods are used in estimating sample size.

4.3.3 Plans to assure representativeness and reliability of sample and to minimize

sampling errors (such as non-response and selection)

4.4 **Controls and case allocation**
- 4.4.1 Matching
- 4.4.2 Random allocation
- 4.4.3 Alternation
- 4.4.4 Population as control
- 4.4.5 Before and after control
- 4.4.6 Analysis of subgroups

4.5 **Plans for testing of equality between sample and control groups**

4.6 **Plans for minimizing non-sampling errors such as observer error; use of volunteers, paid participants or private patients; and-errors of coverage, recording, or data processing.**

4.7 **Development of study instruments**

- 4.7.1 Questionnaires and interviews
  - a. Preparation, precoding and pretesting of questionnaires
  - b. Preparation of instruction manual
  - c. Translation of questionnaires in cross-cultural studies
  - d. Development of editing and coding plans
  - e. Plan for interviews.
  - f. Training of interviewers
- 4.7.2 Other methods of observation
  - a. Medical examination
  - b. Laboratory tests
  - c. Screening procedures
  - d. Unobtrusive observation
- 4.7.3 Design of recording forms (which may be the same as analysis forms)
- 4.7.4 Plans for reliability and validity checks

4.8 **Plans for Analysis**

The investigator is well advised to incorporate plans for analyzing his data in his research proposal to increase the confidence of reviewers in the competence of the research team in handling the data once they are collected.

- 4.8.1 Design of analysis forms
- 4.8.2 Selection of data handling techniques
  - a. Manual sorting
  - b. Machine sorting
  - c. Computer program(s)
  - d. Record linkage
- 4.8.3 Choice of statistical methods to be applied to each hypothesis
  - a. Descriptive statistics
  - b. Tests of significance
  - c. Statistical models and other methods
  - d. Content analysis
  - e. Numerator analysis
  - f. Cohort analysis
- 4.8.4 Design of dummy tables and graphs (skeleton of anticipated tables and graphs)

4.9 **Plans for collecting data**

- 4.9.1 Organization of study and data collection
- 4.9.2 Plans for personnel training (if applicable)
- 4.9.3 Plans for pilot or feasibility studies and pretesting methods

4.10 **Budgeting**

- 4.10.1 Personnel (salaries, fringe benefits, social security, retirement)
- 4.10.2 Consultant fees
- 4.10.3 Equipment and rental of space
- 4.10.4 Supplies
- 4.10.5 Travel, domestic and international
- 4.10.6 Analysis costs (keypunching, verifying, programming and computer time, etc.)
- 4.10.7 Miscellaneous expenditures (for example, illustrations, documents, mail)
- 4.10.8 Overhead

4.11 **Timetable for data collection and analysis**

4.12 **Other administrative details for proposed research**

- 4.12.1 Description of facilities available to investigator (for example, computers, office space)
- 4.12.2 Other sources of research support
- 4.12.3 Biographical sketches of principal investigator and co-investigators with emphasis on previous experience in fields related to research project
- 4.12.4 Permission to investigate human subjects, if applicable

STEP V COLLECTION OF DATA

After a research proposal has been accepted, the investigator should proceed to collect the data following the details of the proposal. Special emphasis should be given to the following items.

5.1 **Personnel recruitment and training**

5.2 **Pilot and feasibility studies and pretesting of instruments**

5.3 **Conduct of study according to design**

5.4 **Record-keeping and editing**

5.5 **Problems related to collection of data**

5.5.1 Sampling problems and errors
5.5.2 Non-sampling problems and errors

STEP VI PROCESSING, CLASSIFICATION AND ANALYSIS OF DATA

Utmost care should be taken in data processing, classification and analysis. Tabulation should be specific and directed towards testing the hypotheses; creating a huge number of tables, many of them irrelevant, is a common error that slows the progress of analysis and cuts into the budget and the time of the investigators. Progress reports should be short but adequate and should meet specific deadlines.

6.1 **Data processing according to plan**

6.2 **Tabulation and graphing**

6.3 **Analysis according to plans**

6.4 **Preparation of progress reports, as required**

STEP VII INTERPRETATION AND CONCLUSIONS

Interpreting data and drawing conclusions form the climax of any research project. Careful, unbiased and critical interpretation of data requires great skill and experience on the part of the investigator. Such skill can be developed by appraising well-designed epidemiologic studies, both classical and modern, by learning from the experioence of others, and by participating in scientific conferences.

7.1 **Interpretation**

7.1.1 Accurate and unbiased evaluation of results
7.1.2 Determination of causation *versus* association
7.1.3 Determination of population to which results may be referred
7.1.4 Checking for ecologic fallacy

7.2 **Drawing conclusions and explaining practical applications of findings**

7.3 **Outlining of future research needs**

STEP VIII REPORTING : OUTLINE FOR FINAL REPORT AND PUBLICATION

In reporting results of anvestigation, one should use the format specified by the journal in which the results will be published. While one journal differs from another in format, especially in such details as footnote or reference form, the following general outline is common to many.

8.1 **Title**
8.2 **Author(s)**
8.3 **Introduction : background, objectives, hypotheses**
8.4 **Material and methods**
8.5 **Results**
8.6 **Discussion**
8.7 **Summary**
8.8 **References, footnotes and/or bibliography**
8.9 **Tables**
8.10 **Graphs**