Research Design in Family Medicine

Jack Froom, MD, Larry Culpepper, MD, Lorne Becker, MD, and Vincenza Boisseau
Rochester, New York

"The unity of all sciences consists alone in its method, not its material . . . it is not facts themselves which make science but the method by which they are dealt with."[1]

The process of research begins with a clear statement of question. One or more hypotheses may be formulated, and a project then designed to test the hypotheses or to describe certain occurrences. Most aspects of research require creative, imaginative activity. The process of question identification derives from knowledgeable curiosity and cannot be taught by traditional pedagogy; however, researchers may be guided to those areas most amenable to productive investigation and given help with both hypothesis formulation and research design. Research topics vary widely within any discipline. The practice of family medicine draws not only from traditional medical and surgical fields but also from the spheres of sociology, psychology, epidemiology, and public health. The content of family medicine is such that its investigators are often obliged to borrow heavily from the research techniques of other fields in order to develop appropriate investigative design for family medicine's questions.

In a discipline which is new, as is family medicine, studies which establish the content and process of that new discipline provide essential background material for future research. In such areas of investigation hypotheses need not be proposed. It is important to question what is happening or what has happened without introduction of the potential problems which accompany the additional queries of How? or Why? Superb documentation of the content of family medicine (or general practice in the United Kingdom) has been provided over the relatively recent past by workers from the United States,[2,3] Canada,[4] and England.[5]

Given such background, current research efforts in family medicine are increasingly directed toward elaboration of causal and comparative factors and, as such, hypotheses must be set forth.

Dr. Jack Froom is Associate Professor of Family Medicine in Medicine, Pediatrics, and Community Health; Dr. Larry Culpepper is Clinical Assistant Professor of Family Medicine in Medicine, Pediatrics and Community Health; Dr. Lorne Becker is Assistant Professor of Family Medicine in Medicine, Pediatrics and Community Health; and Ms. Vincenza Boisseau is Research Associate. All are of The Family Medicine Program, University of Rochester-Highland Hospital, Rochester, New York.

0094-3509/78/0703-0075$04.00

Formulation of appropriate hypotheses presumes working knowledge of the specific problems under investigation. In medical practice physicians often identify questions associated with patient care, and while many may propose creative explanatory hypotheses, relatively few have the research experience essential for development of rational hypothesis testing mechanisms. It is in this area, that of proof (or disproof) of hypotheses, that the family physician most often seeks the assistance of workers from other disciplines more traditionally research-oriented (ie, epidemiology, statistics). In all such interdisciplinary investigation, successful results require active, continuous collaboration from all parties concerned with the project. An excellent example of such collaboration is the work of the Private Practice Research Group in Pennsylvania. This group, which studies pharmacologic agents in the family practice setting, has demonstrated the feasibility of cooperation between "research-experienced clinical investigators and research-interested family physicians."[6] Their productive efforts may serve as a model for a variety of family medicine research endeavors.

The most critical contribution to be made to any investigation is a carefully conceived project design, although a science of investigative procedure does not exist as such. Suchman's analysis of research design[7] includes five penetrating statements which are particularly applicable to family medicine research. These are:

1. *It seems to us futile to argue whether a certain design is 'scientific.' The design is the plan of study and, as such, is present in all studies. It is not a case of scientific or not scientific, but rather of good or less good design.*

Good research design requires that its developers have: knowledge of design principles, clear understanding of the problems under consideration, and the ability to recognize potential sources of bias.

2. *The proof of hypotheses is never definitive. The best one can hope to do is to make more or less plausible a series of alternative hypotheses.*

Hypotheses may be strengthened by well-designed projects relatively free of bias, and confirmatory evidence may be contributed from other studies. Familiarity with statistics in general, and the concepts of probability and significance in particular, are important for interpretation of results.

3. *There is no such thing as a single 'correct' design. Hypotheses can be studied by different methods using different designs.*

There are numerous equally valid technical approaches to solution of any given research problem. The nature of family medicine and the data generated by its practice preclude the use of certain types of procedural design and render others particularly valuable. In some instances, completely new design structures may be needed.

4. *All research design represents a compromise dictated by the many practical considerations that go into social research.*

The choice of design is influenced by the availability of resources. For example, the family physician who engages in clinical practice may find that constraints of patient care dictate the design that is eventually chosen.

5. *A research design is not a highly specific plan to be followed without deviation but rather a series of guideposts to keep one headed in the right direction.*

A pilot project is practically always advisable since a minimum of data from a small project may provoke intelligent modification of the original design. It is not unusual that several design changes are deemed necessary between initiation and completion of a project.

A number of methods and designs have been employed to study problems of interest in family medicine. Several of these will be discussed as they specifically relate to family medicine research projects. An example of one of these, the retrospective study utilizing patient medical records, will be documented in detail. Although, in the past, research employing data generated from records has been hampered by inaccuracies and desultory recording, the introduction of problem-oriented records and specific indices to tally patient population data has created an opportunity to take advantage of this approach in new and exciting ways.

Numerous investigative techniques have been developed by the disciplines of sociology, epidemiology, public health, health care administration, and others interested in issues of health, disease, and medical care delivery. Each has been demonstrated to be useful for the elucidation of specific problems and each has its own particular advantages and disadvantages. Although an exhaustive review of the subject of research de-

sign is beyond the scope of this paper, several aspects of design pertinent to family medicine research will be considered.

Retrospective vs Prospective

Use of the terms, retrospective and prospective, so seemingly simple, is nonetheless confusing within the scientific literature. Although root meanings remain the same (to look back and to look forward), there are three commonly accepted interpretations of each term. Each will be examined in light of its impact upon the conceptualization process so essential for carefully established research design.

1. *One construct of the terms retrospective and prospective, primarily elaborated in epidemiologic studies, may be considered to be fundamentally related to causation and observation. Temporal implications of both terms, in this case, are related to the sequence of events rather than to collection of data.*

For illustrative purposes, a theoretic model may be used: given end result "B" (established), what can be determined concerning causal factor(s) "A?" (retrospective). Or, given causal factor "A" (established), what can be determined concerning end result(s) "B?" (prospective).

Examples of these approaches are abundant. A recent study of the relationship between oral contraceptives and development of thromboembolic conditions[8] was accomplished in a retrospective manner by these definitions. Occurrence of thromboemboli in women (effect "B") was postulated to be related to intake of oral contraceptive agents (cause "A"). A group of women hospitalized with thromboembolic disease was matched with a similar group hospitalized for other conditions. The use of oral contraceptives was established to be higher among the group with thromboemboli.

The Royal College of General Practitioners' Oral Contraceptive Study[9] evaluated the same relationship prospectively. In this instance, women receiving oral contraceptives (cause "A") were matched with a group not receiving oral contraceptives. Development of thromboemboli (effect "B") in each group was recorded and incidence was compared. The hypothesis that intake of oral contraceptives is related to development of thromboemboli was thus further tested and results offered additional confirmatory evidence that the hypothesis was more likely to be valid.

In general, both types of investigations suggest causal relationships; however, prospective studies are usually considered more powerful determinants of causation. Most simplistically, a retrospective study deals with a single effect but may elucidate a number of possible causes, while a prospective study investigates a single hypothesized cause which may provoke a number of different effects. In the usual sequence of events, retrospective studies are first accomplished to determine a hypothetical relationship. Prospective studies frequently follow to establish further evidence for validity of the hypothesized relationship and to calculate incidence.

Given this first interpretation, it is immaterial if data for either type of study are obtained from existing records or are gathered specifically for purposes of the investigation.

2. *The second interpretation of the terms retrospective and prospective does relate exclusively to temporal factors in data collection.*

Briefly, retrospective studies are conceived as those which use available data from existing records, while prospective studies generate new information. The primary disadvantage to retrospective data collection relates to the frequent absence of critical information. Although patient records may provide valuable data, they were initially produced as an aid to patient care and not to answer research questions. Another source of existing information, medical data banks, have customarily been instituted to meet specific needs. Caution should be observed in the use of such information since it may have been gathered in a manner consistent with the initial purpose, but be inappropriate to the needs of the present study. Although retrospective data collection is admittedly cumbersome, there are certain types of information which can be gathered by no other means. When questions are general or when the required populations are far in excess of the available sample, data may be gathered more efficiently from records at hand. An additional negative factor is the inability to construct a well-controlled situation. By definition, a retrospective study implies retrospective control, and the investigative process cannot be so clear cut as a precisely controlled

prospective study. It is perhaps for this reason that statisticians are frequently disdainful of the retrospective process as illustrated by the following vignette from W. G. Cochran:

> *Moreover, I owe my first post in the Depression mainly to the fact that my employers had a large batch of post data which were regarded as a potential mine of information. They hired me to dig it out. I dug furiously, but I doubt they received their money's worth. Fortunately, my salary was so low that this moral problem caused me no loss of sleep.*[10]

Retrospective studies, however, do have certain important advantages. Since the data and records already exist, they are readily accessible. Informed consent and cooperation of subjects are generally not important considerations. Certain sources of bias are more easily avoided, since the initial recordings were accomplished without either patient or recorder awareness of the eventual use of the data. On the other hand, the investigator may be biased by a selective choice of existing data to be used.

Several important family medicine studies have been totally retrospective in nature by this definition. Descriptions of the content of family practice[2,3] were essentially retrospective use of data generated for multiple purposes including office management, continuing education, audit,[11] and outreach.[12] Results of these studies have helped to define the scope of family medicine and have contributed substantially to development of curricula for teaching.[13]

Prospective collection of data was used in a family medicine study to investigate factors associated with elevated lead levels in children.[14] Screening tests were performed on 333 children. The group with elevated lead levels was compared to the group with normal levels. No differences were found in age or sex distribution of the two groups, but a higher proportion of children with elevated lead levels resided in areas designated as "lower," socioeconomically.

3. *The third interpretation will be mentioned only briefly since it applies only to studies of the quality of medical care and has not received wide attention.*

A retrospective study by this definition[15] measures the quality of care received by a group of patients, but makes no immediate attempt to intervene if suboptimal care is detected. At the end of the retrospective study, the areas needing improvement are pointed out to the providers and it is hoped that the quality of care received by subsequent patients improves. In a prospective study, an intervention is made whenever suboptimal care is detected so that the quality of care received by the group being studied improves as a result of the study.

Observational vs Interventive

Observational investigations are those in which no measures are taken to alter either the course of events or the situation under study. Interventive studies may be considered to be those in which a factor(s) operate(s) to potentially alter an outcome. Intervention is customarily a deliberate action on the part of the researcher. In general, observational studies are undertaken by epidemiologists and sociologists or others studying characteristics of large populations, while interventive studies are the primary vehicle for clinical and basic research.

Observational Studies

Some types of observational research designs are particularly suited to the needs of family medicine researchers. These include:

Descriptive Studies

These studies note occurrences but make no attempt to determine causation. Significant contributions can be made to current understanding of the natural history and distribution of various clinical problems by well-constructed descriptive studies. Compared to other specialty practices, family medicine patient rosters are composed of groups most representative of the demographic distribution of the population at large. A question commonly studied in family medicine research relates to the occurrence of disease syndromes or problems within a population. The usual design of such studies involves looking at a sample of the population in detail and making the inference that the same results would have been found had the

entire population been studied. This often involves a sample of a sample; that is, the study is carried out using a sample of the patients in the practice, the patients in the practice being a sample of the larger population (ie, county, state, etc). Although extrapolation of results to these larger populations is not invariably the intent of a descriptive study, it is essential to establish the demographic characteristics of the population under study, making particular note of those parameters relevant to the investigated disorder. Practices which routinely maintain indices permitting description of the patient population by demographic variables (ie, age/sex) are particularly suited to this type of study.

In a descriptive study of acute otitis media (OM) in family practice, for example, approximately 20 percent of individuals with diagnosed OM were over age 15.[16] Since most reports of OM emanate from pediatricians and there are a paucity of reports in adults, OM has been commonly considered to be primarily limited to children.

The Medical College of Virginia's report on the content of family practice[2] and the National Ambulatory Medical Care Survey[17] are excellent large-scale descriptive studies which have contributed valuable basic data for future investigations.

Although descriptive studies may be appropriately employed to study a wide variety of health related factors (eg, practice management issues), the most common application of such investigations in family medicine relates to clinical problem or disease. There are three measures of disease of particular use to the family medicine researcher; each will be briefly defined.

Incidence. Disease incidence may be defined as the rate of appearance of disease in a population. Since incidence is expressed mathematically as a rate, a denominator is needed. Incidence is usually noted as cases per hundred or per thousand population. Population, itself, must be defined since it may be population within a specific age range, total practice population, population of females or males within the practice, or other stated group.

Prevalence. Prevalence differs from incidence in that it is a measure of the existence of any given disease or problem within a given population. Thus, although the incidence of uncomplicated hypertension in family practice may be low (the number of new cases per year per hundred patients), its prevalence may be relatively high (the number of hypertensives per hundred patients).

Workload. Workload is related to patient care generated by a group of patients with a specified disorder. The most frequent measure of workload is that proportion of health care provider's total time dedicated to care of patients with a particular diagnosis.

Analytic Studies

In contrast to descriptive studies, analytic investigation requires comparison and hypothesis formulation. Hypotheses may ultimately involve constructive theorization; but initially, the null hypothesis must be tested. Very simply stated, the null hypothesis makes the assumption that no differences exist between compared groups. The null hypothesis is basic to statistical determination of the significance of determined differences. Although elaboration of statistical principles is beyond the scope of this paper, a list of relevant publications is appended.

Analytic studies involve measurement of variables. When considered in relation to cause and effect, independent variables imply cause, dependent variables imply effect. Most analytic studies in family medicine are concerned with causal or etiologic factors in disease and patterns of occurrence. Three types of analytic studies will be discussed as they pertain to observational research.

Case Control. The typical case control study first identifies a group of cases with a specific problem. A control group is then established with characteristics as close as possible to the case group except for the problem (or disease) under consideration. In epidemiologic terms, a case control study is retrospective. Its dependent variable is the disease, and independent variables are potential causal, or etiologic, factors. Comparisons are then made between the rate of occurrence of the independent variables between the two groups.

An important case control study by Doll and Hill[18] first demonstrated the relationship between smoking and cancer of the lung. A group of 649 males with bronchogenic carcinoma were matched with a control group of males from the same hospi-

tal, and frequency of smoking history was compared between the two groups. Although a potential relationship was thus defined, it was not until later studies demonstrating the carcinogenic effect of the coal tars found in tobacco that cigarette smoking was more firmly established as an etiologic agent in carcinoma of the lung.

Important considerations in the construct of a case control investigation are:

a. *Matching*. The groups must be matched for those factors related to the disorder under study. Validity of results is, to some extent, dependent upon careful, considered matching.

b. *Choice of independent variables*. Informed decision must be made regarding those variables to be measured.

Cohort. A cohort study is, in some regards, an inverse case control study. In one of an excellent series of recent publications on clinical biostatistics, Feinstein proposes use of the term "trohoc" rather than case control; "trohoc" being the familiar cohort spelled backward.[19] Until such picturesque vocabulary becomes more familiar, the term case control seems, unfortunately, to be doomed to persevere. A cohort is literally a group of persons included in an endeavor. One, two, or more cohorts may be studied, usually prospectively, to elucidate factors related to disease development.

The Framingham study is an excellent example of a single cohort study. One group (cohort) of patients was followed longitudinally to study the development of ischemic heart disease. Data from this extensive investigation, reported by several investigators,[20-22] have strengthened the hypothesis of multiple risk factors in the etiology of ischemic heart disease.

When two or more cohorts are followed, matching again becomes an important determinant of the validity of results. Characteristics of cohorts should be similar in all regards save those suspected factor(s) under consideration (independent variables) as being contributory to disease development. The studies subsequent to the initial case control studies of Doll and Hill[18] were performed using, in one instance, smoking vs nonsmoking cohorts to establish incidence of development of carcinoma of the lung.

One of the more obvious limitations of cohort studies is that they tend to be relatively lengthy and are not suited to disorders of low incidence or protracted latency periods between cause and effect.

Cross-sectional. Cross-sectional studies investigate the prevalence of disease states relative to group or subgroup characteristics. In a recent family medicine study, for example, an inverse relationship was postulated between age and bronchial asthma.[23] A 60,000 patient population was divided into subgroups on the basis of age, and prevalence of asthma was determined. Results demonstrated a threefold difference in prevalence between the group 5 to 9 years of age and the group 15 to 24, asthma being more prevalent in the younger age group. Cross-sectional studies are, in many ways, analytic descriptive studies. Studies such as the example may bear no relationship to the prevalence of asthma within the community, however, since only the health-care-seeking population is examined.

Observational studies, whether descriptive or analytic, are useful within acknowledged limits. They are valuable in establishing relationships, but only preliminary in discovering etiology. Interventive studies are the more powerful determinants of causation.

Interventive Studies

Interventive studies are classically experimental investigations. Practically all are prospective in both the epidemiologic and temporal sense, and all are analytic by definition. Few interventive studies are undertaken without prior observational determination of the appropriate factors for investigation. Discussion will center about this most deliberately constructed, carefully controlled, experimental approach to problem solving, and to those variations in design most suited to the needs of family medicine research. Choice of research design among these variations will be dictated largely by constraints of resource and material availability. Illustrative examples have been chosen from research efforts of members of this family medicine program and from recent publications in the medical literature.

Pre- and Post-Intervention

In this type of investigation, each subject/group serves as its own control. Such studies are particularly suited to determination of therapeutic efficacy of specific regimens, and to studies in which there may be wide individual variations in initial conditions (pre) which might obliterate any significant effect of the interventive factor (post). In many instances, individual changes in measurable parameters are pooled and, as is possible with finite determinations, are expressed as mean change (or mean percent change) from pre-intervention values.

In a study of the skeletal response to calcifediol (25-hydroxy-vitamin D) in renal osteodystrophy, sequential bone biopsy specimens were obtained from five patients undergoing long-term hemodialysis before and following therapeutic intervention.[24] Results were analyzed both individually and as a pool. Since skeletal compromise in this disorder varies widely among individual patients, the pre- vs post-intervention design was indicated.

A second example of a pre- and post-intervention study illustrates that this design, as well as others, may be used to evaluate certain practice management and quality of care issues. Patient expectations prior to an office visit were compared with fulfillment of those expectations following the visit. As a measure of patient satisfaction, concordance scores for expectations and fulfillment were calculated and were found to correlate with age, sex, education, marital status, and other demographic variables. Absence of a control group or a group which received a similar intervention in a different setting does not lessen the validity of such a study, but may limit broad extrapolation of results.

Experimental vs Control

In this type of study two or more cohorts are either matched or selected at random. Traditionally, one group is assigned for intervention of some type, the other remains the control with no intervention. Ideally, determinant variables are evaluated prior to (as the baseline measurement) as well as following intervention. Baseline determinations are essential to establish validity of results; those studies in which pre-intervention evaluation is either impossible or overlooked contribute little but suggestive inference to the subjects under consideration.

A recent study by Hulley and co-workers[25] on the effect of risk factor intervention on plasma levels of high-density lipoprotein cholesterol is an excellent example of an experimental vs control study in which baseline determinations were made.

Experimental vs Placebo

Groups are also established in this type of study by either random assignment or selective matchings, and baseline measurement of end-point determinants are routinely made. One group receives the intervention under evaluation and the other a placebo intervention. Measurements made again at the conclusion of the study are compared with pre-intervention values in the same group and with corresponding values in the other group. This is the classical design of drug trials. A double-blind study was recently completed at this family medicine program to evaluate the effectiveness of a gargle to relieve sore throat of diverse etiologies. Both groups received gargle: one containing the pharmacologically active ingredient, the other a placebo. Results indicated a statistically significant advantage of treatment with the active compound in relief of associated pain and dysphagia.

Although this type of investigation avoids many common sources of error, there are, nevertheless, some sources of potential bias. One of these may relate to the voluntary nature of participation in the investigation. Those individuals who declined to enter the study may have exhibited characteristics such that their inclusion would have altered the final outcome.

When relatively subjective evaluations are involved in the foregoing types of interventional studies, results may be complicated by the "Hawthorne effect."[26] This is the phenomenon whereby the process of measurement itself induces changes within the group. For example, during baseline evaluation, an individual who is questioned concerning seat belt use may alter his pattern of use because of the interrogation process itself rather than as a result of any specific intervention. To overcome distortions due to the "Hawthorne effect," Solomon[27] proposes a four-

group design. With this method, in addition to the groups involved in the intervention vs nonintervention designs, two additional groups receive placebo or the active intervention. Baseline measurements are eliminated in these groups. Comparison is then made between the several groups to assess and incorporate effects of the initial measurement into the analyses of final results.

Crossover

This design studies groups which alternately serve control and experimental functions. The intervention, often a medication, literally crosses over from one group to the other at specified intervals. Most frequently the control is a placebo. This design is particularly useful for therapeutic trials in patients with chronic disorders which are subject to exacerbation and remission such as depression or rheumatoid arthritis. Care should be taken in the initial design of such a study that time intervals between crossovers exceed the duration of action of the pharmacologic agent.

Reliability and Validity

The true measure of the quality of research design is ultimately established by the reliability and validity of its results. Reliability is a measure of the accuracy and reproducibility of the observation. Validity relates to the extent that results are indicative of the "real" state; in other words, do the results, no matter how reproducible, actually measure that which they were intended to measure?

Reliability may be affected by any of several factors during the conduct of an investigation. Some of these are:

Subject reliability

The extent to which each subject proffers consistent responses is an important determinant of reliability. Vacillating verbal responses to questions or fluctuating laboratory determinations compromise study reliability. Patient compliance (or noncompliance) may seriously affect reliability of results.

Observer reliability

There are two measures of observer reliability. One relates to concordance of opinion or results among several observers of a specified phenomenon. The second is concerned with consistency of determinations by a single observer of the same occurrence over a period of time.

Situation reliability

The physical conditions of the experimental setting may be variable. For example, a study of depression may show varying results depending upon season of the year. Other situations may exist which grossly alter study outcome, such as disease epidemics or sporadic fads in health care or self-treatment.

Instrument reliability

Any mechanism which provides information is considered to be an instrument. Most laboratory determinations are relatively reproducible and have well-defined limits of variability. As an informational instrument, the questionnaire frequently exhibits considerable variation due, in part, to the ambiguity of questions, and to its compounding influence on subject reliability.

Processing reliability

Data processing is subject to errors of judgment and management. Even automated data processing may be rendered less reliable by errors in coding, keypunching, or programming.

Validity

Validity is the final measure of research endeavor. It is quite apparent that unreliable data will not produce valid conclusions. Not so obvious, but equally important, is the fact that reliable data do not necessarily lead to valid conclusions. There are three basic indices of validity; all are mutually supportive, yet none is definitive.

The first is internal consistency. Repetition of the same study leading to the same results may increase confidence in its validity slightly, but is more reflective of reliability. On the other hand, evaluation of the same hypothesis by another ap-

proach employing other parameters and different protocol but leading to the same conclusions strengthens validity.

External validity is another index of how accurately the results of an investigation reflect a "true" state. This test of validity is one in which criteria other than those employed for the initial investigation are examined as they relate to the initial findings. For instance, if a certain medication has been established in an initial study to consistently reduce blood pressure, external validity of that finding would be enhanced by the observations that patients treated with that medication exhibited significantly lower incidence of stroke or congestive heart failure.

The third, and possibly strongest, index of validity concerns the predictive value of study observations as well as their general applicability. Extrapolation of results from a single study in one setting to a general population is the most difficult index to measure and frequently impossible to accomplish. The large number of variables, often with independent effects, encountered in any medical setting and particularly in family practice frequently preclude the comfort of this measure of validity.

Evaluation of Resources

The precise characteristics of the problems under consideration will, to a large extent, determine the nature of research design. There are other considerations, however, which must enter into a realistic decision, not only of the specific design to be selected, but of the scope and extent of the investigation to be undertaken. The first resource for research endeavor to be considered is usually the patient population, but other sources of information such as patient records, availability of laboratory or x-ray facilities, and access to additional patients via cooperation with other investigators must be considered as well. Evaluation of personnel, space, and financial support will also enter into the rational planning of research effort.

Research in family medicine is frequently designed to answer questions concerning the available patient population. These populations will generally number between 2,000 and 4,000, numbers sufficient to measure variables common to a large proportion of this group. Studies of patient satisfaction, health care utilization, compliance with medical instruction, or other general parameters of health care delivery may be satisfactorily accomplished in populations of these dimensions. Overall morbidity studies and those relating to specific disease entities will, of necessity, be limited to the more common health problems such as upper respiratory tract infections, cystitis, low back pain, anxiety, marital problems, and the like. Access to the entire family allows investigation of the impact of family configuration on such common disease entities and morbidity patterns. Investigations of infrequently occurring disorders, on the other hand, must enlist the collaborative contributions of other practices. Few, if any, family practices contain sufficient numbers to examine the natural history of herpes zoster, for example, or the incidence of pleural effusion.

Researchers associated with a group which maintains indices describing demographic variables of their patient population will enjoy greater ease of entree into individual and collaborative research activity than those who have no such association. The age/sex register,[28] diagnostic index,[11] family information files,[29] and socioeconomic registers[30] are examples of these indices. These registers permit initial selection of subjects for a particular study (eg, females 25 to 44 years old with abdominal pain) and expedite selection of matched groups of subjects. They are also instrumental in the extrapolation of results obtained from a study population to the practice population.

The ambulatory medical record may be a rich source of information on common health problems, but must be approached with caution. In addition to errors introduced by questions of observer reliability which cannot be assessed in the usual retrospective investigation, problems of legibility and organization may further compromise collection of appropriate data. Dictation and subsequent typing of the medical record may improve legibility, but concurrently introduces another potential source of error, that of transcription. The problem-oriented technique of record organization[31] enhances the availability of data for extraction and permits assessment of that health care provider's level of rational thought concerning the patient's health problem and the care rendered.

Computer services can be helpful to any researcher and to the family medicine researcher as

well, although it should be emphasized that excellent research has been accomplished without automation, and lack of access to such services should be no deterrent to research activity. The practice which routinely enters patient information into computer storage is at an obvious advantage; however, information collated by chart review may be entered as a special task and data analysis may proceed with greater ease than manual analysis.

In addition to the previously stated tasks of the researcher, it is essential that he/she familiarize him/herself with previous work performed in that area proposed for investigation. A careful review of the literature, admittedly tedious and time consuming, may be made somewhat more palatable by the recent development of indexing services. Computer searches are available which, if properly used, can provide lists of appropriate references for further pursuit. In establishing a computer search, great care must be taken in generation of key words or phrases since these form the search basis. The investigator must be adequately specific to avoid receiving overwhelmingly lengthy lists, while eschewing such finite specificity that pertinent material may be missed. Once the desired material is listed, Interlibrary Loan Services in the United States make special periodicals and books available to smaller libraries. Computer searches in this country are dependent upon listing in the *Index Medicus*. It is an unfortunate fact that, as yet, some periodicals which publish family medicine research work are not thus indexed. One source of both indexed and nonindexed references is the Canadian Library of Family Medicine. It will conduct literature searches for a fee of $5.00 per hour (a nominal sum). In addition, this library has compiled bibliographies on over 300 topics in family medicine which are available for copying and mailing costs. Both services are available to nonmembers of the College of Family Physicians of Canada. Whatever the source of materials, it is left to the researcher to become thoroughly familiar with their content, and to evaluate his project in the light of prior work within the field.

That research activity is both time consuming and costly is a fact admitted by all who have been involved in its conduct. The practicing physician-researcher must determine early in the research process just who will be involved and how it will be financed. Staff obligated to patient care activities may have neither the time nor the skills required, and it may be appropriate to employ additional personnel. In some circumstances, relatively unsophisticated workers such as high school or college students may be taught, for example, the specific skills involved in collecting data; in others, an interested office nurse can make outstanding contributions to a research effort. In any case, there will be associated costs even if these costs involve only the physician's or other personnel's time. Many physician-researchers finance their activity from practice income. Others may contract with a drug manufacturer to conduct clinical trials of pharmacologic agents. Lastly, research grants are available from numerous sources including diverse branches of the Federal government, private foundations, and service agencies. Physicians should understand, however, that grants from these latter sources require extensive documentation and involve elaborate application procedures, and that competition for funds is intense, creating a relatively formidable barrier.

Finally, research design almost invariably benefits from critical review prior to initiation of a project. Ideally such review should be undertaken by uninvolved parties from other disciplines or by other family medicine researchers. Frequently epidemiologists, sociologists, psychologists, and biostatisticians provide valuable insight concerning adequacy of design and potential sources of bias. Involvement prior to project initiation is particularly important for statisticians, whose services benefit from thorough knowledge of the quality of information to be evaluated, the research setting and patient population, and the overall intent and impact of the study under consideration. By fullest possible understanding of the project's scope, the biostatistician can best anticipate sources of error and assess the several variables required to determine sample size and significance of results.

The Design Process (An Example from Family Medicine Research)

This description of the process of design is taken from a recently completed study conducted

within the Rochester Family Medicine Program. Generically, it is a retrospective, descriptive study of otitis media using patient records as its data base. Each step will be described sequentially as it was performed. Although the limitations of this type of study have been described (see the preceding section, Retrospective vs Prospective), it is particularly attractive for family medicine research since it can be accomplished with limited resources.

Step 1. Creative Thinking

This activity frequently begins with chance or serendipitous observation during the course of patient care, reading, or other research activity. In the routine evaluation of patients, it was observed that a significant portion of cases of otitis media (OM) occurred in adults. This observation, coupled with the realization that there is a paucity of information in the medical literature concerning OM in the adult population, prompted further inquiry. In this phase of study activity, it is important to fully and specifically delineate the questions to be answered. For the OM project the following were proposed for study in this ambulatory setting:

a. Is the presentation of OM different in persons over age 15 from those under 15?

b. Does the mode of therapeutic management of OM vary with age of the patient?

c. Is the course of OM different in patients over age 15 from those under 15?

d. Are there certain disorders which occur with greater frequency in association with OM than in the general population? If so, what are they?

e. Are these associated conditions different in patients over age 15 from those under 15?

f. Is the family structure of patients with OM different from the family structure of patients not receiving that diagnosis?

g. Is the family structure of patients with OM different in patients over age 15 from those under 15?

It may be impossible to ultimately answer all questions originally conceived during this early stage of the investigative process; however, it is important to phrase all questions in specific terms, preferably in writing.

Step 2. Feasibility

The next determination for the investigator is that of study feasibility. The first consideration in the OM project was to establish the total number of OM cases within the practice population. The diagnostic index was consulted and it was found that during a one-year period more than 400 cases of OM had been treated. A second problem related to availability of adequate personnel to perform necessary project tasks, in particular, the time consuming process of extraction of data from the medical records of involved patients. Two senior medical students expressed interest in the project and were enlisted to assist in the data extraction process. Although a complete enumeration of required resource personnel need not be accomplished at this point, the most basic should be identified to determine if the proposed study can be adequately brought to completion.

Step 3. Patient Management Guidelines

For studies evaluating health-care-provider behavior in the diagnosis and management of specific disorders, it is often useful to develop disorder-related protocols. When perfected, these protocols provide a vehicle for quality of care audit and management guides for midlevel providers. When associated with a research project, the specific criteria developed for diagnosis can serve to exclude cases from study which fail to meet the established criteria.

In the OM project, all diagnosed cases of OM were included; of these, 87 percent met the minimal diagnostic criteria. There are three possible explanations for the 17 percent of cases where minimal criteria were not fulfilled. These are:

a. the physician failed to record specific observed data items.

b. the minimal diagnostic criteria should be revised.

c. the physician erred in the diagnosis.

The process employed for development of a protocol for OM, in the course of this research effort, began with a group discussion including family physicians and family physicians-in-training who negotiated minimal diagnostic criteria, and suggested laboratory investigations and appropriate management standards. Discussion focused on personal experience rather than

literature reports. The preliminary protocol thus developed was submitted to other practicing family physicians, pediatricians, and otolaryngologists for suggestion. Incorporation of their input resulted in a revised protocol which appears in Appendix 1. The protocol is not intended as a definitive statement but rather as a reflection of one useful approach to the problem of OM. Although new knowledge from this and other research efforts may produce further revision, the process of protocol production helped define more sharply those areas of controversy and resulted in agreement concerning choice of minimum diagnostic criteria, use of decongestants, and choice of antibiotics in young children.

Step 4. Literature Review

The literature review should focus on current information pertaining to the goals and controversies within the scope of the specific research project. A fruitful approach may be to begin with a recent review article relevant to the issue at hand. Since a thoughtful reviewer will have already sifted through the literature and incorporated the most useful information into his review, some of the selection process may have already been accomplished. Care must, however, be taken in selection of the initial review article. It is generally best to draw from established medical journals of general interest, or from the most credible publications within the discipline under which the disorder in question falls. Selective review of the literature may also help define questions and controversies other than those delineated by the present researcher. Thorough familiarity with the literature may also prevent needless repetition of previous, well-designed, definitive studies. For most types of research design, the literature review is best accomplished at an earlier stage, generally immediately following Step 1, Creative Thinking, and question formulation. In the OM study process it was decided to first elicit family physician experience and then consult the literature. The intent was to prevent dampening of creative input which could have been the case with prior presentation of "expert" findings, since there is a paucity of ambulatory-care-based research on OM (as is true of many other areas). Wherever in the design process a literature review is performed, the research project should be re-evaluated and appropriate changes made in light of this newly gathered information.

Step 5. Crystallization of Design

At this point, specific questions to be addressed by the project are restated or revised as judged necessary from insights gained in Steps 1 through 4. Data items needed to answer these revised questions are listed, and "dummy tables" may be designed to anticipate format needs for analysis and display. Essential data items are then incorporated into a collection instrument, the extraction form.

Step 6. The Extraction Form

Development of an extraction form is a critical step in research design involving the retrospective use of medical records. A well-constructed form will facilitate efficient and accurate data transcription, permit data coding, and be accompanied by precise definitions of ambiguous terms.

First, a glossary of definitions for each data item should be produced. Some of the following areas should be considered:

Registration and Encounter Data

These data are usually collected when the patient or family registers with the practice and are maintained for future health related encounters. Often definition and classification of such items as age, race, socioeconomic status, and family structure have received prior agreement from the practice. If the study sample is to be compared with the practice population, identical definitions should be adopted. For example, standard age groups are used in this practice for routine reporting. These are: <1, 1-4, 5-14, 15-24, 25-44, 45-64, and >64. These are more detailed than those suggested in the *Glossary for Primary Care*,[32] but can be compared to the standard groupings. If special age breakdowns are needed, they should either combine the standard groups or divide them in a fash-

ion permitting subsequent regrouping into the original standard groups.

Age breakdown by decades would be inappropriate for the family medicine practice population. In the OM study we were concerned with only two age groups: less than 15 and over 15. By using this scheme, comparison with the total practice population is possible. Similar considerations are important for all demographic variables studied.

Problem Classification

Whenever possible, a standard classification of health problems or diseases should be used. Currently, most family medicine researchers, especially in North America, use the *International Classification of Health Problems in Primary Care* (ICHPPC) produced under the auspices of the World Organization of National Colleges, Academies, and Academic Associations of General Practitioners/Family Physicians and published by the American Hospital Association.[33] Several family medicine training programs initially used the *Royal College of General Practitioners' Classification of Diseases* (RCGP), but have converted to the ICHPPC. A description of the process and problems of conversion with an item by item list of comparable numbers within the two classifications has recently been published.[34]

Temporal Span and Sequence of Events

It is essential to define the time span to be included in the study. A one-year time period is frequently used; however, there are other intervals equally appropriate. When comparisons are made with the total patient population between patients seen over a one-year period with OM, for example, it is imperative that information on the total patient population be obtained from the same time period.

The investigator must also determine if he will count visits, individuals, or episodes of illness in order to choose the correct denominator. For the OM study, episodes were chosen with full awareness that it is often difficult to determine precisely when one episode ends and a new infection begins.

In some studies, a "zero time point" may be used. This is associated with a number of temporally related items such as onset of symptoms, date of first visit, and date of diagnosis. The zero point is important to define unequivocally since from it length of episodes of illness may be calculated. For studies evaluating outcome it may be necessary to derive data from visits which extend beyond the defined time limits of the study.

Other Operational Definitions

Selltiz suggests that:

Operational definitions ought to (1) assign empirical and logical meaning to concepts in an explicit and precise way; and (2) assign meaning to concepts so that the indicators of the concepts relate to the indicators of other concepts in ways that are predicted by theory. In other words, definitions ought to be unambiguous and clear in what they refer to and definitions ought to be constructed so that concepts fit into theories.[35]

Format

For ease of extraction, the form should be so designed that it reflects the sequential location of data within the medical charts. Chart organization within this practice follows this sequence:

1. Registration and demographic information
2. Initial data base (history and family history)
3. Problem list
4. Medication list
5. Screening sheet
6. Record of visits following the problem-oriented format

Each extracted item requires suitable classification. In the case of OM, the item "management" was classified under the following headings:

1. Antibiotics
2. Decongestants
3. Antihistaminics
4. Antihistaminics plus decongestants
5. Analgesics
6. Myringotomy
7. Referral to consultant
8. Referral to hospital
9. None of the above (Specify)
10. Unrecorded

The antibiotic section (item 1) was subclassified as follows:

a. Penicillin
b. Ampicillin or amoxicillin
c. Erythromycin
d. Gantrisin
e. Penicillin plus gantrisin
f. Tetracycline
g. None of the above (Specify)
h. Unrecorded

Classification facilitates handling of large numbers of discrete and disparate data items. For other items of information it is preferable to record specific values for such items as weight, body temperature, or blood glucose. With numeric data such as this, calculation of means and standard deviations may be made.

Coding

Codification of all information must be accomplished for studies requiring computer analyses. Data processing personnel should be consulted regarding this phase of extraction-form design.

Step 7. The Pilot Study

A trial (pilot study) of the extraction form on a small, usually five to ten, number of charts will generally be adequate to determine the suitability of the form and will frequently highlight problems of definition. Actual experience with the length of time required to locate and extract information from each chart will allow estimation of the magnitude of the study. Not infrequently, a pilot study will indicate redesign of the extraction form and will further clarify the feasibility of accomplishing study aims.

Step 8. Review

At this point, the hypotheses, "dummy tables," extraction form, and general research design should be carefully reviewed. Ideally, consultation would be obtained from a "Research Committee" or other group of research-oriented individuals from family medicine and other related disciplines. The critical decision of whether to proceed or to abandon any study should be made at this point. Since all preceding steps consume only a small portion of the total resources essential for completion of a project, omission of this review step could be, at the least, wasteful. Should any of the foregoing steps contain inaccurately conceived material, considerable numbers of hours could be devoted to the production of worthless, unreproducible, or invalid data. Although all carefully planned studies may not be successful, poor planning invariably results in questionable data and conclusions.

Summary and Conclusions

With its recent emergence as a specific medical discipline, and with increasing definition of its scope and content, research in family medicine presents an exciting and intellectual challenge to its participants. The unique characteristics of the patient population permit investigation of factors in health and disease not ordinarily amenable to research within other disciplines. By participation in the ongoing care of all members of a family, direct assessment may be made of the impact of the family as a functioning unit upon the well-being of its individual members. Much remains to be learned concerning health-related problems as they present for diagnosis and management in the ambulatory setting.

It is only through the application of established principles of research methods and design within the family practice setting that new insights may be gained. This discussion of research design, by no means comprehensive, is intended to give an overview of those methods of research most likely to evoke meaningful answers to some of the questions encountered in the family practice setting. Any single project may employ one or a combination of the design types illustrated here.

Although this paper has focused on technical aspects of investigational procedure, the critical element of all research activity lies with the creative input of the investigator. In the words of P. L. Berger (1969),

> *In science as in love, a concentration on technique is quite likely to lead to impotence.*

Appendix
Protocol for the Diagnosis and Management of Otitis Media

A. Acute Otitis Media
   1. Minimal diagnostic criteria—one or more of the following:
      a. Redness and bulging of tympanic membrane (in the absence of bullous myringitis)
      b. Redness and abnormal mobility of the tympanic membrane
      c. A difference between the two tympanic membranes in redness or mobility
      d. Redness of the tympanic membrane, plus one or more of the following: pain, fever, loss of landmarks
   2. Laboratory work
      a. If patient with fever is less than six weeks of age, do the following:
         i. Hospitalization, lumbar puncture, blood culture, chest x-ray, tympano centesis (optional), urine culture
      b. If patient with fever is six weeks to three months of age, do the following:
         i. Hospitalization is optional, but patient requires lumbar puncture, blood culture, and chest x-ray (particularly if child appears toxic)
      c. If patient is over three months of age:
         No laboratory work is required
   3. Treatment
      a. Antibiotics
         i. If patient is less than six weeks of age, hospitalize. Antibiotics will depend upon current susceptibility of gram negative organisms
         ii. If patient is six weeks to three months of age:
            a. Amoxicillin 30 mg/kg per day
            b. Ampicillin 50 to 75 mg/kg per day
            c. If vomiting, 600,000 units of Bicillin C-R plus oral Gantrisin (if more than two months of age, 150 mg/kg per day in divided doses) or intramuscular ampicillin 125 to 250 mg, then either oral amoxicillin or ampicillin
         iii. If patient is over three months of age: any of the following:
            a. 1,200,000 units Bicillin C-R if greater than 30 kg; 600,000 Bicillin C-R, if less than 30 kg
            b. Amoxicillin 30 mg/kg per day; ampicillin 50 to 75 mg/kg per day
            c. Oral penicillin 1/2 gm/day if less than 30 kg; 1 gm/day if greater than 30 kg
            d. Option (a) or (c), plus Gantrisin 150 mg/kg in divided doses per day
            e. Erthromycin and Gantrisin particularly useful in the penicillin-allergic patient
      b. Decongestants and/or antihistamines: perhaps useful if the child is allergic. There is no evidence of their value in the treatment of otitis media.
      c. Follow-up: two weeks. Check movement and morphology of tympanic membrane. Ask about and/or check hearing. If hearing is impaired at this visit, reschedule visits monthly until hearing returns to normal.

B. Recurrent otitis media
   1. Definition: three or more infections in one year
   2. Treatment: oral Gantrisin 500 mg twice a day (only 'proven' benefit is in males under 6 years old)

C. Prevention
   1. Educate mother to feed baby in upright position.
   2. All children with cleft palate need polyethylene (PE) tubes as soon as feasible (usually about 3 to 5 months of age).

D. When to refer
   1. Incomplete resolution of acute otitis media with appropriate therapy manifested by:
      a. Persistent decreased (>25 decibels) hearing at three months after onset of infection, or severe hearing loss at one month
      b. Persistence of middle ear fluid at three months or hemotympanum at any time
      c. Atelectatic tympanic membrane
      d. For recurrent otitis media which may benefit by the placement of PE tubes
      e. At parents' request
      f. For allergic work-up if this appears to be a prominent feature of the illness

E. Place of tonsillectomy and/or adenoidectomy: probably none. Occasionally, adenoidectomy is useful at the time of PE tube insertion.

F. Use of myringotomy: rarely required.
   Only proven benefit is for acute relief of pain.

References

1. Pearson K: The Grammar of Science. 1911 Edition, reissued by the Meridian Library, New York, 1957
2. Marsland DW, Wood M, Mayo F: The content of family practice. J Fam Pract 3:37, 1976
3. Froom J, Rozzi C, Metcalfe DH: Research comments: Computer analysis of morbidity reports in primary care (Implications for a national morbidity study). J Clin Comput 2:42, 1973
4. Dixon AS: Survey of a rural practice: Rainy River. Can Fam Physician 21:693, 1975
5. Morbidity statistics from general practice: Second National Morbidity Survey 1970-1971. In Office of Population Censuses and Surveys: Studies on Medical and Population Subjects, No. 26. London, Her Majesty's Stationery Office, 1974, p 36
6. Rickels K, Hesbacker P, Fisher EL, et al: The private practice research model from psychopharmacology for clinical research in family practice. J Clin Pharmacol 17:541, 1977
7. Suchman EA: The principles of research design. In Doby JT, Suchman EA, McKinney JC, et al: An Introduction to Social Research. Harrisburg, Penn, The Stackpole Co, 1954
8. Sartwell PE, Masi AT, Arthes FG, et al: Thromboembolism and oral contraceptives: An epidemiologic case-control study. Am J Epidemiol 90:365, 1969
9. Royal College of General Practitioners: Oral Contraceptives and Health. London, Pitman Medical, 1974
10. Cochran WG: Research techniques in the study of human beings. In McKinlay JB: Research Methods in Health Care. New York, Prodist, 1973
11. Froom J, Culpepper L, Boisseau V: An integrated medical record and data system for primary care: Part 3: The diagnostic index: Manual and computer methods and applications. J Fam Pract 5:113, 1977
12. Henk M, Froom J: Outreach by primary care physicians. JAMA 233:256, 1975
13. Froom J, Treat DF, Farley ES, et al: Curriculum for family medicine. NY State J Med 74:1551, 1974
14. Froom J: Lead screening by family physicians. J Fam Pract 4:631, 1977
15. Brook RH, Davies-Avery A, Greenfield S, et al: Assessing the quality of medical care using outcome measures: An overview of the method. Med Care 15(Suppl to No. 9):1-165, 1977
16. Froom J, Mold J, Culpepper L, et al: The spectrum of otitis media in family medicine. Presented at the Eighth World Meeting of the World Organization of National Colleges, Academies, and Academic Associations of General Practitioners/Family Physicians, Montreux, Switzerland, May 14-19, 1978
17. National ambulatory medical care survey: 1973 Summary: United States, May 1973-April 1974. In National Center for Health Statistics (Rockville, Md): Vital and Health Statistics, series 13, No. 21. DHEW publication No. (HRA) 76-1772. Government Printing Office, 1975
18. Doll R, Hill AB: Smoking and carcinoma of the lung: Preliminary report. Br Med J 2:739, 1950
19. Feinstein AR: Clinical biostatistics: Part 20: The epidemiologic trohoc, the ablative risk ratio, and 'retrospective' research. Clin Pharmacol Ther 14:291, 1973
20. Friedman GD, Kannel WB, Dawber TR, et al: An evaluation of followup methods in the Framingham heart study. Am J Public Health 57:1015, 1967
21. Dawber TR, Kannel WB, Lyell LP: An approach to longitudinal studies in a community: The Framingham study. Ann NY Acad Sci 107:539, 1963
22. Kannel WB, Castelli WP, McNamara PM: The coronary profile: 12-year followup in the Framingham study. J Occup Med 9:611, 1967
23. Froom J: Prevalence and natural history of chronic diseases in childhood. In Grave GD, Pless IB (eds): Chronic Childhood Illness: Assessment of Outcome, vol 3 of the National Institutes of Health (Bethesda, Md): Fogarty International Center Series on Teaching of Preventive Medicine. DHEW publication No. (NIH) 76-877. Government Printing Office, 1974
24. Teitelbaum SL, Bone JM, Stein PM, et al: Calcifediol in chronic renal osteodystrophy: Skeletal response. JAMA 235:164, 1976
25. Hulley SB, Cohen R, Widdowson G: Plasma high-density lipoprotein cholesterol level: Influence of risk factor intervention. JAMA 238:2269, 1977
26. Suchman EA: Evaluative Research. New York, Russell Sage Foundation, 1967
27. Solomon RL: An extension of control group design. Psychol Bull 46:137, 1949
28. Froom J: An integrated medical record and data system for primary care: Part 1: The age-sex register: Definition of patient population. J Fam Pract 4:951, 1977
29. Froom J, Culpepper L, Kirkwood R, et al: Family information. J Fam Pract 5:265, 1977
30. Farley ES, Boisseau V, Froom J: Implications of filing family folders by area of residency. J Fam Pract 5:427, 1977
31. Weed LL: Medical Records, Medical Education and Patient Care. Cleveland, Ohio, The Press of Case Western Reserve University, 1969
32. A glossary for primary care. Report of the North American Primary Care Research Group (NAPCRG). J Fam Pract 5:633, 1977
33. International Classification of Health Problems in Primary Care. Report of the Classification Committee of the World Organization of National Colleges, Academies, and Academic Associations of General Practitioners/Family Physicians. Chicago, American Hospital Association, 1975
34. Schneeweiss R, Stuart HW Jr, Froom J, et al: A conversion code from the RCGP to the ICHPPC classification system. J Fam Pract 5:415, 1977
35. Selltiz C, Wrightsman LS, Cook SW: Research Methods in Social Relations. New York, Holt, Reinhart, and Winston, 1976

Suggested Reading

Blalock HM: Causal Inferences in Nonexperimental Research. Chapel Hill, NC, University of North Carolina Press, 1961

Cutler SJ, Greenhouse SW, Cornfield J, et al: The role of hypothesis testing in clinical trials. J Chronic Dis 19:857, 1966

Kilpatrick SJ Jr: Statistical Principles in Health Care Information. Baltimore, Md, University Park Press, 1977

Snedecor GW, Cochran WG: Statistical Methods. Ames, Iowa, Iowa State University Press, 1967

Susser MW: Causal Thinking in the Health Sciences: Concepts and Strategies of Epidemiology. New York, Oxford University Press, 1973