Evaluating the Medical Literature Part I: Abstract, Introduction, Methods

[Cuddy PG, Elenbaas RM, Elenbaas JK: Evaluating the medical literature. Part I: Abstract, introduction, methods. Ann Emerg Med 12:549-555, September 1983.]

Paul G. Cuddy, PharmD*†
Robert M. Elenbaas, PharmD*‡
Julia K. Elenbaas, PharmD*
Kansas City, Missouri

From the Schools of Pharmacy and Medicine, University of Missouri-Kansas City;* and the Departments of Medicine† and Emergency Health Services,‡ Truman Medical Center, Kansas City.

Received for publication December 23, 1982. Revision received April 6, 1983. Accepted for publication June 29, 1983.

Address for reprints: Robert M. Elenbaas, PharmD, Department of Emergency Health Services, Truman Medical Center, 2301 Holmes Street, Kansas City, Missouri 64108.

The ability to critically evaluate medical literature is becoming more important as an increasing number of articles describing innovations in health care are published.[1] Only after a critical review can the clinician derive valid and useful conclusions from an article and appropriately incorporate its findings into his practice. While the majority of medical investigators do not intentionally attempt to deceive, they are subject to influence by their biases or convictions. Occasionally, simple mistakes will be made. While the editorial review process should prevent very poorly conducted studies from being published, the system is not a perfect screen. Thus the reader must not accept the written word on face value, but must form his own conclusions after a careful, critical review.

Our purpose in this three-part series is to review the steps used to critically evaluate the medical literature. This will be accomplished by dissecting representative sections of typical research publications (Figure 1). Each section will be addressed by describing its contribution to the research article, identifying information that should appear, and emphasizing common pitfalls. Examples taken from published or fictional studies will be used to clarify central concepts. While most examples will use studies involving the evaluation of drugs, this simply reflects the authors' areas of professional expertise. The concepts presented are applicable to any study.

Part 1 of this series will address the Abstract, Introduction, and Methods sections of research articles. Part 2 (*Annals*, October 1983) will provide an overview of the selection and interpretation of statistical tests. The concluding segment (*Annals*, November 1983) will address the Results and Discussion sections. The series follows the same format as a typical research publication, which will be helpful in allowing the reader to systematically evaluate the medical literature.

There is probably no such thing as a perfect study. On critical examination all will likely be found to have some flaws. The reader must decide the degree to which the study conclusions are validated, and must form his own conclusion.

THE ABSTRACT

The purpose of the Abstract is to provide an abbreviated summary of the article. Although generally restricted in length (eg, 250 words),[2,3] a complete abstract should include information that identifies the study purpose, research design, methodology, and results; it may contain a brief statement of conclusions or recommendations. The good abstract is informative, and should be able to stand apart from the article.

Abstracts frequently are prepared by the author, and thus may reflect his biases. While it is possible that the author's abstract may be misleading or inaccurate, it is a given that study methods cannot be sufficiently explained due to space limitations. A critical examination of the study methods is crucial to any review. Thus after reading only the abstract, one cannot form a

critical opinion of the study's validity. Yet another potential limitation of reading the abstract before the remainder of the paper is that the reader may become biased toward the study if its conclusions are in agreement or at variance with his own beliefs. Despite these limitations, the abstract serves a useful and time-saving function.

THE INTRODUCTION

The purpose of the Introduction is to acquaint the reader with the problem under study and convey the reasons for conducting the investigation. Background information should also be presented. The author may review the historical developments which have led to the study. This should enable the reader to understand the problem and evaluate the results of the study without referring to earlier publications.

The most important portion of the Introduction is identification of the study objective. Typically this is located in the last paragraph of the Introduction. The study objective identifies the specific questions to be answered. Subsequent examination of the Methods, Results, and Discussion is done with the study objective in mind; the Methods must be designed to answer the question, and the author's conclusions should not extend beyond the stated objective.

A common flaw in medical research publications is failure to present a clearly defined objective. The recent publication by Albert et al[4] contains a good illustration of a clearly stated study objective: "The following experiment was done to ascertain whether corticosteroids improve results of spirometry or arterial blood gas levels, or both, in patients with chronic obstructive pulmonary disease and acute respiratory failure." In addition to identifying the objective, the authors have defined the study population and specific methods for evaluation of efficacy. An example of a poorly stated objective for the same publication might read as follows: "We studied the use of corticosteroids in patients with lung disease." In this instance one does not know the target population, what the investigators are attempting to study, or how they will measure success.

Reviewing an article that does not contain a clearly stated study objective can be very difficult because the reader has no map of the author's intentions or yardstick to measure the study's methods or conclusions. Identifying the study objective is the most important thing to accomplish when reading the Introduction, for analysis of the balance of the paper will determine whether the investigators have answered the question they set out to answer.

THE METHODS

The Methods section is the single most important section of any research publication. It is unfortunate that many clinicians omit the Methods from their review, for the nuts and bolts of the study can only be found here.[5]

An investigator is usually trying to prove a cause and effect relationship between two items when conducting a study. He should attempt to prove this by controlling as many extraneous, confounding variables as possible, so that the agent tested is the only variable that could have accounted for the results. A study does not provide direct proof that, for example, the drug being investigated actually caused the effect being observed. Instead, if one has ruled out an inherent (baseline) difference between the study groups, a difference in handling or evaluating the study groups, and the likelihood that the results were due to chance, it is reasonable to assume the drug was responsible. While the study methods may not always make for exciting or rapid reading, only by reviewing this segment can one assess the validity of the study.

The Methods should be presented in sufficient clarity and detail so that the reader could reproduce the investigator's work if desired. A typical Methods section will include information concerning experimental design, study sample, treatment allocation, index of accomplishment, and statistical test selection.

Experimental Design

In general, the research design of medical articles will be of two types: cross-sectional or longitudinal (Figure 2).[6,7] *Cross-sectional* studies involve observations made at one point in time (for example, the prevalence of diabetes mellitus in the United States on the date this article was published). They usually have an epidemiologic purpose. The results of a cross-sectional study may confirm that two factors are likely to be present together, and thus suggest a relationship. The disadvantage of the cross-sectional study is that the importance of the relationship cannot be determined; a cause and effect relationship cannot be established. The advantage, however, is that they can be done simply and economically.

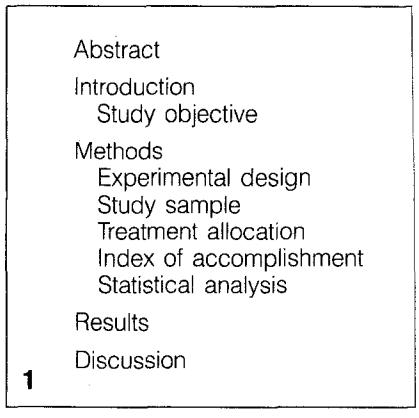

Fig. 1. *Sections of the research paper.*

Longitudinal studies, on the other hand, involve observations over a prolonged period. This type of study can be further subdivided into two major designs: retrospective and prospective. *Retrospective* studies, sometimes called "trohoc" studies, are studies in which subjects are examined for some common factor in their history (eg, from recognition of a disease to identification of its suspected cause).[7] If the retrospective study also attempts to identify the suspected etiologic factors in patients who do not have the disease in question (controls), it is a *case-control* study. Unfortunately retrospective studies generally rely on medical records or subject recall. For this reason, it is quite difficult to verify the existence of a factor or other condition to the degree that can be done with prospective designs. Likewise, it is more difficult to establish the comparability of cases and controls in the retrospective study and it is impossible to intervene in the patient's care. Because of these deficiencies, it is difficult to establish a cause and effect relationship with the retrospective study design.

The retrospective design does, however, serve a useful purpose: it is generally well suited to the study of rare diseases or conditions. Retrospective studies may have epidemiologic importance or serve to identify problems for subsequent prospective evaluation. They commonly involve more patients, are less expensive to perform, and can be conducted in a

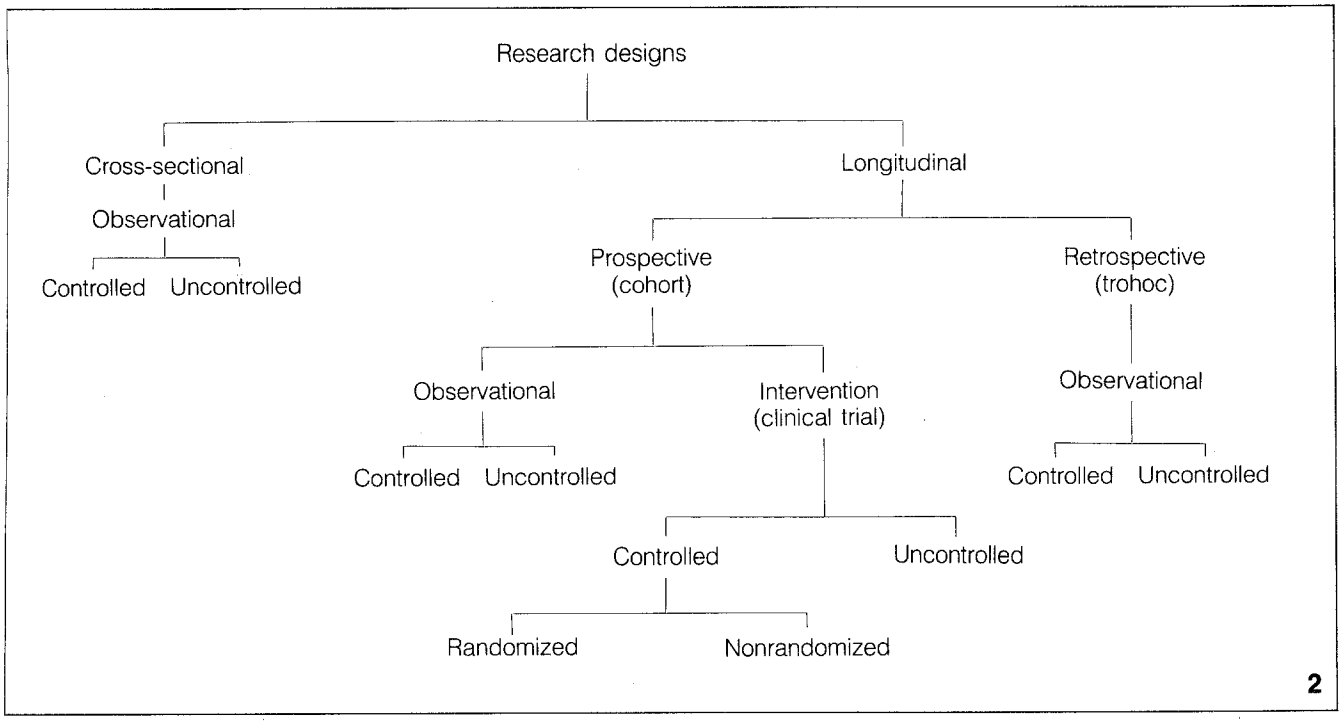

Fig. 2. *Basic research designs.*

shorter time interval than can prospective studies.

Prospective studies, also called "cohort" studies, utilize a design in which subjects are followed forward in time from identification of the characteristic (eg, etiologic agent) to outcome (eg, disease).[7] In many instances, a prospective design will be used to simply observe the effects of a treatment or event which ordinarily would have taken place. Typically, however, the prospective design involves observation of a group of subjects who have the risk factor or trait in question and a group comparable in all respects except for the absence of the risk factor or characteristic (controls). One such example could be the assessment of the long-term outcome of patients with out-of-hospital cardiopulmonary arrest with and without bystander-initiated CPR.

Prospective studies are generally more expensive to conduct than are retrospective studies, and may involve very long periods of follow up. The advantages of prospective design include the ability of the investigator to verify comparability of subjects prior to beginning the study; establish more rigid criteria for the existence of the trait, disease, or condition before the study begins; and execute closer follow up of subjects. The prospective study design more effectively establishes the cause and effect relationship because the investigator can verify that the characteristic preceded its alleged effect or condition.

The *prospective clinical trial* is a study design commonly encountered in medical research publications.[8] The clinical trial is a prospective study in which the investigator intervenes in the patient's care specifically for the purposes of evaluation, and not as part of the routine care plan. Clinical trials can either be controlled (in which a comparable group of subjects is also studied but without receiving the treatment intervention) or uncontrolled (in which no comparison group is involved). In a controlled clinical trial, the method of treatment allocation can be either random or nonrandom.

The principal difference between the clinical trial and other designs is the ability of the investigator to actively intervene, rather than simply to observe. Major limitations include its expense, the ethical problems inherent in testing new therapies in human beings or creating control groups, the time necessary to perform the study, and complexity. The prospective, randomized, controlled clinical trial, however, is probably the most reliable technique for evaluating treatments, and ultimately serves to resolve many therapeutic controversies.

Study Sample

The reviewer must be able to precisely visualize the sample under study. This requires that the author clearly identify both the entry and exclusion criteria of the study. Generally *entry criteria* serve two purposes: they describe the population the patients represent, and they enable the reader to determine whether the study sample resembles his clinical practice sufficiently well so that study findings can be extrapolated to it. *Exclusion criteria* serve three major functions: they help to ensure that the study sample is as homogeneous as possible, they specifically identify patient subsets to which study results should not be extrapolated, and they assure patient safety in the study by excluding individuals for whom some aspect of participation would be contraindicated or dangerous.

The study by Albert et al[4] will be used to demonstrate the precise, clear statement of entry and exclusion criteria. Within Methods is the following statement of entry criteria: "We accepted all hospitalized male and female patients with acute respiratory insufficiency, acute and chronic bronchitis, and chronic airflow obstruc-

tion." If no additional definition of these classifications were proffered, we would view a classic illustration of vague entry criteria. Fortunately, however, the authors have provided an explicit definition for each term: "Acute respiratory insufficiency was defined by an arterial PO_2 of 65 mm Hg or lower on room air (after correcting for hyperventilation assuming a respiratory quotient [R] of 0.08) or an arterial PCO_2 of 50 mm Hg or higher accompanied by a pH of 7.35 or lower, or both." Comprehensive descriptions are also provided for acute bronchitis, chronic bronchitis, and chronic airflow obstruction. Having read their entry criteria, it is possible to define precisely the study sample and separate these individuals from those with other forms of lung disease.

These authors also provided a complete listing of exclusion criteria. As one example of these, we find excluded from the study patients with a "personal or family history of asthma." Once again this vague statement could be misleading if further definition were not provided. Fortunately, however, this is defined as "episodic wheezing or dyspnea that rapidly reversed with treatment and was separated by asymptomatic periods." We now know that whatever the study's conclusions regarding the efficacy of corticosteroids, they apply only to patients with COPD meeting specified criteria (entry criteria) and not to individuals with asthma (exclusion criteria).

Other common exclusion criteria for research protocols include concomitant drug therapy, presence of contraindications to study therapy, sex, age, and pregnancy. Particularly pertinent to the placebo-controlled trial are exclusion criteria that exclude patients in whom placebo therapy might be dangerous. For example, the efficacy of prophylactic oxacillin in dog bite wounds was investigated.[9] Patients with orthopedic injuries (eg, open fracture) were specifically excluded from the study because they would likely have been put in inappropriate jeopardy if randomly placed in the placebo group.

Entry and exclusion criteria help both to define the patient sample and to eliminate complicating or hazardous situations. Only when these criteria are present and explicitly defined will one be able to visualize the study sample and extrapolate the investigator's results to one's own practice.

Withdrawal of patients from a prospective investigation represents one additional factor which may markedly influence the eventual study sample.[10] In contrast to exclusion criteria, which prevent a patient from participating in the investigation, withdrawals represent those patients who were initially deemed study candidates only to be withdrawn at a later date. There are three major reasons for patient withdrawals: poor patient compliance with the study protocol prohibiting accurate data collection, data that are uninterpretable (eg, poor-quality roentgenogram), and eventual identification of factors that describe ineligible subjects (eg, the patient develops a new disease or problem contraindicating participation).

Treatment Allocation: Sample Types

Clinical studies can contain an essentially endless number of groups. We will address the simplest case of two treatment groups. Studies with two treatment groups can be divided into those with paired and those with independent samples. *Paired samples* exist when each subject in one group has only one matching mate in the other group. Examples of paired samples include the case of identical twins, left versus right eye, different patients specifically matched for key characteristics, or the same individual studied under both treatment conditions.

When the same subject can serve as both members of the pair, a special case exists. The "cross-over" design in clinical trials makes use of this special type of paired sample in which the patient receives first one treatment and then the other. In this instance, it is important to consider the potential for treatment carry over, that is, the effects of the first treatment persist into the second treatment period, making measurement of the second agent's true effect difficult. If possible, a "washout period" should be a part of the design between the two treatments to eliminate this potential interference. The length of this washout period will vary depending on the duration of action of the agents under investigation.

Another special concern in the cross-over study is to avoid introducing a systematic bias by placing all the patients on one treatment and then giving them the other. For example, suppose an investigator compared nifedipine and propranolol in the chronic management of angina and all the patients received propranolol first followed by nifedipine. A systematic bias that may invalidate the study results has been introduced. Natural progression of the disease, changes in patient behavior, or physical conditioning may serve to modify the perceived effect of the second drug. To avoid such bias, half the patients should first receive nifedipine, while the other half first receive propranolol. Allocation to the two groups should be done randomly. While this cross-over design is the strongest method available for a clinical trial, not all problems can be studied in this way. Cross-over design is especially valuable for studying chronic, stable diseases.

Independent samples differ from paired samples because they include different subjects that are not matched as pairs for key characteristics, and are also often of different size. Generally larger numbers of subjects will be required to conduct a study using the independent, in contrast to the paired, sample design.

Ideally the two samples, when taken as a whole, will be similar in all important characteristics except exposure to the study maneuver. This is not always possible, however, and it is this variability that reduces the investigator's ability to attribute observed response to the treatment intervention. The reader will want to ascertain that all treatment groups are equivalent with respect to key baseline or concurrent variables when independent samples are used.

Treatment Allocation: Randomization

Random assignment in clinical trials involves a chance procedure by which patients are assigned to receive the modality under investigation.[8] Successful randomization is achieved when every subject has an equal chance of receiving any available intervention. Randomization is important for two reasons: it forms the basis for the validity of many statistical tests used to analyze study results; and it helps to ensure that investigator and patient bias do not influence treatment assignment and, therefore, study results.[11]

An example of a desirable randomization technique would include the use of a table of random numbers. Assigning treatment by day of the week the patient presents, date of birth, medical record number, or other similar schemes is undesirable because each involves a systematic mechanism instead of being truly random. For example, suppose an investigator comparing penicillin and killsallicillin in pneumonococcal pneumonia assigned his patients to the penicillin group if they had an even medical record number and to killsallicillin if the medical record number were odd. Knowing that penicillin has much lower minimum inhibitory concentrations than does killsallicillin, the investigator might hesitate to enroll in the study a particularly ill patient with an odd medical record number because he wants the patient to receive the "best therapy possible." The systematic bias introduced could create treatment groups with different disease severity, and thus invalidate the study's results. If an investigator feels he cannot ethically place a certain type of patient in either study group, the study's exclusion criteria should prevent such a patient from participating at all.

Randomly assigning patients to treatment groups tends to create groups that are comparable in key baseline or concurrent variables when large sample sizes are involved. It is possible, however, that unbalanced treatment groups could be created despite appropriate randomization, especially when treatment groups are relatively small. The cross-over design eliminates this problem, but in studies with independent samples, comprehensive comparative data that allow the evaluator to determine whether the samples are truly comparable in baseline or other important variables should be presented. The ultimate responsibility for ascertaining that randomization occurred adequately rests with the reader.

Assignment of treatment is possible only with the clinical trial. The allocation of patients to study groups is predetermined in retrospective and prospective studies other than the prospective clinical trial. For example, in retrospective studies, those patients possessing an attribute or disease form one group, and those lacking it are the controls. Biased samples may exist in such studies. This can be illustrated in a prospective study comparing surgery to drug therapy in the treatment of cancer: it is likely that the decision to operate or treat with a drug involved the physician's bias.

Treatment Allocation: Control Groups

The careful reader will examine research publications to verify that an investigator has expended reasonable effort to control for factors that may confound the results. The investigator can accomplish this objective by including a control group within the study design. However, not all studies employ control groups.[8] These types of designs are called "uncontrolled" trials. When the investigator includes a control group within the design, the reader should assess its effectiveness: the control and study groups should be similar in all important baseline characteristics and in the procedures performed on them, except for the variable being tested.

Many types of controls can be used in research studies. Historical controls are used occasionally, and might be appropriate when medical ethics preclude inclusion of a "non-treatment" control. In this case, data from the current study are compared to those collected in previously published investigations. For example, it would be inappropriate to withhold cardiopulmonary resuscitation and advanced cardiac life support from patients with cardiopulmonary arrest; however, an investigator may wish to compare outcome of patients with out-of-hospital cardiac arrest managed with and without paramedic-initiated advanced cardiac life support. A historical control of patients treated before development of a sophisticated EMS program could be compared to the current system involving paramedics. However, certain criteria must be satisfied before use of a historical control can be accepted. The distribution of baseline patient characteristics and patient management, other than the study variable itself, should be similar in both groups. General advances in medical care thus usually invalidate the use of historical controls.

A concurrent control group exists in a study in which at least two groups of patients with different interventions are being compared. This is commonly done, for example, when new drugs are compared to standard, "state of the art" drug therapy (the penicillin versus killsallicillin example). In this instance, the reader should verify that both treatments are administered in appropriate dosage and under the same guidelines. Here the reader should also determine whether a fixed dosage regimen or a titrated schedule is indicated; that is, should all patients receive a standard dosage or should the dose be titrated to the measure of some clinical response? Titrated schedules are indicated when significant inter-individual variability exists in patient response. Examples of this situation include such problems as obstructive lung disease, hypertension, or cardiac dysrhythmias.

The third type of control used in clinical trials is the placebo, which can be defined as the effect of a therapeutic intervention that does not have any specific, objective activity for the condition being treated.[8,12] Use of a placebo control group represents one of the most powerful techniques available to the investigator to identify the degree of patient suggestibility, assess the doctor's personal influence on the patient, and control for other sources of potential bias in research studies. The importance of the placebo control is that placebo responders can positively or negatively influence the results of research studies. For this reason the reader should always recognize when inclusion of a placebo is necessary. A placebo group is particularly indicated when studying diseases in which spontaneous variation or remission is common and when measuring therapeutic response is highly subjective. Examples of these situations include pain, rheumatoid arthritis, ulcerative colitis, angina, and peptic ulcer disease.

Occasionally the reader will find clinical trials with a concurrent control group in which no treatment is given. This may occur in the study of conditions for which no treatment was available previously. The absence of treatment might bias the observer, and therefore a placebo should be given in such circumstances. A placebo is not necessary, and is in fact unethical, in studies of treatment of infectious diseases or contraception.

The placebo must be as identical as possible to the study intervention in all aspects.[13] Appearance, route of administration, color, odor, and cost all contribute to creation of the placebo effect. To establish a cause and effect

relationship between the study medication and observed response, one must be satisfied that a placebo response is not being interpreted as an effect of active therapy.

Treatment Allocation: Blinding

Controlled clinical trials should be blinded to prevent the potential biases of investigator and patient from influencing the study results. Trials can be conducted in a single- or double-blind fashion.[8,14] A *single-blind trial* describes a situation in which either the subject or the investigator is unable to identify the assigned treatment. A *double-blind trial*, on the other hand, describes a situation in which neither the investigator nor the subject knows the treatment assignment.

Not all controlled studies must be blinded, although it is usually a desirable characteristic which tends to enhance study validity. The more objective the measurement of treatment response, the less critical blinding becomes. However, even when highly objective measures of response are used, blinding the investigator and/or patient to the treatment allocation will help to eliminate potential sources of bias. For example, suppose that an investigator is comparing subcutaneous epinephrine and terbutaline in acute asthma, and the clinical response is being assessed by measuring FEV_1 (a patient-dependent test). If, for some reason, the investigator believed prior to initiating the study that terbutaline would prove to be the superior agent, he might unconsciously encourage terbutaline patients to perform with maximum effort and accept sub-maximum effort from epinephrine-treated individuals. Blinding the investigator to the treatment would eliminate this potential source of bias (variability).

Having established the need for a control group, there are few rational reasons for not conducting the study in a single- or double-blind fashion. Legitimate exceptions are studies involving surgical procedures or physical manipulation, situations that make evaluation of their efficacy especially difficult.[15]

Thus there are two general indications for the use of blinding techniques. Investigations should be blinded when the measurement of response to treatment is subjective, and when knowledge of the assigned treatment risks the introduction of bias by the subject or investigator. Blinding is also important when ancillary care may influence results. For example, wound infections can be influenced by general wound care, and one would want to be confident that dressing changes were handled similarly in both the antibiotic-treated and control groups.

The reader should always verify that blinding maneuvers have been accomplished successfully when they are used, ie, treatment and control dosage forms should be made to look identical and should be administered using a similar regimen. It should be difficult for the patient or investigator to deduce which treatment is being given. For example, a hypothetical double-blind study could be conducted comparing subcutaneous epinephrine and nebulized terbutaline in acute asthma. Because the drugs would be administered by different routes, blinding could be maintained by giving each patient both a nebulization and an injection: some patients receiving epinephrine subcutaneously and normal saline placebo by nebulization, while others receive normal saline placebo subcutaneously and terbutaline by nebulization following the same administration schedule. This procedure to maintain blinding is sometimes called the "double dummy" technique.

A break in blinding procedures can destroy the study's effectiveness. This would be particularly likely if a characteristic taste, smell, shape, or color of one treatment were not present in the control. Side effects and characteristic laboratory alterations represent other potential leaks in study blinding. It may sometimes be necessary, therefore, to prohibit access to this particular information by the individual assessing patient response. Assuming patient safety can be maintained, this is highly desirable.

Index of Accomplishment

An appropriate index of accomplishment should be selected to assess outcome. The author should specifically state those parameters he intends to measure when assessing response to treatment. Before the index of accomplishment can be judged to be appropriate, the reader must recall the study objective. This can be illustrated by example. A hypothetical study is designed in which an anticonvulsant is administered to acutely treat alcohol withdrawal seizures; the study objective is to determine whether the drug decreases seizure activity. An appropriate index might be the frequency of seizures following anticonvulsant administration. An inappropriate index might involve the use of an EEG to assess efficacy.

The index should always be as precise and reproducible as possible. This implies that it be free of potential subject recall error, instrument error, and evaluator bias. Subject recall error can occur when subjects vary in their ability to recall past events (eg, the number of anginal episodes in the preceding year). Instrument error occurs when the measuring device lacks accuracy and precision (eg, many laboratory tests may lose precision when concentrations are extremely low or very high). Lastly, the investigator may unknowingly observe a measurement in a biased fashion. This occurrence can be avoided by having another individual who is blinded from the treatment allocation assess outcome (eg, in the report by McHardy comparing sucralfate and placebo in duodenal ulcer, healing was assessed by an independent panel using endoscopic photographs[16]).

The last consideration in evaluating an author's index of accomplishment is recognition of the multiple variables that can influence the assessment process. When factors during the assessment process influence the subjects, the final composition of the study sample may differ from the group of subjects prior to treatment. For this reason factors other than the study intervention might contribute to the study outcome. One illustration of this potential pitfall involves long-term antihypertensive drug studies. When subjects return for blood pressure determination the investigator may influence some of the subjects' behavior patterns, eg, smoking, exercise, or weight. If this influence is widespread, the characteristics of the study sample may change from their baseline distribution. In this instance, factors other than the treatment intervention may be responsible for the perceived treatment-induced decrease in pressure, thereby complicating the establishment of the cause and effect relationship.

Another variable that influences the assessment process is the tendency of a study subject to execute the assessment process either more or less effectively as the maneuver is repeated.

Situations in which this is likely to occur include the performance of pulmonary function testing and cardiovascular treadmill testing. The anxiety and fear that may initially surround these tests may decrease with repeated performance of the test. For this reason the outcome can change without reflecting intervention effect. This factor can complicate the establishment of the cause-effect relationship. If at all possible, indices of accomplishment not subject to such influences should be chosen. When no effective alternative exists, the reader must consider this potential source of bias and recognize the importance of a concurrent control group.

SUMMARY

In this installment on critical evaluation of medical literature, we have discussed the Abstract, Introduction, and Methods sections. A great deal of information is placed in the methodology section of any paper. To critically evaluate a research publication, one must evaluate the methodology, for evaluation of the results of an investigation is meaningless if the methodology is unsound.

REFERENCES

1. Durack DT: The weight of medical knowledge. *N Engl J Med* **298**:773-775, 1978.

2. Mattox KL: Scientific communication: The abstract. *JACEP* **6**:465-466, 1977.

3. Pal AC: Writing good abstracts. It's a tough job. *Phys Ther* **58**:175-177, 1978.

4. Albert RK, Martin TR, Lewis SW: Controlled clinical trial of methylprednisolone in patients with chronic bronchitis and acute respiratory insufficiency. *Ann Intern Med* **92**:753-758, 1980.

5. Der Simonian R, Charette LJ, McPeek B, et al: Reporting on methods in clinical trials. *N Engl J Med* **306**:1332-1337, 1982.

6. Fletcher RH, Fletcher SW: Clinical research in general medical journals: A 30-year perspective. *N Engl J Med* **301**:180-183, 1979.

7. Colton T: Medical surveys, in Colton T (ed): *Statistics in Medicine.* Boston, Little, Brown and Company, 1974, p 271-288.

8. Colton T: Clinical trials, in Colton T (ed): *Statistics in Medicine.* Boston, Little, Brown and Company, 1974, p 253-269.

9. Elenbaas RM, McNabney WK, Robinson WA: Prophylactic oxacillin in dog bite wounds. *Ann Emerg Med* **11**:248-251, 1982.

10. May GS, DeMets DL, Fiedman LM, et al: The randomized clinical trial: Bias in analysis. *Circulation* **64**:669-673, 1981.

11. Gore SM: Assessing clinical trials — why randomize? *Br Med J* **282**:1958-1960, 1981.

12. Vrhovac B: Placebo and its importance in medicine. *Int J Clin Pharm* **15**:161-165, 1977.

13. Hill LE, Nunn AJ, Fox W: Matching quality of agents employed in "double-blind" controlled clinical trials. *Lancet* **1**:352-356, 1976.

14. Gore SM: Assessing clinical trials — double-blind trials. *Br Med J* **283**:122-124, 1981.

15. Spodick DH: Randomized controlled clinical trials — the behavioral case. *JAMA* **247**:2258-2260, 1982.

16. McHardy GG: A multicenter, double-blind trial of sucralfate and placebo in duodenal ulcer. *J Clin Gastroenterol* **3**(Supp 2):147-152, 1981.