Critically Reading a Research Article*

DENNIS L. HART,† MPA, PT, W. ROGER POSTON II,‡ MS, EdD, JAN F. PERRY,‡ MA, PT

This paper presents a method for critically reading a research article. Emphasis is placed on the specific format of the question being asked, the control of variables to ensure internal and external validity, selected basic rules for applying statistics, and the justification for specific conclusions drawn by the investigator. The method is designed to be clinically relevant with the final evaluation stressing improved patient care. This format has been effectively used by students of the Department of Physical Therapy at the Medical College of Georgia for critiquing literature.

The techniques of research are frequently foreign to clinicians, and yet research offers the justification and validation for our livelihood. We should make every effort to investigate our practice scientifically, validate the effective, eliminate the ineffective, develop the undiscovered, and critically analyze the published literature. To that end, therefore, this article presents a method which will provide clinicians with the basic steps necessary to analyze an article and determine whether its results are worthy of clinical application in their own practices.

OVERVIEW

The critical review of research must proceed along a specific path which may vary slightly depending upon the type of research conducted. Most reports of inferential investigations should start with an introduction that includes, most importantly, the statement of purpose or the statistically testable research question which the study was designed to answer. The clinical appropriateness of this question should be justified by a review of the literature, which is part of the introduction. The methodology, the next section of the report, should explain the research design and the steps taken to answer the question. Since the experiment should be reproducible, the report should be clearly and precisely stated.

The results section, which follows the methodology, should concisely state the statistical outcome of the data that was collected. No opinion should be introduced at this point, only the pertinent data.

In the next section, the conclusion or discussion, the author should relate the data to the previously mentioned pertinent literature, and clinically relevant conclusions should be offered. If the author has asked a clinically important question, implemented an unbiased plan to answer the question (internal validity), and statistically analyzed the data appropriately, then and only then can the reader begin to generalize the information to patient populations outside the subject sample of the investigation (external validity). If appropriate, the reader can then use the information in the treatment of patients.

Let us offer a word of caution. Medicine is still an art, and therefore many gray areas exist. The intelligent clinician will take pieces of information from many articles and combine them in the treatment of patients. Your own particular clinical situation will influence how the information obtained in the literature will be used.

The various parts of a research article have been quickly reviewed. A more detailed examination of what should be in those parts and how they can be read with a critical eye is now in order.

THE QUESTION

Every article published should have a stated purpose. The reader should not have to search for the purpose but should find it stated explicitly
in the introduction of the paper. The purpose may take the form of a statistically testable research question, a research hypothesis, or a null hypothesis or it can be editorialized into the statement of purpose. If the statement is not specific, the investigator either was not clear about what he or she was asking or did not know how to ask it. If the question is not stated, you cannot be sure that the study pursued a specific question.

Once you know what is being asked, identify the question as descriptive, correlational, or comparative. The descriptive question is the most basic question an investigator can ask. It seeks to describe or define something not yet studied or something that has been studied from another point of view. The descriptive study forms the base of the pyramid of research knowledge (Fig. 1). Each of the other questions is based on a description of what is occurring. Description questions frequently begin with the words "what," "how," or "how much." They usually do not lend themselves to anything other than descriptive statistics. The descriptive question "What is the effect of heat on nerve excitability?" cannot be answered by either a yes or no and is not statistically testable. Be cautious of the author who reports statistical significance for a descriptive question. If this occurs, the author has either applied inappropriate statistics to the original question or tested a question that was not originally asked. Analysis of a question not proposed prior to the collection of data easily lends itself to bias.

Correlational and comparative questions, the next two levels of the pyramid of research, are inferential and statistically testable. The correlational question is used to relate independent characteristics in the same subject. The question of whether the characteristics are or are not related can be answered with a yes or no and, therefore, can be analyzed statistically. The analysis will provide the reader with the type of relationship, direct or inverse. The typical correlational question is stated: "Is there a relationship between X and Y?" For the purpose of drawing conclusions from any study, the correlational question is stronger than the descriptive but not as strong as the comparative.

The comparative question is used to determine cause and effect relationships. It is based on the statistical assumption that there is no difference between any two treatments or between one treatment and no treatment at all (the null hypothesis). The basic question asked then is: "Is there a difference between X and Y?" This question can be answered by a yes or no and can be analyzed statistically.

In reviewing the question, make sure that it is justified by appropriate literature review. Pertinent references should be cited to support the concepts being studied, the reasons for the study, and the appropriateness of the study.

After a question is well formulated, the author should present the plan for answering it. This is done in the methodology.

**METHODOLOGY**

The methodology is a precise plan of action which the investigator follows to answer the research question. It should not reflect any opinions and should be written in language exacting enough to allow precise duplication of the study. If it is not reproducible, the method by which the authors attempted to answer their question is not totally known and is, therefore, difficult to assess.

The methodology should state exactly how specific variables were controlled. This control is one of the most important aspects of the study. Good control of all the variables that can influence the outcome (high internal validity) results in a meaningful study. If important variables are not controlled (low internal validity), the findings of the study are greatly compromised. Readers must be extremely wary of poor control of variables for all studies.

The methodology should also state precisely how the characteristics or variables were measured. Any methodology that does not provide the reader with an exact description of what was measured as well as how well it was measured should be questioned. In a descriptive study, the author must have precisely identified the variable or variables under examination.
To gain a general idea of the characteristics or variables that should have been considered, the reader should look back to the research question and ask: “What logically can affect the outcome of this study?” Items such as age, sex, past medical history, and cultural background may or may not affect various studies. The reader should be aware of how the author has controlled the impact that each trait has on the characteristics being studied. One means of controlling variables is to select subjects with similar characteristics (age, weight, sex). Anyone not fitting the predetermined criteria would not be included in the study, since they may adversely affect the measured characteristic. In general, greater control of variables in a study leads to stronger conclusions that can be made from that study.

To review a study for internal validity (the control of variables), the reader should ask several questions. First, “How many subjects are there?” The number of subjects should be sufficiently large to represent the size of the population from which the sample was drawn. For example, if the population is extremely small and homogeneous, the sample may be correspondingly small and still yield significant data. Conversely, if the population is large and heterogeneous, the sample must be correspondingly large to yield significant data. Generally, more subjects are considered appropriate for most unbiased statistical analyses.

Another question is: “How were the subjects selected?” Subject selection is based on criteria predetermined by the author. If not well planned, the process of subject selection may introduce bias into the study. Bias is evident in the following example. A researcher wants to determine whether or not total knee replacement surgery will reduce knee pain and selects only patients with severe rheumatoid arthritis to participate. The data from the study will not demonstrate the effectiveness of the total knee replacement surgery in general but only for the patients with severe rheumatoid arthritis. Based on this study, a statement of the general effectiveness of total knee replacement surgery would be erroneous and biased. The study may have good internal validity (control of variables); however, the external validity (ability to generalize) is poor, since the subject sample was narrow. Hence, the criteria for subject selection must be carefully considered, determined ahead of time, and related to the population from which the sample is to be drawn and the question being asked.

A third question is: “Were extraneous variables controlled?” As discussed earlier, this is a critical concern in reviewing an article. Consider the reasonableness of comparing coordination of a 20-year-old to a geriatric patient, cardiac work capacity of men and women, or interactive styles of inner-city children and rural children. In each case, at least one obvious variable was not controlled, i.e., age, sex, and socioeconomic interaction, respectively. These variables will bias the data.

The reader of investigative studies must differentiate between characteristics that influence the collection of data and those that may not. Generally, if the subjects are homogeneous in demographic data, then the study will be more specific and generalizable to a smaller, particular population. If the subjects represent a more varied, heterogeneous population, the data should be generalizable to a larger but still particular population.

The next question a reader might ask in reviewing an article could be: “Were the subjects placed into groups, and if so, how?” This question applies specifically to comparative studies. Usually, at least two groups receiving different treatments are compared. In some studies, an extra group, a “control group,” is used. The purpose of having groups is to enable the investigator to treat the groups exactly alike except for the treatment variable. This increases internal validity. If a difference in the outcomes between the groups occurs, the conclusion could be drawn that the treatment probably caused the difference. (One can never be absolutely sure, however, that the differences were not caused by chance alone.)

The method by which subjects were selected for each group is an important consideration. Was it a random selection, was selection based on the population being studied, or was bias introduced into the selection? If, for instance, all the subjects older than 60 years of age were placed into group A and those younger than 40 years of age were placed into group B, and age was not the variable being studied, the data could be biased. When at all possible, selection should be made at random and based on the population being studied to reduce bias.

A final question to ask is: “How were the subjects treated?” The general rule is that each subject or group of subjects should be handled exactly the same except for the variable being tested. Many variations to this rule exist. Each variation demands additional control to offset the
chance of introducing another confounding or uncontrolled variable.

The concerned investigator will state exactly how each subject or group was treated. While reviewing the study, recall the precise research question asked, the characteristics measured, the variables that were controlled, and those that were not controlled. Then ask yourself if the investigator really was measuring what he or she thought he or she was measuring. For instance, if the internal validity (control) was poor, the possibility that the investigator did not measure the characteristic necessary to answer the original question is fairly strong.

RESULTS

The results section of the article must report the data in an unbiased manner and without opinion. The data that supports or rejects the original research question should be presented along with the results of the statistical analyses. It should be presented in a clear, concise, and usable form. The reader should look for data that the author did not report but which he or she said were collected. That data may influence the answer to the original question.

Before exploring the results further, the statistical methods should be examined. A few simple rules should be kept in mind to guide your decisions concerning the appropriateness of the statistical analyses.

If the subjects cannot be assumed to represent the larger population, the study will require the application of a nonparametric testing procedure ($\chi^2$, rank correlation, etc.)

If however the sample does represent the population, statistics related to the normal distribution or parametric tests can be used ($t$ test, analysis of variance, correlation).

If the groups had fewer than 30 subjects, parametric tests must be applied carefully since the small number of subjects may inadequately represent the total population.

In a well-designed and well-controlled study, an isolated, aberrant value may occur. This may demand a change in the statistical analysis to control for the abnormal or skewed data. If this was done, the reasons for the change should be clearly stated.

As you review the results of the statistical analysis, refer back to the research question. The author should state the significant findings of the study. Again, they should be related to the research question. Keep in mind any conclusions you drew while reviewing the methodology, particularly the representativeness of the sample and the appropriateness of the statistical analysis. Many times authors use parametric statistics when the sample selection, size, or demographic characteristics would have dictated nonparametric statistics.

DISCUSSION

In this section, the author has the opportunity and is expected to offer an educated, justified opinion about the results of the study. Many authors overstep their responsibilities and offer opinions that were not justified by their study. To separate the opinions that were justified from the ones that were not, review the research question and relate it to the supporting literature review. Then, relate both of those to the results of the present study.

The results should state whether or not the research question was answered. The type of answer will be dependent on the type of question asked. A descriptive study will present a description of the subject studied. A correlational study should state whether or not the characteristics being studied were correlated and, if they were, how they were. Similarly, a comparative study should state how the items being studied did or did not compare.

Regardless of what the study revealed, the results should be related to the question, methodology, and results. If any conclusions are drawn that were indirectly but not specifically related to the question, methodology, and results, the strength of those conclusions is low.

Finally, did the author overlook any important analyses or conclusions? The reader's knowledge of the subject may allow viewing the study in a different manner that would generate different results and conclusions. That's good! Compare the two—yours and the author's. If there is a large discrepancy, the reader may wish to study the problem further.

Remember, the key to a sound clinically applicable study is the control of variables and justification of all conclusions. If these two qualifications are met, the value of the study in general increases. When unsupported conclusions are made, they must be examined for bias and may be purely conjectural. Some conjectures are educated assumptions and deserve consideration. They should be tested by an appropriate study. They should not be accepted as true based on the reported study. Finally, research
does not prove anything. It merely generates data that either supports or does not support the null hypothesis.

CONCLUSIONS

The final evaluation of any medically oriented study lies in its contribution to the advancement of patient care. If the study is not clinically applicable, its value is greatly reduced.

To get an impression of the overall evaluation of a study, the following questions should be answered:

1) Did the author answer the question that was asked?
2) Was the methodology appropriate and unbiased with strong internal validity?
3) Were the results unbiased?
4) Was the review of literature related to the results of the study and was it used to draw sound, logical conclusions?
5) Are the results applicable to the clinical situation?

If the questions can all be answered in the affirmative, the study is valid. However, no study can be applied to all clinical problems. Reflect upon the study when confronted with a similar patient pathology. Use the conclusions as a base of knowledge for a clinical frame of reference when reviewing a specific patient problem, but also keep in mind the limitations of the study. Never consider any results as proof of positive, definitive answers for any clinical situation. Patients are people and are always unique; they will always break the rules.

SUMMARY

Using the method presented in this paper to review research articles will allow students and clinicians to make better judgments about the volumes of research that are being published. This article has been directed toward research applied to clinical situations, review of which should improve patient care.

SELECTED BIBLIOGRAPHY