Research Session "임상약학 연구체계의 정립 및 활성화를 위하여"

임상약학 연구의 이론과 실제

Clinical Pharmacy Research: Theory and Practice

Charles D. Sands, Pharm.D.
Clinical Research Center
Wallace Memorial Baptist Hospital
Pusan, Korea
&
Associate Professor, College of Pharmacy
Samford University
Birmingham, Alabama 35229 USA

"The formulation of a research problem is far more often essential than its solution, which may be merely a matter of mathematical or experimental skill. To raise new questions, new possibilities, to regard old problems from a new angle requires creative imagination and marks real edvance in science."

Albert Einstein
The Evolution of Physics
Simon and Schuster, 1938

CLINICAL PHARMACY RESEARCH: THEORY AND PRACTICE

INTRODUCTION

An Overview of the research process
Challenge to the clinical pharmacist

THE ANATOMY OF RESEARCH: WHAT IT'S MADE OF

The research question
The significance (rationale)
The design
The subjects
The variables
Statistical analysis

THE PHYSIOLOGY OF RESEARCH: HOW IT WORKS

Designing the study
Implementing the study
Drawing causal inference
The errors of research

DESIGNING THE STUDY

The study protocol

SUMMARY

REFERENCES

APPENDIX

INTRODUCTION

An overview of the research process.

The philosophic cornerstone of research: drawing inferences about truth in the universe from events observed in the study sample.

Designing and implementing a research project can be divided into six steps.

Step 1. Choice of the research question.

a. A question about a health problem that the investigator wants to answer.
b. Choosing the right research question is the biggest challenge.
c. Review the existing medical literature.

Step 2. Developing the study protocol.

Objective: to design a feasible and inexpensive study that will produce a correct answer to the research question.

a. Subject selection
b. Variable measurements
c. Study design
d. Sample size
e. Ethical concerns
f. Data management and analysis
g. Statistical analysis
h. Limitations

Step 3. Pretesting the study.

Step 4. Implementation of the study.

Step 5. Analyzing the results of the study.

Step 6. Drawing conclusions and making inferences from the study.

THE ANATOMY OF RESEARCH: WHAT IT'S MADE OF

1. The study protocol.

    a. The written plan of the study.
    b. Helps organize the research in a logical, focused and efficient way.

2. The research question.

    a. The objective of the study.
    b. A problem that the investigator wants to solve.

        Example: Does ranitidine effect the metabolism of theophylline?

3. The significance.

    a. The rationale for doing the study.
    b. Why is the research question important?
    c. What kind of answers will the study provide?
    d. Review the medical literature.

        Example: Ranitidine and theophylline are often prescribed together. If ranitidine significantly effects theophylline metabolism there is the potential for either subtherapeutic or toxic theophylline serum concentrations. There are several isolated reports in the medical literature of theophylline toxicity possibly related to the addition to ranitidine to the drug regimen. Therefore it would important to know if this interaction actually exists.

4. The design.

    a. Observational study: the investigator observes the events without altering them.

        Example: A case-control study comparing the needle-sharing history of i.v. drug-abusers who have AIDS virus antibodies with the history of those who do not.

        (1). Cross-sectional study: Each subject is examined on only occasion.
        (2) Longitudinal study: each subject is followed over a period of time.

    b. Experimental study: the investigator applies an intervention (indendent variable), and

observes the effect on the outcome (dependent variable).
[Randomized clinical trial; controlled clinical trial]

Example: A randomized trial of the effect of ranitidine on theophylline serum concentrations.

c. Retrospective study (case-control study): deals with past events.

d. Prospective study (cohort study): deals with events that have not yet occurred when the study begins.

No one design is always better than the others. For each research question the investigator must make a judgement as to which design is the most efficient way to get a satisfactory answer. The randomized trial is often held up as the ultimate standard, but there are many situations for which an observational study is a better choice.

5. The Subjects

a. Selection criteria

(1) The process of defining the study population.
(2) The kinds of subjects best suited to the research question.
Normal (healthy) subjects; patients.
(3) Where and how to recruit the subjects.

b. Sampling

(1) The process of picking the subgroup of the population that will be the subjects of the study.
(2) Various sampling techniques.

6. Measurement of Variables

a. Variables: characteristics of the subjects to be measured.

b. Predictor variable:

(1) age, height, weight, race, smoker, non-smoker, etc.
(2) Intervention variable (independent variable).
(a) The investigator manipulates and observes the effect on the outcome variable.

Example: dose of drug, type of drug.

(3) Confounding variable: other predictors that may confuse the interpretation of the outcome.
Example: socioeconomic status, race, age

c. Outcome variable (response variable, dependent variable).
Example: SDC's, pharmacokinetic parameters, BP, HR, etc.

7. Statistical Analysis

Hypothesis Testing.

a. Experimental studies:

(1). Hypothesis: A hypothesis is a version of the research question that provides the basis for testing the statistical significance of the study results (outcomes).

Example: Ranitidine has a significant effect on theophylline serum concentrations when compared to placebo.

(2). Null hypothesis: the negative statement of the research hypothesis

Example: There is no difference between ranitidine vs placebo in their effects on theophylline serum concentrations

b. Descriptive studies

(1). Do not need a hypothesis.

(2). Because their purpose is to describe how variables are distributed (e.g., the prevalence of fast metabolizers in the population, etc.).

8. Sample Size Estimation

a. Experimental studies (those with a hypothesis)

(1) Estimating the number of subjects to observe the expected difference in outcomes between study groups.
(2) Power - The ability of a study to detect a false null hypothesis (Type II error)

b. Descriptive studies.

(1) Estimating the number of subjects needed to produce descriptive statistics (means, proportions, etc) of adequate precision.

9. Ethical Considertions.

a. General considerations.
b. International Ethical Guidelines for Biomedical Research (WHO).
c. Institutional Review Board (IRB).
d. Informed Consent.

THE PHYSIOLOGY OF RESEARCH: HOW IT WORKS

1. Designing the Study

a. The research question.

(1) What the investigator wants to answer (e.g., Does ranitidine effect the metabolism of theophylline?).
(2) Impossible to study all subjects in the world (e.g., all subjects taking ranitidine and theophylline together).
(3) Transformation of the research question to the study plan:

(a) Must change the question to that which can be answered by the study.

Example: Does ranitidine, compared to placebo, effect the metabolism of theophylline in twelve normal Korean adults at the Baptist Hospital in Pusan?

(b) Must choose the variables that will represent the phenomena of interest.

Example: How do we measure the metabolism of theophylline? One way might be to measure the pharmacokinetic parameter, clearance, and consider it a measurement of metabolism.

(c) Must now express the simple research question in terms of the objective of the study.

General outline to transform the research question into the study objective: Don't ask is A better than B? Be specific, ask:

"In population W, is drug A, at a daily dose X, more efficacious in reducing Z over a period of time T, than drug B at daily dose Y?"

Example:

Study Objective: To determine if concurrent administration of ranitidine (150mg bid X 7 days) will impair the total body clearance (Cl) of orally administered SR-theophylline (10mg/kg/day X 7 days) in healthy adult ethnic Koreans in China?

2. Implementing the Study

Research Question	→	Study Protocol	→	Actual Study
	Design		Implement	
	External Validity		Internal Validity	

---

3. The Errors of Research

a. All studies have some error.
b. Goal: to maximize internal and external validity...
   so that the inferences can be applied to the study population.
c. Types of error.
   (1) Random error: a wrong result due to chance (unknown sources of variation)
   (2) Systematic error: a wrong result due to bias
   (3) Sampling error
   (4) Measurement error
   (5) Stistical errors
      (a) Type I error: the rejection of a true null hypothesis
      (b) Type II error: the failure to reject a false null hypothesis.
   (5) Inferential error: all of the above can lead to errors in making inferences about the study results.
d. To get the right answer to the research question:
   Design and implement the study to minimize inferential errors.

DESIGNING THE STUDY PROTOCOL

1. Developing the Study Protocol

a. First step: establish the research question (previously discussed).
b. Four versions (steps) in the development of the study protocol:

Step (1) The Research Question: a one sentence statement of the research question

Step (2) Study Outline: a 1-2 page outline of the elements of the study.

Element	Purpose
Research question (objectives)	What questions will the study address?
Significance (background, rationale)	Why are these questions important? Literature review.
Design	How will the study be carried out?
Time frame	
Epidemiologic approach	
Subjects	Who are the subjects and how will they be selected?
Selection criteria	
Sampling design	
Variables	What measurements will be made?
Predictor variables	
Outcome variables	
Statistical Analysis	How large is the study? How will the data be analyzed? What statistical tests will be used?
Hypothesis	
Sample size estimation	
Power	
Analytic approach	
Ethical Considerations	

Step (3) Study Protocol:

A fleshed in version of the 1-2 page outline.
Usually 5 to several hundred pages.
The main document used to plan the study.

Step (4) Operational Manual (Procedure Manual)

The main document used to guide clinical and laboratory procedures and record the data.
A collection of:

procedures
instructions
data collection forms

2. Making Inferences.

Inference: an important statistical concept.

The data obtained from the study of a small number of subjects are then used to make an "educated guess" (inference) about the effects on the larger population of interest.

SUMMARY

1. Anatomy of Research: the elements that make up the study plan
   - Research question
   - Design
   - Study subjects
   - Measurement approaches
   - Statistical Analysis

2. Physiology of Research: how the study works.

   Inferences: The study results are used to make inferences about (1) what happened in the study sample (internal validity) and about (2) events in the outside world (external validity).

   Error: The challenge is to design and implement a study plan with adequate control to minimize random error (chance) and systematic error (bias).

3. Developing the Study Protocol.

   Research question: a one sentence statement of the main question the investigator is trying to answer.

   Study Outline: a 1-2 page outline that lists the elements of the study.

   Study Protocol: the completed study plan for carrying out the research.

   Operational Manual: list of various clinical procedures, data collection form, etc.

4. Inferences.

   Consider the main inferences that will be drawn from
   a. the study subjects to the population
   b. the study measurements to the phenomena of interest

Consider the relationships between

a. the research question (what the investigator wants to answer in the world outside)

b. the study plan (what the study is designed to answer)

c. the actual study (what the study actually finds, given the errors of implementation)

REFERENCES

Books

1. Gehlback SH. Interpreting the Medical Literature: Practical Epidemiology for Clinicians, 2nd Edition. New York: Macmillan Publishing Co., 1988.
2. Hawkins C, Sorgi M. Research: How to Plan, Speak and Write About it. New York: Springer-Verlag, 1985.
3. Hulley SB, Cummings SR. Designing Clinical Research: An Epidemiologic Approach. Baltimore: Williams & Wilkins, 1988.
4. Infelfinger JA, Mosteller F, Thigodeau LA, Ware JH. Biostatistics in Clinical Medicine, 2nd Edition. New York: Macmillan Publishing Co., 1987.
5. Spilker B. Guide to Clinical Interpretation of Data. New York: Raven Press, 1986.
6. Spilker B. Multinational Drug Companies: Issues in Drug Discovery and Development. New York: Macmillan Publishing Co., 1989.
7. Riegelman RK, Hirsch RP. Studying a Study and Testing a Test: How to Read the Medical Literature, 2nd Edition. Boston: Little, Brown and Company, 1989.
8. Bolton S. Pharmaceutical Statistics: Practical and Clinical Applications, 2nd Edition. New York: Marcel Dekker, 1990.
9. Marks, RG. Designing a Research Project: The Basics of Biomedical Research Methodology. New York, Van Nostrand Reinhold Company, 1982.
10. Spilker B. Guide to Clinical Studies and Developing Protocols. New York: Raven Press, 1984.
11. Shapiro SH, Louis TA. Clinical Trials: Issues and Approaches. New York: Marcel Decker, 1983.
12. CR, Tsay JY. Statistics in the Pharmaceutical Industry. New York: Marcel Dekker, 1981.
13. Bailar JC, Mosteller F. Medical Uses of Statistics. Waltham, Massachusetts: NEJM Books, 1986.
14. Spilker, B. Guide to Planning and Managing Multiple Clinical Studies. New York: Raven Press, 1987,
15. Friedman LM, Furberg CD, DeMets DL. Fundamentals of Clinical Trials, 2nd Edition. Littleton, Massachusetts: PSG Publishing Co., 1985.
16. Norman GR. Streiner DL. PDQ Statistics. Philadelphia: B.C. Decker, 1986.
17. Smith RB, Kroboth PD, Juhl RP. Pharmacokinetics and Pharmacodynamics: Research Design and Analysis. Cincinnati: Harvey Whitney Books, 1986.
18. Schefler WC. Statistics for Health Professionals. Reading, Massachusetts: Addison-Wesley Publishing Co., 1984.
19. Press SJ. Bayesian Statistics: Principles, Models, and Applications. New York: John Wiley & Sons, 1989.
20. Lwanga SK, Tye CY. Teaching Health Statistics. Geneva: World Health Organization, 1986.
21. Day RA. How to Write and Publish a Scientific Paper, 3rd Edition. Phoenix: Oryx Press, 1988.
22. International Ethical Guidelines for Biomedical Research Involving Human Subjects. Geneva:

World Health Organization, 1993.
23. Grantsmanship: Research Planning and Proposal Writing. Bethesda: American Society of Hospital Pharmacists, 1993.

Journal Articles

1. Bailar JC, Mosteller F. Guidelines for Statistical Reporting in Articles for Medical Journals: Amplifications and Explanations. Annals of Internal Medicine 1988;108:266-73.
2. International Committee of Medical Journal Editors. Uniform Requirements for Manuscripts submitted to Biomedical Journals. New England Journal of Medicine 1991;324:424-428.
3. Stolley PD, Strom BL. Sample aize calculations for clinical pharmacology studies. Clin Pharmacol Ther 1986;May:489-490.
4. Levine RJ. New International ethical guidelines for research involving human subjects. Annals of Internal Medicine 1993;119:339-341.

APPENDIX A: Protocol Outline (Example)

Title of Research: The Effect of Ranitidine on Theophylline Metabolism in Ethnic Koreans in China

I. INTRODUCTION

1. Synopsis of the medical problem to be evaluated.
2. Literature review.
3. Rationale for the present study.
4. Research question.
5. Hypothesis.

II. EXPERIMENTAL DESIGN

1. Research Objective.
2. Measurement Variables.
    a. Predictor Variables
        (1) Intervention variables (independent variables)
        (2) Confounding variables
    b. Response Variables (outcome variables, dependent variables)
3. Study design
4. Research definitions
    a. Universe of interest.
    b. Experimental unit.
    c. Population of interest.
    d. Sample
5. Subject Selection.
    a. Selection criteria.
    b. Sampling technique.
6. Sample size estimation.
7. Control of bias.
8. Randomization schedule (random assignment of subjects to treatment groups).
9. Drug Dosing regimen.
10. Adverse effects assessment.
11. Pharmacokinetic analysis.
12. Description of the data.
13. Statistical analysis.
    a. Power (retrospective)
14. Measurement of theophylline serum concentrations.
15. Clinical procedures.
16. Facilities and equipment.

III. ETHICAL CONSIDERATIONS.

1. General considerations.
    a. International Ethical Guidelines for Biomedical Research
2. Institutional Review Board (IRB)
3. Informed Consent.

IV. FINANCES

1. Budget
2. Submitting a grant proposal.
3. Sources of income.
4. Expenses

V. LIMITATIONS OF THE STUDY.

VI. REFERENCES

VII. APPENDIX.

A. Study design: flow chart.
B. Data collection form
C. Time and event schedule.
D. Informed consent form.
E. Adverse effects assessment form
F. Description of data (tables, charts, figures, etc.)
G. Statistical Data

Reprinted from Annals of Internal Medicine 1988;108:266-73.
Reprinted with permission.

Guidelines for Statistical Reporting in Articles for Medical Journals

Amplifications and Explanations

JOHN C. BAILAR III, M.D., Ph.D.; and FREDERICK MOSTELLER, Ph.D.; Boston, Massachusetts

The 1988 edition of the *Uniform Requirements for Manuscripts Submitted to Biomedical Journals* includes guidelines for presenting statistical aspects of scientific research. The guidelines are intended to aid authors in reporting the statistical aspects of their work in ways that are clear and helpful to readers. We examine these guidelines for statistics using 15 numbered statements. Although the information presented relates to manuscript preparation, it will also help investigators in earlier stages make critical decisions about research approaches and protocols.

[MeSH terms: clinical protocols; clinical trials; eligibility determination; manuscripts, medical; probability; random allocation; statistics. *Other indexing terms:* blinding; blocking; confidence intervals; International Committee of Medical Journal Editors; matching; *P* values; software; statistical methods; stratification; study design; treatment complications; Uniform Requirements for Manuscripts]

In 1979, the group now known as the International Committee of Medical Journal Editors first published a set of uniform requirements for preparing manuscripts to be submitted to their own journals. These uniform requirements have been revised several times (1), and have been widely adopted by other biomedical journals. In the 1988 revision (2), the Committee added guidelines for presenting and writing about statistical aspects of research. The purpose of these guidelines is to assist authors in reporting statistical aspects of their research in ways that will be responsive to the queries of editors and reviewers and helpful to readers.

We present the statistical guidelines as a sequence of 15 numbered statements, and amplify and explain some of the reasoning behind the guidelines. The material focuses on manuscript preparation, but it should also be helpful at earlier stages when critical decisions about research approaches and protocols are made. This article does not provide a short course in statistics because we can deal with only a few important aspects of what should be reported in publications about work already done, but we provide references to general statistical texts. The International Committee is not responsible for these amplifications; however, we have tried to present the spirit of the Committee's discussions as well as our own views.

The International Committee's statistical guidelines are as follows:

From the Department of Health Policy and Management, School of Public Health, Harvard University; Boston, Massachusetts; Office of Disease Prevention and Health Promotion, U.S. Dept. of Health and Human Services, Washington, D.C.; Department of Epidemiology and Biostatistics, McGill University; Montreal, Quebec, Canada.

Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results. When possible, quantify findings and present them with appropriate indicators of measurement error or uncertainty (such as confidence intervals). Avoid sole reliance on statistical hypothesis testing, such as the use of *P* values, which fails to convey important quantitative information. Discuss eligibility of experimental subjects. Give details about randomization. Describe the methods for, and success of, any blinding of observations. Report treatment complications. Give numbers of observations. Report losses to observation (such as dropouts from a clinical trial). References for study design and statistical methods should be to standard works (with pages stated) when possible, rather than to papers where designs or methods were originally reported. Specify any general-use computer programs used.

Put general descriptions of methods in the Methods section. When data are summarized in the Results section specify the statistical methods used to analyze them. Restrict tables and figures to those needed to explain the argument of the paper and to assess its support. Use graphs as an alternative to tables with many entries; do not duplicate data in graphs and tables. Avoid non-technical uses of technical terms in statistics, such as "random" (which implies a randomizing device), "normal," "significant," "correlation," and "sample." Define statistical terms, abbreviations, and most symbols.

Our general approach is that scientific and technical writing should be comprehensible at the first reading for the average reader who is knowledgeable about the general area but not a subspecialist in the specific topic of investigation.

1. Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results.

Authors should report which statistical methods they used, and why. In many instances they should also report why other methods were not used, although this is rarely done.

Readers must be told about weaknesses in study design and about study strengths in enough detail to form a clear and accurate impression of the reliability of the data, as well as any threats to the validity of findings and interpretations. Such details are often omitted, although investigators probably know them (3, 4).

The researcher must decide which statistical measures and methods are appropriate, given that a statistical goal has been defined. Investigators often have a choice: Mean or median? Nonparametric test or normal approximation? Adjustment, matching, or stratification to deal with confounding factors? Choosing statistical methods generally requires an appreciation of both the problem and the data, and an experienced biostatistician, statistician, or

epidemiologist can often provide substantial help. This help ideally begins before the study, because the foundation for reporting one's findings is laid before the study even begins.

Trying several reasonable statistical methods is often appropriate, but this strategy must be disclosed so that readers can make their own adjustments for the authors' industriousness or skill in fishing through the data for a favorable result. Whatever statistical task is defined, it is inappropriate, and indeed unethical, to try several methods and report only those results that suit the investigator. Results of overlapping methods need not be presented separately when they largely agree, but authors should state what additional approaches were tried, and that they did agree. Of course, results that do not agree also should be given, and investigators may sometimes find that such disagreements arise from important and unexpected aspects of the data.

Units should always be specified in text, tables, and figures, although not necessarily every time a number appears if the unit is clear to the reader. Often, careful choice of units of measurement can help clarify and unify the study question, biological hypothesis, and statistical analysis. Careful reporting of units can also help to avoid serious misunderstanding. Are quantities in milligrams or millimoles? Are rates per 10 000 or per 100 000? Does a figure show number of different patients, or number of myocardial infarcts among those patients (including second infarcts), or number of admissions to a given hospital (including readmissions)? Research investigators often use an abbreviated language that is clear to their colleagues, but they may have to make a special effort to assure that such usage will not confuse nonspecialists, or even other experts.

2. When possible, quantify findings and present them with appropriate indicators of measurement error or uncertainty (such as confidence intervals).

Investigators have to choose a way to report their findings. The most useful ways give information about the actual outcomes, such as means and standard deviations as well as confidence intervals. The tendency to report a test of significance alone—rather than with this additional information—should be resisted, although a significance test in the context of other information may be helpful.

Readers have many reasons for studying a research report. One reason is to find out how a particular treatment does in its own right, not just in comparison with another treatment. At a minimum, readers should be offered the mean and standard deviation for every appropriate outcome variable. Significance levels (P values), such as $P = 0.03$, are often reported to show that the difference seen or some other departure from a standard (a null hypothesis) had little probability of occurring if chance alone was the cause. Merely reporting a P value from a significance test of differences loses the information about both the average level of performance and the variability of individual outcomes for the separate treatments.

Exact P values rather than statements like "$P < 0.05$" or "P not significant" should be reported where possible so that readers can compare the calculated value of P with their own choice of critical values. In addition, other investigators may need exact values of P if they are to combine results of several separate studies.

In independent samples, information about means, standard deviations, and sample sizes can often be readily converted to a significance test and thus into a P value. From the P value alone, none of the others can be reconstructed, so that important information is lost when only a P value is reported (5, 6).

Make clear whether a reported standard deviation is for the distribution of single observations, or for the distribution of means (standard errors), or for the distribution of some other statistic such as the difference between two means. If the standard deviation for single observations is given, together with sample sizes, then in independent samples the reader can compute the other standard deviations.

Each statistical test of data implies both a specific null hypothesis about those data (such as "The 60-day survival rate in Group A equals that in Group B," so that the difference is zero) and a specific set of alternative hypotheses (such as "The survival rate is different in Group B," which allows for a range of values for the difference). It is critical that both the null hypothesis and the alternatives be clearly stated, although many authors fail to do so. Clear reporting will not only help readers, it is also likely to reduce the frequency of abuse of P values.

It is critical also that authors specify how and when they developed each null hypothesis in relation to their consideration of the data. Statistical theory requires that null hypotheses be fully developed before the data are examined—indeed, before even the briefest view of preliminary results. Otherwise, P values cannot be interpreted as meaningful probabilities.

Authors should always specify whether they are using two-tail or one-tail tests.

3. Avoid sole reliance on statistical hypothesis testing, such as the use of P values, which fails to convey important quantitative information.

Confidence intervals offer a more informative way to deal with the significance test than does a simple P value. Confidence intervals for a single mean or a proportion provide information about both level and variability. Confidence intervals on a difference of means or proportions provide information about the size of difference and its uncertainty, but not about component means, and these should be given.

A significance test of observed data, generally to determine whether the (unknown) means of two populations are different, usually winds up with a score that is referred to a table, such as a t-, normal-, or F-table. The table then presents the P value.

Although confidence limits offer appraisals of variability and uncertainty, in some studies, such as certain large epidemiologic and demographic studies, biases are often greater threats to the validity of inferences than ordinary random variability (expressed in the standard deviation). Coding or typing errors may exaggerate the number of

deaths from a cause, nonresponse to treatment may be selective (those patients more ill being less likely to respond), and so on. Although the potential sources of bias are many, books on applied statistics, epidemiology, and demography alert the research worker to common difficulties, and often to steps that may be taken toward their amelioration.

4. Discuss eligibility of experimental subjects.

Reasons for and methods of selecting patients or other study units should always be reported, and if the selection is likely to matter, the reasons should be reported in detail. The full range of potentially eligible subjects, or the scope of the study, should be precisely stated in terms that readers can interpret. It is not enough to say that the natural history of a condition has been seen in "100 consecutive patients." How do these patients compare with what is already known about the condition in terms of age, sex, and other factors? Are patients from an area or population that might be special? Are patients from an "unselected" series with an initial diagnosis, or do they include referral patients (weighted with less serious or more serious problems)? In comparing outcomes for patients who underwent surgery to outcomes for patients treated medically, were the groups in similar physical condition initially? What about probable cases not proved? Many other questions will arise in specific instances. Sometimes information is obvious (for example, if the investigator studied patients from one hospital because that is where he or she practices). Other questions about scope need answers. (Why begin on 1 January 1983? Why include only patients admitted through the emergency room?) Authors should try to imagine themselves as readers who know nothing about the study.

Although every statistically sound study has such "scope" criteria to determine the population sampled by the investigator, many also have more detailed "eligibility" criteria. Medical examples include the possible exclusion of patients outside a specified age range, those previously treated, those who refuse randomization or are too ill to answer questions, and other groups.

Which criteria are used to establish scope and which are used to establish eligibility may be uncertain, although both must be reported. Scope pushes study boundaries outward, toward the full range of patients or other study units that might be considered as subjects, whereas eligibility rules narrow the scope by removing units that cannot be studied, that may give unreliable results, that are likely to be atypical (for example, the extremes of age), that cannot be studied for ethical reasons (for example, pregnant women in some drug studies), or that are otherwise not appropriate for individual study.

The first goal is to state both scope and eligibility so that another knowledgeable investigator, facing the same group of patients or other study units, would make nearly the same decisions about including patients in the study.

The second goal is to provide readers with a solid link between the patients or cases studied and the population for which inferences will be made. Both scope and eligibility constraints can introduce substantial bias when results are generalized to other subjects, and readers need enough information to make their own assessment of this potential. Thus, reasons for each eligibility criterion should be stated. The two critical elements in setting the base for generalization are first to document each exclusion under the eligibility criteria with the reasons for that exclusion; and second, to present an accounting (often in a table) of the difference between patients falling within the scope of the study and those actually studied. The article should also say how patients excluded for more than one reason are handled; common approaches are to show specific combinations or to use a priority sequence. Such information helps the reader better understand how the study group is related to the population it came from, and also helps to assure that all omissions are accounted for. It should be so stated if no subject was ineligible for more than one reason.

Another critical element in reporting is to say how and when the scope and eligibility criteria were devised. Were scope and eligibility criteria set forth in a written protocol before work was started? Did they evolve during the course of the study? Were some eligibility criteria added at the end to deal with some problems not foreseen? For example, a written protocol might call for the study of "all" patients, but if only 5% of patients were female, they might be set aside at this point—especially if they are thought to differ from male patients in ways relevant to the subject of the study.

5. Give details about randomization.

The reporting of randomization needs special attention for two reasons. First, some authors incorrectly use "random" as a synonym for "haphazard." To prevent misunderstanding, simply tell readers how the randomization was done (coin toss, table of random numbers, cards in sealed envelopes, or some other method). Readers will then know that a random mechanism was in fact applied, and they can also judge the likelihood that it was subject to bias or abuse (such as peeking at cards). Second, randomization can enter in many ways. For example, a sample may be selected from a larger population at random, or study patients may be randomly allocated to treatments, or treated patients may be randomly given one or another test. Thus, it is not enough just to say that a study was "randomized." The many possible roles of randomization can be dealt with by careful reporting to assure there is no ambiguity.

Even with randomization, imbalances occur, with their predicted frequency, and these may need attention even if they do not call for special steps in the analysis. Stratification or matching may be used in combination with randomization to increase the similarity between the treated and control groups, and should be reported. Sometimes an assessment of the efficacy of stratification or matching in overcoming the imbalance is feasible; if so, it should be done and reported.

If the randomization was "blocked" (for example, by arranging that within each successive group of six patients, three are assigned to one treatment and three to another), reasons for blocking and the blocking factors should be given. Blocking should ordinarily affect statisti-

cal analysis, and authors should say how they used blocking in their analysis or why they did not.

6. Describe the methods for, and success of, any blinding of observations.

"Blinding," sometimes called "masking," is the concealment of certain information from patients or members of the research team during phases of a study. Blinding can be used to good effect to reduce bias, but because it can be applied in different ways, a research report should be explicit about who was blinded to what. An unadorned statement that a study was "blind" or "double blind" is rarely enough.

Patients may be blinded to treatment, or to the time that certain observations are made, or to preliminary findings regarding their progress. A decision to admit a patient to a study may be made blind to that patient's specific circumstances, and a decision that a patient randomized to treatment was not eligible may be made blind to the assigned treatment. The observer who classifies clinical outcomes may be blinded to the treatment, as may be the pathologist who interprets specimens or the technician who measures a chemical substance. These and other efforts to prevent bias by blinding should be reported in enough detail for readers to understand what was done.

The effectiveness of blinding should also be discussed in any situation where the person who is blinded may learn or guess the concealed information, such as by side effects that may accompany one treatment but not another. Such discoveries are particularly important for observations reported by patients themselves and for third-party observations of endpoints with a subjective component, such as level of patient activity.

A particularly critical aspect of blinding is whether the decision to admit a patient to a study was made before (or otherwise entirely and demonstrably independent of) any decision about choice of treatment to be used or offered. Where random allocation to treatments is used, the timing of randomization in relation to the decision to admit a patient should always be stated.

7. Report treatment complications.

Any intervention, or treatment, has some likelihood of causing unintended effects, whether the study is of a cell culture, a person, an ecologic community, or a hospital management system. Side effects may be good (quitting smoking reduces the risk of heart disease as well as the risk of lung cancer) or bad (drug toxicity). Side effects may be foreseen or unexpected. In most studies side effects will be of substantial interest to readers. Does a drug cause so much nausea that patients will not take it? If we stock an ecologic area with one species, what will happen to a predator? Does a new system for scheduling the purchase of hospital supplies at lower overall cost change the likelihood that some item will be exhausted before the replacement stock arrives?

Nearly every medical treatment carries some risk of complications—that is, of unintended adverse effects. Such effects should be sought at least as assiduously as beneficial effects, and they should be reported objectively and in detail. Treatment failure often gives the most useful information from a study. If no adverse effects can be found, the report should say so, with an explanation of what was done to find them.

8. Give numbers of observations.

The basic observational units should be clearly specified, along with any study features that might cause basic observations to be correlated. A study of acid rain might take samples of water from five different depths in each of seven different lakes—35 measurements in all. But the relevant sample size for one or another purpose may be five (depths), or seven (lakes), or 35 (depths in different lakes). In a metaanalysis of such work (7) the whole study may count as only a single observation. Lake water may tend to mix, so that five samples from different depths tell little more about acidity than a single sample; or lake-to-lake differences may be small within a geographic region, so that the study of one lake effectively studies them all.

Similarly, a study in several institutions of rates of infection after surgery may be considered to have a sample size of three hospitals, 15 surgeons, 600 patients, or 3000 days of observation after surgery. But infection rates may differ so much by hospital or surgeon that it is more important to include many hospitals or surgeons, perhaps with only a few patients from each, than to have large samples per surgeon.

Reporting decisions about the basic unit of observation and about sample size, as well as proper method of analysis, may require an informed understanding of statistics as well as the subject matter. The analysis and reporting of correlated observations, such as the water samples and the infection rates described above, raise difficult issues of statistical analysis that often require expert statistical help.

A different kind of problem arises from ambiguity in reporting ratios, proportions, and percents, where the denominator is often not specified and may be unclear to readers. Authors should be meticulous about specifying which study units are included in denominators (which then specifies the group examined) each time there may be any uncertainty.

Whatever the investigators adopt as their basic unit of observation, relationships to and possible correlations with other units must be discussed. Such internal relationships can sometimes be used to strengthen an analysis (when a major source of difference is balanced or held constant), and sometimes they weaken the analysis (by obscuring a critical limitation on effective sample size). Complicated data structures require special attention in study reporting, not just in study design, performance, and analysis.

9. Report losses to observation (such as dropouts from a clinical trial).

When the sample size for a table, graph, or text statement differs from that for a study as a whole, the difference should be explained. If some study units are omitted (for example, patients who did not return for 6-month follow-up), the reduced number should be reconciled with the number eligible or expected by readers. Reporting of losses is often easiest in tables, where entries such

as "patients lost," "samples contaminated," "not eligible," or "not available" (for example, no 15-meter sample from a lake with a maximum depth of 10 meters) can account for each study unit.

Loss of patients to follow-up, including losses or exclusions for noncompliance, should generally be discussed in depth because of the likelihood that patients lost are atypical in critical ways. Have patients not returned for examination because they are well? Because they are still sick and have sought other medical care? Because they are dead? Because they do not wish to burden a physician with a bad outcome? Failure to discuss both reasons for loss (or other termination of follow-up) and efforts to trace lost patients are common and serious. Similarly, issues of noncompliance (reasons, as well as numbers) are often slighted by authors.

10. References for study design and statistical methods should be to standard works (with pages stated) when possible rather than to papers where designs or methods were originally reported.

An original paper can have great value for the methodologist, but often does little to explain the method and its implications or the byways of calculation or meaning that may have emerged since the method was first reported. Standard works such as textbooks or review papers will usually give a clearer exposition, put the method in a larger context, and give helpful examples. The notation will be the current standard, and the explanation will orient readers to the general use of the method rather than the specific and sometimes peculiar use first reported. For example, it would be hard to recognize Student's t-distribution in his original paper; indeed, "t" was not even mentioned. Exceptions to the general advice about using textbooks, review papers, or other standard works occur where the original exposition is best for communication and where it is the only one available.

11. Specify any general-use computer programs used.

General-purpose computer programs should be specified, with the computer that ran them, because such programs are sometimes found to have errors (8). Readers may also wish to know about these programs for their own use. In contrast, programs written for a specific task need not be documented, because readers should already be alert to the likelihood of errors in ad hoc or "private" programs, and because they will not be able to use the same programs for their own work.

12. Put general descriptions of statistical methods in the Methods section. When data are summarized in the Results section, specify the statistical methods used to analyze them.

Where should statistical methods be described? There are good arguments for putting such material in one place, usually in the Methods section of a paper, but our preference (9) is generally to specify statistical methods at the places where their uses are first presented. Methods may differ slightly from one to another application within a given paper; and decisions about which results to report in full, or which methods to use in exploring critical or unexpected findings, generally depend on the data and earlier steps in the analysis. Keeping the specification of statistical methods close to their point of application will sometimes lead to more thought about choices and to better discussion of why a particular method was used in a particular way. Some editors, as well as some of our statistical colleagues, disagree, and authors should follow the instructions of the journal to which they submit their work.

Statements such as "statistical methods included analysis of variance, factor analysis, and regression, as well as tests of significance," when divorced from the outcomes or reasons for their use, give the reader little help. On the other hand, if the only method was the use of chi-squared tests for 2×2 contingency tables, that fact might be sufficiently informative.

Some general suggestions about reporting clinical trials have been discussed by Mosteller and associates (10).

13. Restrict tables and figures to those needed to explain the argument of the paper and to assess its support. Use graphs as an alternative to tables with many entries; do not duplicate data in graphs and tables.

Authors have an understandable wish to tell readers everything they have learned or surmised from their data, but economy is much prized by scientific readers as well as editors. A basic point is that economy in writing and exposition gives an article its best chance of being read. Although many tables may help support the same basic point, and might be appropriate in a monograph, an article generally requires only enough information to make its point—the mathematician's concept of "necessary and sufficient."

There are occasional exceptions. Sometimes the study generates data that have consequences beyond the article. For example, if information about certain biological or physical constants is obtained, it should be retained in the article. An author should inform the editor of this situation in a cover letter. Sometimes such data need to be preserved, but not in the article itself; many journals have some plan for the preservation and documentation of unpublished supporting material. Such plans are often mentioned in a journal's instructions to authors.

Whether tables or graphs better present material is sometimes a vexing question. Some readers go blind when faced with a table of numbers; others have no idea how to read graphs; unfortunately, these groups are not mutually exclusive, and some users of statistical data need to see quantitative findings in text. Overall there is a general failure to tolerate or understand the problems of any group that does not include oneself. Most of what we know about tables and graphs comes from the personal experiences of a few scholars, and little scientific information has been gathered on these subjects. Cleveland (11) has begun some scientific studies of what information can be communicated with graphs (for example, many people read bar charts better than pie charts). Tufte (12) has a beautiful book on the art of graphics.

In the field of tabular presentation, even less scientific investigation has been done, but there seems to be much value in some rules proposed by Ehrenberg (13): Give marginal (row and column) averages to provide a visual focus. Order the rows and columns of the table by the

Table 1. Infant Mortality Rates in the United States, All Races, 1964 to 1966, by Geographic Region and Level of Father's Education*

Region	Education of Fathers (in Years of Schooling)				
	≤ 8	9-11	12	13-15	≥ 16
Northeast	25.32	25.29	18.26	18.29	16.34
North Central	32.09	29.04	18.78	24.32	19.02
South	38.81	31.02	19.33	15.66	16.79
West	25.37	21.09	20.29	23.97	17.52

* Data given as number of deaths per 1000 live births. Adapted from a report of the U.S. Department of Health, Education, and Welfare (14).

marginal averages or some other measure of size or other logical order (keeping to the same order if there are many similar tables). Put figures to be compared into columns rather than rows (with larger numbers on top if possible). Round to two effective (significant) digits. Use layout to guide the eye and facilitate comparisons. In the text give brief summaries to lead the reader in the main patterns and exceptions.

To show the effect of Ehrenberg's rules, we devised Table 1 showing data on infant mortality, and we used Ehrenberg's rules to produce Table 2. Our primary interest is in the association of the father's education with infant mortality, with a secondary interest in region.

Table 1 is obviously "busy" with four-digit numbers, and we have reduced them to two digits. Table 2, with fewer digits, is easier to read although it has more numbers.

Because our primary interest is in the father's education, we put years of education in the rows.

We want the big numbers at the top of the table, so in arranging the rows we started with the lowest level of education rather than the highest. We did not reorder the rows because years of education already provided an order. The regions were reordered according to their average values. The issue of whether to put northeast or north central first depends on whether we want to emphasize what is best or what is poorest. Some people like to have numbers rising as the eye goes from left to right.

We have added averages for the rows and for the columns, and given the grand mean without additional decimals to keep the table simple.

The text might read as follows: "The table shows that the infant death rate has a grand mean of 23 per 1000 live births. Lower education of the father is associated with higher infant mortality, but education beyond the completion of high school (12 years) seems to have no further beneficial effect on the infant mortality rate. The northeast and west have the lowest rates, and the south did slightly better than the north central region. Father's education seems to matter more than region of the country, a variation of 13 deaths per 1000 births for education (range, 30 to 17) compared with 5 for regions (range, 20 to 25). The highest rate seen was in Southern families whose father had no more than a grammar school education (no more than 8 years). The lowest rate was 16, the highest 39, a ratio of nearly two and a half."

14. Avoid non-technical uses of technical terms in statistics, such as "random" (which implies a randomizing device), "normal," "significant," "correlation," and "sample."

Many words in statistics, and in mathematics more generally, come from everyday language and yet have specialized meanings. Thus, when statistical reporting is an important part of a paper, the author should not use statistical terms in their everyday meanings.

The family of *normal* (or Gaussian) distributions refers to a collection of probability distributions described by a specific formula. The distribution of *usual* or *average* values of some quantity found in practice is rarely "normal" in the statistical sense, even when the data have a generally bell-shaped distribution. *Normal* also has many other mathematical meanings, such as a line perpendicular to a plane. When we mix these meanings with the meaning of "normal" for a patient without disease, we have the makings of considerable confusion.

Significance and related words are used in statistics, and in scientific writing generally, to refer to the outcome of a formal test of a statistical hypothesis or test of significance (essentially the same thing). *Significant* means that the outcome of such a test fell outside a chosen, predetermined region. Careful statisticians and other scientists often distinguish between statistical and medical or social significance. For example, a large enough sample might show statistically significant differences in averages on the order of one tenth of a degree in average body temperature of groups of humans. Such a difference might be

Table 2. Infant Mortality Rates in the United States, All Races, 1964 to 1966, by Geographic Region and Level of Father's Education*

Education	Region				Average
	Northeast	West	South	North Central	
≤ 8	25	25	39	32	30
9-11	25	21	31	29	27
12	18	20	19	19	19
13-15	18	24	16	24	21
≥ 16	16	18	17	19	17
Average	20	22	24	25	23

* Data given as number of deaths per 1000 live births. Adapted from a report of the U.S. Department of Health, Education, and Welfare (14).

regarded as of no biological or medical significance. In the other direction, a dietary program that reduces weight by an average of 5 kg might be regarded as important to health, and yet this finding may not be well established, as expressed by statistical significance. Although the 5 kg is important, the data do not support a firm conclusion that a difference has actually been achieved.

Association is a usefully vague word to express a relation between two or more variables. *Correlation,* a more technical term, refers to a specific way to measure association, and should not be used in writing about statistical findings except in referring to that measure.

Sample usually refers to an observation or a collection of observations gathered in a well-defined way. To describe a sample as having been *drawn at random* means that a randomizing device has been used to make the choice, not that some haphazard event has created the sample, such as the use of an unstructured set of patient referrals to create the investigator's control group.

15. Define statistical terms, abbreviations, and most symbols.

Although many statistical terms such as mean, median, and standard deviation of the observations have clear, widely adopted definitions, different fields of endeavor often use the same symbols for different entities. Authors have extra difficulty when they need to distinguish between the true value of a quantity (a parameter such as a population mean, often symbolized by the Greek letter μ) and a sample mean (often written as $\bar{x}$).

We usually take for granted the mathematical symbols $=$, $+$, $-$, and $/$, as well as the usual symbols for inequalities (greater than or less than); we do the same for powers such as x^3, and for the trigonometric and logarithmic abbreviations such as sin, cos, tan, and log, although it is well to report what base the logarithms are using. Typography for ordinary multiplication differs, but is rarely a problem. Generally, symbols such as r for the correlation coefficient should be defined, as should n or N for the sample size, even though these are widely used.

Terms like *reliability* and *validity* are much more difficult, and they should always be defined when they are used in a statistical sense.

One difficulty with an expression such as $a \pm b$, even when a is a sample mean, is that b has many possibilities. (Some journals prefer the notation $a(b)$, but the ambiguities remain unchanged.) The author may use b for the observed sample standard deviation of individual measurements, or the standard error of the mean, or twice the standard error of the mean, or even the interquartile range, depending on the situation. The commonest ambiguity is not knowing whether b represents the standard deviation of individual observations or the standard error of the statistic designated by a. And no single choice is best in all situations. If the measure of variability is used only to test the size of its associated statistic, as for example in a P value to test whether a correlation coefficient differs from zero, then use the standard error. If the measure of variability needs to be combined with other such measures, the standard deviation of single observations is often more useful.

The same difficulty occurs with technical terms. A danger is that a special local language will become so ingrained in a particular research organization that its practitioners find it difficult to understand that their use of words is not widespread. Nearly every laboratory has special words that need to be defined or eliminated in reports of findings.

When one or two observations, terms, or symbols are not defined, readers may be able to struggle along. When several remain uncertain, readers may have to give up because the possibilities are too numerous.

A well-established convention is that mathematical symbols should be printed in italics (15-17). This practice has many advantages, including the reduction of ambiguity when the same character is commonly used to designate both a physical quantity and a mathematical or statistical quantity. In typescripts, an underline is generally used to indicate that a character is to be printed in italics, and authors may need to give special instructions to editors or printers if underlines are used for other purposes, such as to designate a mathematical vector (which might be printed both underlined and in italics).

ACKNOWLEDGMENTS: The authors thank Marcia Angell, Alexia Anctzak, M. Anthony Ashworth, John J. Bartko, Thomas Chalmers, Eli Chernin, Victor Cohn, David Hoaglin, Susan Horn, Edward J. Huth, Deborah A Lambert, Kathleen N. Lohr, Thomas Louis, Thomas R. O'Brien, Kenneth Rothman, Stephen B. Thacker, Anne A. Scitovsky, Barbara Starfield, Wallace K. Waterfall, Jonathan Weiner, Alfred Yankauer, Cleo Youtz, and two reviewers, for their thoughtful and helpful advice.

Grant support: in part by grant GAHS-8414 from the Rockefeller Foundation, grant S8406311 from the Josiah Macy Jr. Foundation, and the Methods Panel of the Institute of Medicine's Council for Health Care Technology.

▶ Requests for reprints should be directed to John C. Bailar III, M.D., Ph.D.; Department of Epidemiology and Biostatistics, McGill University School of Medicine, 1020 Pine Avenue West; Montreal, PQ, H3A 1A2, Canada.

References

1. INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS. Uniform requirements for manuscripts submitted to biomedical journals. *Ann Intern Med.* 1982;96(6 Pt 1):766-71.
2. INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS. Uniform requirements for manuscripts submitted to biomedical journals. *Ann Intern Med.* 1988;108:258-65.
3. DERSIMONIAN R, CHARETTE LJ, MCPEEK B, MOSTELLER F. Reporting on methods in clinical trials. *N Engl J Med.* 1982;306:1332-7.
4. EMERSON JD, COLDITZ GA. Use of statistical analysis in the *New England Journal of Medicine. N Engl J Med.* 1983;309:709-13.
5. SIMON R. Confidence intervals for reporting results of clinical trials. *Ann Intern Med.* 1986;105:429-35.
6. ROTHMAN KJ. Significance questing. *Ann Intern Med.* 1986;105:445-7.
7. LOUIS TA, FINEBERG H, MOSTELLER F. Findings for public health from meta-analyses. *Annu Rev Public Health.* 1985;6:1-20.
8. EMERSON JD, MOSES L. A note on the Wilcoxan-Mann-Whitney test for 2 × k ordered tables. *Biometrics.* 1985;41:303-9.
9. BAILAR JC, MOSTELLER F, eds. *Medical Uses of Statistics.* Waltham, Massachusetts: NEJM Books; 1986.
10. MOSTELLER F, GILBERT JP, MCPEEK B. Reporting standards and research strategies for controlled trials: agenda for the editor. *Controlled Clinical Trials.* 1980;1:37-58.
11. CLEVELAND WS. *The Elements of Graphing Data.* Monterey, California: Wadsworth Advanced Books and Software; 1985.
12. TUFTE ER. *The Visual Display of Quantitative Information.* Cheshire, Connecticut: Graphic Press; 1983.
13. EHRENBERG AC. The problem of numeracy. *Am Statistician.* 1981;35:67-71.
14. U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE. *Infant Mortality Rates: Socioeconomic Factors, United States.* Rockville, Maryland: National Center for Health Statistics; 1972; DHEW publication no. (HSM) 72-1045. (*Vital and Health Statistics;* series 22; no. 14).
15. HUTH EJ. Mathematics and statistics. In: HUTH EJ. *Medical Style &*

Format: An International Manual for Authors, Editors, and Publishers. Philadelphia: ISI Press; 1987:170-6.
16. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. *Units of Measurement.* 2nd ed. (ISO standards handbook 2). Geneva: International Organization for Standardization; 1982.
17. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. *Statistical Methods: Handbook on International Standards for Statistical Methods.* (ISO standards handbook 3). Geneva: International Organization for Standardization; 1979:287-8.

Additional References

The following references may be helpful to readers who want to pursue these topics further.

1. COCHRAN WG. *Sampling Techniques.* 3rd ed. New York: Wiley and Sons; 1977.
2. COHN V. *News and Numbers: A Guide to Reporting Statistical Claims and Controversies in Health and Other Fields.* Ames, Iowa: Iowa State University Press; 1988. (In press).
3. COLTON T. *Statistics in Medicine.* Boston: Little, Brown and Co.; 1974.
4. COMMITTEE FOR EVALUATING MEDICAL TECHNOLOGIES IN CLINICAL USE. *Assessing Medical Technologies.* Washington, DC: National Academy Press; 1985.
5. FLEISS JL. *Statistical Methods for Rates and Proportions.* 2nd ed. New York: John Wiley and Sons; 1981.
6. GARDNER MJ, MACLURE M, CAMPBELL MJ. Use of check lists in assessing the statistical content of medical studies. *Br Med J.* 1986;292:810-2.
7. GORE S, ALTMAN DG. *Statistics in Practice.* London: Taylor and Francis; 1982.
8. INGELFINGER JA, MOSTELLER F, THIBODEAU LA, WARE JH. *Biostatistics in Clinical Medicine.* 2nd ed. New York: Macmillan; 1987.
9. MEINERT CL. *Clinical Trials: Designs, Conduct, and Analysis.* New York: Oxford University Press; 1986.
10. MIKÉ V, STANLEY KE, eds. *Statistics in Medical Research: Methods and Issues, with Applications in Cancer Research.* New York: John Wiley and Sons; 1982.
11. MOSES LE, MOSTELLER F, eds. *Planning and Analysis of Observational Studies.* New York: Wiley and Sons; 1983.
12. SHAPIRO SH, LOUIS TA, eds. *Clinical Trials: Issues and Approaches.* New York: Marcel Dekker; 1983.
13. SACKETT DL, HAYNES RB, TUGWELL P. *Clinical Epidemiology: A Basic Science for Clinical Medicine.* Boston: Little, Brown & Co.; 1985.
14. SNEDECOR GW, COCHRAN WG. *Statistical Methods.* 7th ed. Ames, Iowa: Iowa State University Press; 1980.

SPECIAL REPORT

UNIFORM REQUIREMENTS FOR MANUSCRIPTS SUBMITTED TO BIOMEDICAL JOURNALS

INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS*

In the 12 years since it was first published, the "Uniform Requirements for Manuscripts Submitted to Biomedical Journals" (the Vancouver style), developed by the International Committee of Medical Journal Editors, has been widely accepted by both authors and editors; over 400 journals have stated that they will consider manuscripts that conform to its requirements. This is the fourth edition of the Uniform Requirements, the first to be published in the Journal, *which now serves as coordinator of the ICMJE in North America.*

IN January 1978 a group of editors from some major biomedical journals published in English met in Vancouver, British Columbia, and decided on uniform technical requirements for manuscripts to be submitted to their journals. These requirements, including formats for bibliographic references developed for the Vancouver group by the National Library of Medicine, were published in three of the journals early in 1979. The Vancouver group evolved into the International Committee of Medical Journal Editors. Over the years, the group has revised the requirements slightly; this is the fourth edition.

Over 400 journals have agreed to receive manuscripts prepared in accordance with the requirements. It is important to emphasize what these requirements imply and what they do not.

First, the requirements are instructions to authors on how to prepare manuscripts, not to editors on publication style. (But many journals have drawn on these requirements for elements of their publication styles.)

Second, if authors prepare their manuscripts in the style specified in these requirements, editors of the participating journals will not return manuscripts for changes in these details of style. Even so, manuscripts may be altered by journals to conform with details of their own publication styles.

Third, authors sending manuscripts to a participating journal should not try to prepare them in accordance with the publication style of that journal but should follow the "Uniform Requirements for Manuscripts Submitted to Biomedical Journals."

Nevertheless, authors must also follow the instructions to authors in the journal as to what topics are suitable for that journal and the types of papers that may be submitted — for example, original articles, reviews, or case reports. In addition, the journal's instructions are likely to contain other requirements unique to that journal, such as number of copies of manuscripts, acceptable languages, length of articles, and approved abbreviations.

Participating journals are expected to state in their instructions to authors that their requirements are in accordance with the "Uniform Requirements for Manuscripts Submitted to Biomedical Journals" and to cite a published version.

This document will be revised at intervals. Inquiries and comments from Central and North America about these requirements should be sent to Editor, the *New England Journal of Medicine,* 10 Shattuck St., Boston, MA 02115; those from other regions should be sent to Editor, *British Medical Journal,* British Medical Association, Tavistock Sq., London WC1H 9JR, United Kingdom. Note that these two journals provide secretariat services for the International Committee of Medical Journal Editors; they do not handle manuscripts intended for other journals. Papers intended for other journals should be sent directly to the offices of those journals.

SUMMARY OF REQUIREMENTS

Type the manuscript double-spaced, including title page, abstract, text, acknowledgments, references, tables, and legends.

Each manuscript component should begin on a new page, in the following sequence: title page; abstract and key words; text; acknowledgments; references; tables (each table complete with title and footnotes on a separate page); and legends for illustrations.

Illustrations must be good-quality, unmounted glossy prints, usually 127 × 173 mm (5 × 7 in.), but no larger than 203 × 254 mm (8 × 10 in.).

Submit the required number of copies of manuscript and figures (see journal's instructions) in a heavy paper envelope. The submitted manuscript should be accompanied by a covering letter, as described under Submission of Manuscripts, and permissions to reproduce previously published material or to use illustrations that may identify human subjects.

Follow the journal's instructions for transfer of copyright. Authors should keep copies of everything submitted.

PRIOR AND DUPLICATE PUBLICATION

Most journals do not wish to consider for publication a paper on work that has already been reported in a published paper or is described in a paper submitted or accepted for publication elsewhere. This policy does not usually preclude consideration of a paper that has been rejected by another journal or of a complete report that follows publication of a preliminary report, usually in the form of an abstract. Nor does it prevent consideration of a paper that has been presented at a scientific meeting if not published in full in a proceedings or similar publication. Press reports of the meeting will not usually be considered as breaches of this rule, but such reports should not be amplified by additional data or copies of tables and illustrations. When submitting a paper an author should always make a full statement to the editor about all submissions and previous reports that might be regarded as prior or duplicate publication of the same or very similar work. Copies of such material should be included with the submitted paper to help the editor decide how to deal with the matter.

Multiple publication — that is, the publication more than once of

*Members of the committee are Suzanne and Robert Fletcher (*Annals of Internal Medicine*), Laurel Thomas (*Medical Journal of Australia*), Stephen Lock (*British Medical Journal*), George D. Lundberg (*Journal of the American Medical Association*), Robin Fox (*Lancet*), Magne Nylenna (*Tidsskrift for den Norske Laegeforening*), Lois Ann Colaianni (*Index Medicus*), Arnold S. Relman and Marcia Angell (*New England Journal of Medicine*), Povl Riis (*Journal of the Danish Medical Association, Danish Medical Bulletin*), Richard G. Robinson (*New Zealand Medical Journal*), Bruce P. Squires (*Canadian Medical Association Journal*), and Linda Clever (*Western Journal of Medicine*). Address correspondence to Editor, the *New England Journal of Medicine,* or Editor, *British Medical Journal.*

the same study, irrespective of whether the wording is the same — is rarely justified. Secondary publication in another language is one possible justification, provided the following conditions are met.

(1) The editors of both journals concerned are fully informed; the editor concerned with secondary publication should have a photocopy, reprint, or manuscript of the primary version.
(2) The priority of the primary publication is respected by a publication interval of at least two weeks.
(3) The paper for secondary publication is written for a different group of readers and is not simply a translated version of the primary paper; an abbreviated version will often be sufficient.
(4) The secondary version reflects faithfully the data and interpretations of the primary version.
(5) A footnote on the title page of the secondary version informs readers, peers, and documenting agencies that the paper was edited, and is being published, for a national audience in parallel with a primary version based on the same data and interpretations. A suitable footnote might read as follows: "This article is based on a study first reported in the [title of journal, with full reference]."

Multiple publication other than as defined above is not acceptable to editors. If authors violate this rule they may expect appropriate editorial action to be taken.

Preliminary release, usually to public media, of scientific information described in a paper that has been accepted but not yet published is a violation of the policies of many journals. In a few cases, and only by arrangement with the editor, preliminary release of data may be acceptable — for example, to warn the public of health hazards.

Preparation of Manuscript

Type the manuscript on white bond paper, 216 × 279 mm (8½ × 11 in.) or ISO A4 (212 × 297 mm), with margins of at least 25 mm (1 in.). Type only on one side of the paper. Use double-spacing throughout, including title page, abstract, text, acknowledgments, references, tables, and legends for illustrations. Begin each of the following sections on separate pages: title page, abstract and key words, text, acknowledgments, references, individual tables, and legends. Number pages consecutively, beginning with the title page. Type the page number in the upper or lower right-hand corner of each page.

Title Page

The title page should carry (a) the title of the article, which should be concise but informative; (b) first name, middle initial, and last name of each author, with highest academic degree(s) and institutional affiliation; (c) name of department(s) and institution(s) to which the work should be attributed; (d) disclaimers, if any; (e) name and address of author responsible for correspondence about the manuscript; (f) name and address of author to whom requests for reprints should be addressed or statement that reprints will not be available from the author; (g) source(s) of support in the form of grants, equipment, drugs, or all of these; and (h) a short running head or foot line of no more than 40 characters (count letters and spaces) placed at the foot of the title page and identified.

Authorship

All persons designated as authors should qualify for authorship. The order of authorship should be a joint decision of the coauthors. Each author should have participated sufficiently in the work to take public responsibility for the content.

Authorship credit should be based only on substantial contributions to (a) conception and design, or analysis and interpretation of data; and to (b) drafting the article or revising it critically for important intellectual content; and on (c) final approval of the version to be published. Conditions (a), (b), and (c) must all be met. Participation solely in the acquisition of funding or the collection of data does not justify authorship. General supervision of the research group is also not sufficient for authorship. Any part of an article critical to its main conclusions must be the responsibility of at least one author.

A paper with corporate (collective) authorship must specify the key persons responsible for the article; others contributing to the work should be recognized separately (see Acknowledgments).

Editors may require authors to justify the assignment of authorship.

Abstract and Key Words

The second page should carry an abstract (of no more than 150 words for unstructured abstracts or 250 words for structured abstracts). The abstract should state the purposes of the study or investigation, basic procedures (selection of study subjects or laboratory animals; observational and analytical methods), main findings (give specific data and their statistical significance, if possible), and the principal conclusions. Emphasize new and important aspects of the study or observations.

Below the abstract provide, and identify as such, 3 to 10 key words or short phrases that will assist indexers in cross-indexing the article and may be published with the abstract. Use terms from the medical subject headings (MeSH) list of *Index Medicus*; if suitable MeSH terms are not yet available for recently introduced terms, present terms may be used.

Text

The text of observational and experimental articles is usually — but not necessarily — divided into sections with the headings Introduction, Methods, Results, and Discussion. Long articles may need subheadings within some sections to clarify their content, especially the Results and Discussion sections. Other types of articles such as case reports, reviews, and editorials are likely to need other formats. Authors should consult individual journals for further guidance.

Introduction

State the purpose of the article. Summarize the rationale for the study or observation. Give only strictly pertinent references, and do not review the subject extensively. Do not include data or conclusions from the work being reported.

Methods

Describe your selection of the observational or experimental subjects (patients or laboratory animals, including controls) clearly. Identify the methods, apparatus (manufacturer's name and address in parentheses), and procedures in sufficient detail to allow other workers to reproduce the results. Give references to established methods, including statistical methods (see below); provide references and brief descriptions for methods that have been published but are not well known; describe new or substantially modified methods, give reasons for using them, and evaluate their limitations. Identify precisely all drugs and chemicals used, including generic name(s), dose(s), and route(s) of administration.

Ethics

When reporting experiments on human subjects indicate whether the procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional or regional) or with the Helsinki Declaration of 1975, as revised in 1983. Do not use patients' names, initials, or hospital numbers, especially in any illustrative material. When reporting experiments on animals indicate whether the institution's or the National Research Council's guide for, or any national law on, the care and use of laboratory animals was followed.

Statistics

Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results. When possible, quantify findings and present them with appropriate indicators of measurement error or uncertainty (such as confidence intervals). Avoid sole reliance on statistical hypothesis testing, such as the use of P values, which fails to convey important quantitative information. Discuss eligibility of experi-

mental subjects. Give details about randomization. Describe the methods for and success of any blinding of observations. Report treatment complications. Give numbers of observations. Report losses to observation (such as dropouts from a clinical trial). References for study design and statistical methods should be to standard works (with pages stated) when possible rather than to papers in which the designs or methods were originally reported. Specify any general-use computer programs used.

Put general descriptions of methods in the Methods section. When data are summarized in the Results section specify the statistical methods used to analyze them. Restrict tables and figures to those needed to explain the argument of the paper and to assess its support. Use graphs as an alternative to tables with many entries; do not duplicate data in graphs and tables. Avoid nontechnical uses of technical terms in statistics, such as "random" (which implies a randomizing device), "normal," "significant," "correlations," and "sample." Define statistical terms, abbreviations, and most symbols.

Results

Present your results in logical sequence in the text, tables, and illustrations. Do not repeat in the text all the data in the tables or illustrations; emphasize or summarize only important observations.

Discussion

Emphasize the new and important aspects of the study and the conclusions that follow from them. Do not repeat in detail data or other material given in the Introduction or the Results section. Include in the Discussion section the implications of the findings and their limitations, including implications for future research. Relate the observations to other relevant studies. Link the conclusions with the goals of the study but avoid unqualified statements and conclusions not completely supported by your data. Avoid claiming priority and alluding to work that has not been completed. State new hypotheses when warranted, but clearly label them as such. Recommendations, when appropriate, may be included.

Acknowledgments

At an appropriate place in the article (title-page footnote or appendix to the text; see the journal's requirement) one or more statements should specify (a) contributions that need acknowledging but do not justify authorship, such as general support by a departmental chairman; (b) acknowledgments of technical help; (c) acknowledgments of financial and material support, specifying the nature of the support; (d) financial relationships that may pose a conflict of interest.

Persons who have contributed intellectually to the paper but whose contributions do not justify authorship may be named and their function or contribution described — for example, "scientific adviser," "critical review of study proposal," "data collection," or "participation in clinical trial." Such persons must have given their permission to be named. Authors are responsible for obtaining written permission from persons acknowledged by name, because readers may infer their endorsement of the data and conclusions.

Technical help should be acknowledged in a paragraph separate from those acknowledging other contributions.

References

Number references consecutively in the order in which they are first mentioned in the text. Identify references in text, tables, and legends by Arabic numerals in parentheses. References cited only in tables or in legends to figures should be numbered in accordance with a sequence established by the first identification in the text of the particular table or illustration.

Use the style of the examples below, which are based with slight modifications on the formats used by the U.S. National Library of Medicine in *Index Medicus*. The titles of journals should be abbreviated according to the style used in *Index Medicus*. Consult *List of Journals Indexed in Index Medicus*, published annually as a separate publication by the library and as a list in the January issue of *Index Medicus*.

Try to avoid using abstracts as references; "unpublished observations" and "personal communications" may not be used as references, although references to written, not oral, communications may be inserted (in parentheses) in the text. Include among the references papers accepted but not yet published; designate the journal and add "In press." Information from manuscripts submitted but not yet accepted should be cited in the text as "unpublished observations" (in parentheses).

The references must be verified by the author(s) against the original documents.

Examples of correct forms of references are given below.

Articles in Journals

(1) *Standard journal article* (List all authors, but if the number exceeds six give six followed by et al.)

You CH, Lee KY, Chey RY, Menguy R. Electrogastrographic study of patients with unexplained nausea, bloating and vomiting. Gastroenterology 1980 Aug; 79(2):311-4.

As an option, if a journal carries continuous pagination throughout a volume, the month and issue number may be omitted.

You CH, Lee KY, Chey RY, Menguy R. Electrogastrographic study of patients with unexplained nausea, bloating and vomiting. Gastroenterology 1980;79:311-4.

Goate AM, Haynes AR, Owen MJ, Farrall M, James LA, Lai LY, et al. Predisposing locus for Alzheimer's disease on chromosome 21. Lancet 1989;1: 352-5.

(2) *Organization as author*

The Royal Marsden Hospital Bone-Marrow Transplantation Team. Failure of syngeneic bone-marrow graft without preconditioning in post-hepatitis marrow aplasia. Lancet 1977;2:742-4.

(3) *No author given*

Coffee drinking and cancer of the pancreas [editorial]. BMJ 1981;283:628.

(4) *Article in a foreign language*

Massone L, Borghi S, Pestarino A, Piccini R, Gambini C. Localisations palmaires purpuriques de la dermatite herpetiforme. Ann Dermatol Venereol 1987;114:1545-7.

(5) *Volume with supplement*

Magni F, Rossoni G, Berti F. BN-52021 protects guinea-pig from heart anaphylaxis. Pharmacol Res Commun 1988;20 Suppl 5:75-8.

(6) *Issue with supplement*

Gardos G, Cole JO, Haskell D, Marby D, Paine SS, Moore P. The natural history of tardive dyskinesia. J Clin Psychopharmacol 1988;8(4 Suppl):31S-37S.

(7) *Volume with part*

Hanly C. Metaphysics and innateness: a psychoanalytic perspective. Int J Psychoanal 1988;69(Pt 3):389-99.

(8) *Issue with part*

Edwards L, Meyskens F, Levine N. Effect of oral isotretinoin on dysplastic nevi. J Am Acad Dermatol 1989;20(2 Pt 1):257-60.

(9) *Issue with no volume*

Baumeister AA. Origins and control of stereotyped movements. Monogr Am Assoc Ment Defic 1978;(3):353-84.

(10) *No issue or volume*

Danoek K. Skiing in and through the history of medicine. Nord Medicinhist Arsb 1982:86-100.

(11) *Pagination in Roman numerals*

Ronne Y. Ansvarsfall. Blodtransfusion till fel patient. Vardfacket 1989;13: XXVI-XXVII.

(12) *Type of article indicated as needed*

Spargo PM, Manners JM. DDAVP and open heart surgery [letter]. Anaesthesia 1989;44:363-4.

Fuhrman SA, Joiner KA. Binding of the third component of complement C3 by Toxoplasma gondii [abstract]. Clin Res 1987;35:475A.

(13) *Article containing retraction*

Shishido A. Retraction notice: Effect of platinum compounds on murine lymphocyte mitogenesis (Retraction of Alsabti EA, Ghalib ON, Salem MH. In: Jpn J Med Sci Biol 1979;32:53-65). Jpn J Med Sci Biol 1980;33:235-7.

(14) *Article retracted*

Alsabti EA, Ghalib ON, Salem MH. Effect of platinum compounds on murine lymphocyte mitogenesis [Retracted by Shishido A. In: Jpn J Med Sci Biol 1980; 33:235-7]. Jpn J Med Sci Biol 1979;32:53-65.

(15) *Article containing comment*

Piccoli A, Bossatti A. Early steroid therapy in IgA neuropathy: still an open question [comment]. Nephron 1989;51:289-91. Comment on: Nephron 1988; 48:12-7.

(16) *Article commented on*

Kobayashi Y, Fujii K, Hiki Y, Tateno S, Kurokawa A, Kamiyama M. Steroid therapy in IgA nephropathy: a retrospective study in heavy proteinuric cases [see comments]. Nephron 1988;48:12-7. Comment in: Nephron 1989;51:289-91.

(17) *Article with published erratum*

Schofield A. The CAGE questionnaire and psychological health [published erratum appears in Br J Addict 1989;84:701]. Br J Addict 1988;83:761-4.

Books and Other Monographs

(18) *Personal author(s)*

Colson JH, Armour WJ. Sports injuries and their treatment. 2nd rev. ed. London: S. Paul, 1986.

(19) *Editor(s), compiler as author*

Diener HC, Wilkinson M, editors. Drug-induced headache. New York: Springer-Verlag, 1988.

(20) *Organization as author and publisher*

Virginia Law Foundation. The medical and legal implications of AIDS. Charlottesville: The Foundation, 1987.

(21) *Chapters in a book*

Weinstein L, Swartz MN. Pathologic properties of invading microorganisms. In: Sodeman WA Jr, Sodeman WA, editors. Pathologic physiology: mechanisms of disease. Philadelphia: Saunders, 1974: 457-72.

(22) *Conference proceedings*

Vivian VL, editor. Child abuse and neglect: a medical community response. Proceedings of the First AMA National Conference on Child Abuse and Neglect; 1984 Mar 30-31; Chicago. Chicago: American Medical Association, 1985.

(23) *Conference paper*

Harley NH. Comparing radon daughter dosimetric and risk models. In: Gammage RB, Kaye SV, editors. Indoor air and human health. Proceedings of the Seventh Life Sciences Symposium; 1984 Oct 29-31; Knoxville (TN). Chelsea (MI): Lewis, 1985: 69-78.

(24) *Scientific and technical report*

Akutsu T. Total heart replacement device. Bethesda (MD): National Institutes of Health, National Heart and Lung Institute; 1974 Apr. Report No.: NIH-NHLI-69 2185-4.

(25) *Dissertation*

Youssef NM. School adjustment of children with congenital heart disease [dissertation]. Pittsburgh (PA): Univ. of Pittsburgh, 1988.

(26) *Patent*

Harred JF, Knight AR, McIntyre JS, inventors. Dow Chemical Company, assignee. Epoxidation process. US patent 3,654,317. 1972 Apr 4.

Other Published Material

(27) *Newspaper article*

Rensberger B, Specter B. CFCs may be destroyed by natural process. The Washington Post 1989 Aug 7;Sect. A:2 (col. 5).

(28) *Audiovisual*

AIDS epidemic: the physician's role [videorecording]. Cleveland (OH): Academy of Medicine of Cleveland, 1987.

(29) *Computer file*

Renal system [computer program]. MS-DOS version. Edwardsville (KS): Medi-Sim, 1988.

(30) *Legal material*

Toxic Substances Control Act: Hearing on S. 776 Before the Subcomm. on the Environment of the Senate Comm. on Commerce, 94th Cong., 1st Sess. 343 (1975).

(31) *Map*

Scotland [topographic map]. Washington: National Geographic Society (US), 1981.

(32) *Book of the Bible*

Ruth 3:1-18. The Holy Bible. Authorized King James version. New York: Oxford Univ. Press, 1972.

(33) *Dictionary and similar references*

Ectasia. Dorland's illustrated medical dictionary. 27th ed. Philadelphia: Saunders, 1988: 527.

(34) *Classical material*

The Winter's Tale: act 5, scene 1, lines 13-16. The complete works of William Shakespeare. London: Rex, 1973.

Unpublished Material

(35) *In press*

Lillywhite HB, Donald JA. Pulmonary blood flow regulation in an aquatic snake. Science. In press.

Tables

Type each table double-spaced on a separate sheet. Do not submit tables as photographs. Number tables consecutively in the order of their first citation in the text and supply a brief title for each. Give each column a short or abbreviated heading. Place explanatory matter in footnotes, not in the heading. Explain in footnotes all nonstandard abbreviations that are used in each table. For footnotes use the following symbols, in this sequence: *, †, ‡, §, ‖, ¶, **, ††, ...

Identify statistical measures of variations such as standard deviation and standard error of the mean.

Do not use internal horizontal and vertical rules.

Be sure that each table is cited in the text.

If you use data from another published or unpublished source obtain permission and acknowledge fully.

The use of too many tables in relation to the length of the text may produce difficulties in the layout of pages. Examine issues of the journal to which you plan to submit your paper to estimate how many tables can be used per 1000 words of text.

The editor, on accepting a paper, may recommend that additional tables containing important backup data too extensive to publish be deposited with an archival service, such as the National Auxiliary Publication Service in the United States, or made available by the authors. In that event an appropriate statement will be added to the text. Submit such tables for consideration with the paper.

Illustrations

Submit the required number of complete sets of figures. Figures should be professionally drawn and photographed; freehand or typewritten lettering is unacceptable. Instead of original drawings, roentgenograms, and other material send sharp, glossy black-and-white photographic prints, usually 127 × 173 mm (5 × 7 in.), but no larger than 203 × 254 mm (8 × 10 in.). Letters, numbers, and symbols should be clear and even throughout and of sufficient size that when reduced for publication each item will still be legible. Titles and detailed explanations belong in the legends for illustrations, not on the illustrations themselves.

Each figure should have a label pasted on its back indicating the number of the figure, author's name, and top of the figure. Do not write on the back of figures or scratch or mar them by using paper clips. Do not bend figures or mount them on cardboard.

Photomicrographs must have internal scale markers. Symbols, arrows, or letters used in the photomicrographs should contrast with the background.

If photographs of persons are used, either the subjects must not be identifiable or their pictures must be accompanied by written permission to use the photograph.

Figures should be numbered consecutively according to the order in which they have been first cited in the text. If a figure has been published acknowledge the original source and submit written permission from the copyright holder to reproduce the material. Permission is required irrespective of authorship or publisher, except for documents in the public domain.

For illustrations in color, ascertain whether the journal requires color negatives, positive transparencies, or color prints. Accompanying drawings marked to indicate the region to be reproduced may be useful to the editor. Some journals publish illustrations in color only if the author pays for the extra cost.

Legends for Illustrations

Type legends for illustrations double-spaced, starting on a separate page, with Arabic numerals corresponding to the illustrations. When symbols, arrows, numbers, or letters are used to identify parts of the illustrations, identify and explain each one clearly in the legend. Explain the internal scale and identify method of staining in photomicrographs.

Units of Measurement

Measurements of length, height, weight, and volume should be reported in metric units (meter, kilogram, or liter) or their decimal multiples.

Temperatures should be given in degrees Celsius. Blood pressures should be given in millimeters of mercury.

All hematologic and clinical-chemistry measurements should be reported in the metric system in terms of the International System of Units (SI). Editors may request that alternative or non-SI units be added by the authors before publication.

Abbreviations and Symbols

Use only standard abbreviations. Avoid abbreviations in the title and abstract. The full term for which an abbreviation stands should precede its first use in the text unless it is a standard unit of measurement.

Submission of Manuscripts

Mail the required number of manuscript copies in a heavy paper envelope, enclosing the manuscript copies and figures in cardboard, if necessary, to prevent bending of photographs during mail handling. Place photographs and transparencies in a separate heavy paper envelope.

Manuscripts must be accompanied by a covering letter signed by all coauthors. This must include (a) information on prior or duplicate publication or submission elsewhere of any part of the work as defined earlier in this document; (b) a statement of financial or other relationships that might lead to a conflict of interest; (c) a statement that the manuscript has been read and approved by all authors, that the requirements for authorship as previously stated in this document have been met, and furthermore, that each coauthor believes that the manuscript represents honest work; and (d) the name, address, and telephone number of the corresponding author, who is responsible for communicating with the other authors about revisions and final approval of the proofs. The letter should give any additional information that may be helpful to the editor, such as the type of article in the particular journal the manuscript represents and whether the author(s) will be willing to meet the cost of reproducing color illustrations.

The manuscript must be accompanied by copies of any permissions to reproduce published material, to use illustrations or report sensitive personal information of identifiable persons, or to name persons for their contributions.

Participating Journals

Journals that have notified the International Committee of Medical Journal Editors of their willingness to consider for publication manuscripts prepared in accordance with earlier versions of the committee's uniform requirements identify themselves as such in their information for authors. A full list is available on request from the *New England Journal of Medicine* or the *British Medical Journal*. Citations of this document should be to one of the sources listed below.

International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. N Engl J Med 1991; 324:424-8.

International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. BMJ 1991 Feb 9;302(6772).

Sample size calculations for clinical pharmacology studies

Paul D. Stolley, M.D., M.P.H., and Brian L. Strom, M.D., M.P.H. *Philadelphia, Pa.*

Numerous manuscripts are submitted to CLINICAL PHARMACOLOGY AND THERAPEUTICS that report on studies that are hopelessly too small—which is a common reason manuscripts are rejected for publication. An old joke implied that a study should have either 10 or 100 subjects. Why? Because "both numbers are round and divisible by 10 and don't leave you with those messy fractions." The point of the joke is that there is a better way to decide on the appropriate study size, so that the study is neither so small that it cannot reliably answer the questions nor so large that it wastes resources. Such sample size calculations are critically important in the design of a study.

Sample size determination for a given experiment requires that the investigator specify four parameters in order to calculate a fifth, i.e., n, or sample size. The approach is slightly different depending on whether the outcome variable is: (1) a continuous variable, a number (e.g., serum concentration of a drug) for which one usually studies the mean value of that outcome variable in a group of patients; or (2), a dichotomous variable, yes or no (e.g., a patient does or does not have a particular adverse reaction), for which one usually studies the proportion of patients who have the outcome variable. Inasmuch as the former situation is likely to be of interest to readers of this JOURNAL, it will be discussed first, followed by the latter situation.

CONTINUOUS OUTCOME VARIABLES

First, the investigator selects an alpha (α) level, specifying the level of type I error he or she is willing to tolerate. A type I error is the probability of concluding, because of chance alone, that a difference exists when in fact there is no real difference between the study groups. This is conventionally selected to be 0.05, although it can be larger or smaller.

Second, the investigator also selects a beta (β) level, specifying the level of type II error he or she is willing to tolerate. A type II error is the probability of concluding, because of chance alone, that a difference does not exist when in fact there is a real difference between the study groups. This is closely related to the concept of power, which is the probability of detecting a difference between two study groups if one truly exists. Mathematically, power = $(1 - \beta)$. The value for β is conventionally selected to be 0.1 or 0.2, although it can be larger or smaller. Notice that it is conventional to have a less strict criterion for a type II error than for a type I error; in most scientific investigations it is thought more important to avoid claiming a difference that does not truly exist than to miss a true difference. This does not have to be the case, however, depending on the purposes of the study.

The third parameter the investigator must specify is the difference between the groups to be detected in the experiment (Δ), i.e., the difference in means thought biologically important. If the investigator will settle for detecting only a large difference, then fewer subjects will be needed. If the difference between study groups sought is small, however, then a larger study group will be needed. It is axiomatic that for large differences a smaller number of subjects will suffice; small differences require large numbers of subjects. It is important to realize, however, that we are not specifying here the difference expected, but the smallest difference we would like to be able to detect.

Finally, the fourth parameter that must be specified is the standard deviation of the outcome variable used in the study (σ). A very precise method of measurement (small σ) will permit detection of any given difference with a much smaller sample size than would be required with a less precise measurement.

Once one has specified the above parameters, cal-

From the Clinical Epidemiology Unit, Section of General Medicine, Department of Medicine, University of Pennsylvania School of Medicine.

Supported in part by Grant No. FD-U-000079 from the United States Food and Drug Administration and by grants from The Rockefeller Foundation and The Andrew W. Mellon Foundation.

Received for publication Jan. 26, 1986; accepted Jan. 26, 1986.

Reprint requests to: Dr. Paul D. Stolley, M.D., M.P.H., 229L NEB/S2, University of Pennsylvania School of Medicine, Philadelphia, PA 19104.

culation of the required sample size is a simple matter of inserting the values into the following equation:

$$n = \frac{2(Z_\alpha + Z_\beta)^2\sigma^2}{\Delta^2}$$

where n is the sample size needed in the experimental group, Z_α is the Z value corresponding to the two-tailed α ($Z_\alpha = 1.645$ for a two-tailed $\alpha = 0.10$ or a one-tailed $\alpha = 0.05$; $Z_\alpha = 1.96$ for a two-tailed $\alpha = 0.05$ or a one-tailed $\alpha = 0.025$; and $Z_\alpha = 2.576$ for a two-tailed $\alpha = 0.01$ or a one-tailed $\alpha = 0.005$), Z_β is the Z value corresponding to the one-tailed β ($Z_\beta = 0.842$ for $\beta = 0.2$, $Z_\beta = 1.282$ for $\beta = 0.1$), σ is the standard deviation of the control sample, and Δ is the smallest difference between the two study groups thought important to detect. This assumes a study with one control subject per study subject.

DICHOTOMOUS OUTCOME VARIABLES

When the outcome variable is dichotomous (e.g., a patient either improves or does not improve), the outcome variable of interest is the proportion of people who have the disease, rather than the mean of a specified measurement. To calculate the required sample size, one focuses on the difference in proportions one would like to be able to detect (Δ, as above). Also, in this situation one can take advantage of the fact that the standard deviation can be expressed mathematically in terms of the proportion of subjects with the outcome variable in the control group. Therefore, for a cohort (prospective) study one can calculate sample size by specifying α, β, the smallest proportion developing the disease in the exposed study group that one considers important to detect (p_1), and the proportion expected to develop the disease in the unexposed control group (p_2), using the formula:

$$n = \frac{\left[Z_\alpha\sqrt{2p(1-p)} + Z_\beta\sqrt{p_1(1-p_1) + p_2(1-p_2)}\right]^2}{(p_1 - p_2)^2}$$

where n, Z_α, and Z_β are as above and $p = (p_1 + p_2)/2$. Again, this assumes one control subject per study subject. Note that p_1/p_2 is the relative risk, i.e., the incidence rate in the exposed group divided by the incidence rate in the control group. A relative risk >1.0 indicates that the exposure appears to increase the risk of the outcome. A relative risk <1.0 indicates that the exposure appears to decrease the risk of the outcome. A relative risk of 1.0 indicates that there is no association between the exposure and the outcome.

For a case-control (retrospective) study, one can calculate the sample size using the same formula, replacing p_1 with p_3, which represents the smallest proportion exposed to the risk factor of interest that one would consider important to detect in the diseased (case) group, and replacing p_2 with p_4, which represents the proportion expected to experience the exposure of interest in the undiseased (control) group.

More complicated designs generally require statistical or epidemiologic consultation.

When an investigator finally completes and publishes an investigation, if no difference is detected it is useful to give the reader some notion of the power of the study, i.e., how big a difference could have been detected in the study. A sentence such as "we found no difference between the two treatments and could have detected an X% difference with a power of Y" is useful information. The above formulas can also be used for this purpose.

Finally, if one reads an article and wants quickly to estimate the power of the study to detect a difference one believes important to detect, a recent article presents nomograms that provide a simpler, albeit less precise, alternative approach, as well as another discussion of some of the issues covered in this commentary.[1] Other useful references for readers interested in a more in-depth discussion of this material are provided as references 2 and 3.

References

1. Young MJ, Bresnitz EA, Strom BL. Sample size nomograms for interpreting negative clinical studies. Ann Intern Med 1983;99:248-51.
2. Schlesselman JJ. Sample size requirements in cohort and case-control studies of disease. Am J Epidemiol 1974; 99:381-4.
3. Cohen J. Statistical power analysis for the behavioral sciences. New York: Academic Press, 1977.

References

1. **Cryer B, Feldman M.** Effects of nonsteroidal anti-inflammatory drugs on endogenous gastrointestinal prostaglandins and therapeutic strategies for prevention and treatment of nonsteroidal anti-inflammatory drug-induced damage. Arch Intern Med. 1992;152:1145-55.
2. **Allison MC, Howatson AG, Torrance CJ, Lee FD, Russell RI.** Gastrointestinal damage associated with the use of nonsteroidal antiinflammatory drugs. N Engl J Med. 1992;327:749-54.
3. Risk of GI ulceration, bleeding and perforation with nonsteroidal anti-inflammatory therapy. In: Physicians Desk Reference. 47th ed. Montvale, New Jersey: Medical Economics Data; 1993:2352.
4. **Cryer B, Goldschmiedt M, Redfern JS, Feldman M.** Comparison of salsalate and aspirin on mucosal injury and gastroduodenal mucosal prostaglandins. Gastroenterology. 1990;99:1616-21.
5. **Brooks PM, Day RO.** Nonsteroidal antiinflammatory drugs—differences and similarities. N Engl J Med. 1991;324:1716-25.
6. **Fuster V, Dyken ML, Vokonas PS, Hennekens C.** Aspirin as a therapeutic agent in cardiovascular disease. Circulation. 1993;87: 659-75.
7. **Thun MJ, Namboodiri MM, Heath CW.** Aspirin use and reduced risk of fatal colon cancer. N Engl J Med. 1991;325:1593-1600.
8. **Graham DY, Agrawal NM, Roth SH.** Prevention of NSAID-induced gastric ulcer with misoprostol: multicentre, double-blind, placebo-controlled trial. Lancet. 1988;ii;277-80.
9. **Agrawal NM, Roth S, Graham DY, White RH, Germain B, Brown JA, et al.** Misoprostol compared with sucralfate in the prevention of nonsteroidal antiinflammatory drug-induced gastric ulcer. A randomized, controlled trial. Ann Intern Med. 1991;115:195-200.
10. **Robinson MG Jr, Griffin JW, Bowers J, Kagan DG, Kugut DO, et al.** Effect of ranitidine on gastroduodenal mucosal damage induced by nonsteroidal antiinflammatory drugs. Dig Dis Sci. 1989;34:424-8.
11. **Ehsanullah RS, Page MC, Tidesley G, Wood JR.** Prevention of gastroduodenal damage induced by nonsteroidal antiinflammatory drugs: controlled trial of ranitidine. BMJ. 1988;297:1017-21.
12. **Walt RP.** Misoprostol for the treatment of peptic ulcer and antiinflammatory drug-induced gastroduodenal ulceration. N Engl J Med. 1992;327:1575-80.
13. **Graham DY, White RH, Moreland LW, Schubert TT, Jaszewski R, Tindall E, et al.** Duodenal and gastric ulcer.prevention.with misoprostol in arthritis patients taking NSAIDs. Ann Intern Med. 1993; 119:257-262.
14. **Raskin JB, White R, Jaszewski R, Teoh LS, Sue SO.** Misoprostol and ranitidine in the prevention of NSAID-induced gastric and duodenal ulcer disease: a multicenter trial [Abstract]. Gastroenterology. 1992;102:A151.
15. **Davies GR, Wilkie ME, Rampton DS.** Effects of metronidazole and misoprostol on indomethacin-induced changes in intestinal permeability. Dig Dis Sci. 1993;38:417-25.
16. **Davies GR, Perry JD, Rampton DS.** Treatment of intestinal damage associated with long-term NSAID use: a six week double-blind, controlled trial of misoprostol. Gastroenterology. 1993;104:A688.
17. **Bjarnason I, Hayllar J, Somasundaram S, et al.** Misoprostol in the treatment of NSAID-induced enteropathy: a double blind, placebo controlled, parallel group study [Abstract]. Gastroenterology. 1993; 104:A669.
18. **Roth S, Agrawal N, Mahowald M, Montoya H, Robbins D, Miller S, et al.** Misoprostol heals gastroduodenal injury in patients with rheumatoid arthritis receiving aspirin. Arch Intern Med. 1989;149:775-9.
19. **Raskin JB, White R, Sue SO.** Efficacy and safety of misoprostol in the prevention of NSAID-induced gastric ulcers—preliminary findings. Gastroenterology. 1993;104:A177.

New International Ethical Guidelines for Research Involving Human Subjects

In March 1993, the Council of International Organizations of Medical Sciences (CIOMS), in collaboration with the World Health Organization (WHO), issued their *International Ethical Guidelines for Biomedical Research Involving Human Subjects* (1). This document, an extensive revision of the CIOMS-WHO guidelines published in 1982 (2), provides guidance for the proper application of the principles of the Declaration of Helsinki (3, 4) and focuses particularly on research sponsored by or initiated in developed countries and carried out in developing countries.

The first international code of ethics for research involving human subjects, the Nuremberg Code, was developed from 1947 to 1949 by the Nuremberg Military Tribunals during their trial of the Nazi physician-researchers; the Code's immediate purpose was to provide a set of standards for judgment of outrages committed in the name of science by the Nazi physician-researchers (3).

In 1964, the World Medical Association issued the Declaration of Helsinki (3, 4), which adapted the principles of the Nuremberg Code to the existential realities of medical research (3). For example, the Declaration replaced the Nuremberg Code's first principle ("The voluntary consent of the human subject is absolutely essential") (4) with a recognition of the legitimacy of proxy consent for research involving children and persons with cognitive impairment. Unlike the Nuremberg Code, the Declaration of Helsinki reflects the fact that research protocols might, and often do, include components expected to provide direct therapeutic, diagnostic, or prophylactic benefit to individual subjects and that ethical justification of such beneficial modalities should differ from that of nonbeneficial procedures.

The 1982 CIOMS-WHO *Guidelines* added, among other things, a requirement for review and approval of all proposed research by an "ethical review committee" (2). The Declaration of Helsinki, even in its third and most recent revision in 1989 (5), calls only for "consideration, comment and guidance" by "a specially appointed [independent] committee."

Until very recently, all international ethical codes and national laws and regulations were based on the assumption that medical research was hazardous to and exploitative of the subjects; this assumption reflects their origins as responses to the atrocities committed by Nazi physician-researchers, the calamitous experience with thalidomide, and, in the United States, exposés of such scandals as the Tuskegee Syphilis Study (5). Since the mid-1980s, this assumption has been largely displaced by the equally incorrect perception of participation in research and access to investigational drugs as benign and beneficial; this shift was due primarily to the successful efforts of highly articulate and influential AIDS activists (5). The writers of the 1993 *Guidelines*, recognizing the need for a balanced perspective, sought to encourage investigators to conduct ethical and beneficial research while maintaining necessary vigilance to safeguard the rights and welfare of research subjects (5).

When research is carried out in one country by investigators from another country, whose ethics should apply? This question, the most difficult one that con-

fronted the writers of the 1993 *Guidelines*, is subsidiary to the larger question of whether ethics are universal. Ethical universalists hold that ethical standards, properly understood, are and ought to be the same everywhere (6, 7). Ethical pluralists, by contrast, maintain that because systems of ethics are socially constructed within particular cultures and necessarily reflect their histories and traditions, variations across cultures are both expected and legitimate (8–10). The 1993 *Guidelines* reflect a compromise. Although some ethical principles, such as respect for persons, beneficence, and distributive justice, are considered universal, the legitimacy of ethical pluralism within specified limits is acknowledged.

The 1993 document is 52 pages long and consists of 15 statements of guidelines, each of which is followed by a detailed commentary on how the guideline should be applied in specific contexts. Novel features can be found in several categories, of which three are informed consent, ethical review, and obligations of sponsors and host countries.

The *Guidelines* state that in some cultures, when

> investigators cannot make prospective subjects sufficiently aware of the implications of participation [in research] to give adequately informed consent, the decision...should be elicited through a reliable intermediary such as a trusted community leader. In some cases other mechanisms...may be more suitable. However consent is obtained, all prospective subjects must be clearly told that their participation is entirely voluntary, and that they are free to refuse to participate or to withdraw...at any time without loss of any entitlement. (1)

Each individual must receive all information that would be conveyed if the study were to be conducted in a developed country; "otherwise, assurance of freedom to refuse or withdraw...would be meaningless" (1). Furthermore, "In some cultures...women's rights to exercise self-determination...are not acknowledged. In such cases, women should not normally be involved in research for which societies that recognize these rights require informed consent" (1). The *Guidelines* make provisions for affording women in such cultures access to investigational drugs and other therapies for which there are no equal or superior alternatives, even though they cannot give formal consent.

When research is initiated or sponsored by agencies in a developed country and carried out in a developing country, "the ethical standards applied should be no less exacting than they would be in the case of research carried out in [the initiating] country" (1). Committees in the initiating country are assigned special responsibility for determining that

> the scientific methods are sound and suitable for the aims of the research, [that] the drugs, vaccines or devices to be studied meet adequate standards of safety, [that] there is sound justification for conducting the research in the host country rather than in the country of the external sponsoring agency, and that the proposed research does not in principle violate the ethical standards of the external sponsoring country.... Committees in the host country have the special responsibility to determine whether the goals of research are responsive to the health needs and priorities of the host country. Moreover, because of their better understanding of the culture in which the research is proposed to be carried out, they have special responsibility for assuring the equitable selection of subjects and the acceptability of plans to obtain informed consent, to respect privacy, to maintain confidentiality, and to offer benefits that will not be considered excessive inducements to consent.
>
> In short, ethical review in the external sponsoring country may be limited to ensuring compliance with broadly stated ethical standards, on the understanding that ethical committees in the host country will have greater competence in reviewing the detailed plans for compliance in view of their better understanding of the cultural and moral values of the population in which the research is proposed to be conducted. (1)

The obligations of external sponsors of research to be conducted in developing countries are, in general, presented as *prima facie* obligations; that is, they may not apply in specific cases if competent authorities in both countries agree not to invoke them. For instance, research designed to develop a product (for example, a drug or vaccine) should be conducted only in host countries in which the disease for which the product is indicated is an important problem; any such product should be made reasonably available to inhabitants of the host country or community at the completion of successful testing. To help develop the capacity of the host country to carry out similar projects independently, sponsors are expected to employ and, if necessary, train local persons to function as investigators, research assistants, or data managers or to serve in other similar capacities. Sponsors are further expected, when appropriate, to make necessary health care services available to the population from which research subjects are to be recruited and to provide reasonable amounts of financial, educational, and other assistance to enable the host country to form independent and competent scientific and ethical review committees.

International declarations and guidelines issued by such agencies as the World Medical Association or CIOMS-WHO do not have the force of law. The extent to which they have any effect on the conduct of biomedical research involving human subjects depends entirely on the influence they have on the development of national regulations and the policies of professional organizations and agencies that provide funding for biomedical research. United States federal regulations for the protection of human research subjects have been influenced substantially by the Nuremberg Code and the Declaration of Helsinki. According to a survey conducted by CIOMS, the 1982 CIOMS-WHO *Guidelines* have had a major influence on research policies and practices around the world (1). The 1982 *Guidelines* had no perceptible effect on U.S. federal regulations because they were largely compatible with the regulations and because they were issued after the last major revision of the federal regulations was completed in 1981. The 1993 CIOMS-WHO *Guidelines* similarly do not create a need to consider substantial revision of the U.S. federal regulations; rather, they call attention to points that ought to be considered in addition to those covered in federal regulations—particularly when research is carried out in a developing country by investigators from a developed country.

Are these guidelines to be considered the final word on the ethics of multinational research? As Co-chair of

the Steering Committee that supervised their development, I can say most emphatically that they are not. Ethical codes, guidelines, and regulations, if they are to remain vital and valid, require constant interpretation, reinterpretation, and occasional revision in the light of practical experience in the field.

Robert J. Levine, MD
Yale University School of Medicine
New Haven, CT 06510

Acknowledgments: Several passages in this editorial are derived from a lecture presented at the Eleventh Annual Meeting of the International Clinical Epidemiology Network (INCLEN), Cairo, Egypt, 26 January 1993, and from a commentary to be published in the *INCLEN Newsletter*. The CIOMS/WHO *International Ethical Guidelines for Biomedical Research Involving Human Subjects* can be obtained from CIOMS, *c/o World Health Organization, Avenue Appia, 1211 Geneva 27, Switzerland.*

Ann Intern Med. 1993;119:340-341.

References

1. **CIOMS/WHO.** International Ethical Guidelines for Biomedical Research Involving Human Subjects. Geneva: CIOMS; 1993.
2. **CIOMS/WHO.** Proposed International Guidelines for Biomedical Research Involving Human Subjects. Geneva: CIOMS; 1982.
3. **Perley S, Fluss SS, Bankowski Z, Simon F.** The Nuremberg Code: an international overview. In: The Nazi Doctors and the Nuremberg Code: Human Rights in Human Experimentation. Annas GJ, Grodin MA, eds. New York: Oxford University Press; 1992: 149-73.
4. **Levine RJ.** Ethics and Regulation of Clinical Research. 2d edition. Baltimore, Maryland: Urban & Schwarzenberg; 1986.
5. **Levine, RJ.** Keynote address, XXVIth CIOMS Conference. In: Ethics and Research on Human Subjects: International Guidelines. Bankowski Z, Levine RJ, eds. Geneva: CIOMS; 1993: 4-10.
6. **Angell, M.** Ethical imperialism? Ethics in international collaborative clinical research [Editorial]. N Engl J Med. 1988;319:1081-3.
7. **IJsselmuiden CB, Faden RR.** Informed consent and public health research in Africa—another look. N Engl J Med. 1992;326:830-4.
8. **Levine RJ.** Informed consent: some challenges to the universal validity of the Western model. Law, Med Health Care. 1991;19:207-13.
9. **Christakis NA.** Ethics are local: engaging cross-cultural variation in the ethics for clinical research. Soc Sci Med. 1992;35:1079-91.
10. **Barry M.** Ethical considerations of human investigations in developing countries: the AIDS dilemma. N Engl J Med. 1988;319: 1083-5.