The Risk of Determining Risk with Multivariable Models

John Concato, MD, MS, MPH; Alvan R. Feinstein, MD, MS; and Theodore R. Holford, PhD

■ *Purpose:* To review the principles of multivariable analysis and to examine the application of multivariable statistical methods in general medical literature.

■ *Data Sources:* A computer-assisted search of articles in *The Lancet* and *The New England Journal of Medicine* identified 451 publications containing multivariable methods from 1985 through 1989. A random sample of 60 articles that used the two most common methods—logistic regression or proportional hazards analysis—was selected for more intensive review.

■ *Data Extraction:* During review of the 60 randomly selected articles, the focus was on generally accepted methodologic guidelines that can prevent problems affecting the accuracy and interpretation of multivariable analytic results.

■ *Results:* From 1985 to 1989, the relative frequency of multivariable statistical methods increased annually from about 10% to 18% among all articles in the two journals. In 44 (73%) of 60 articles using logistic or proportional hazards regression, risk estimates were quantified for individual variables ("risk factors"). Violations and omissions of methodologic guidelines in these 44 articles included overfitting of data; no test of conformity of variables to a linear gradient; no mention of pertinent checks for proportional hazards; no report of testing for interactions between independent variables; and unspecified coding or selection of independent variables. These problems would make the reported results potentially inaccurate, misleading, or difficult to interpret.

■ *Conclusions:* The findings suggest a need for improvement in the reporting and perhaps conducting of multivariable analyses in medical research.

Annals of Internal Medicine. 1993;**118**:201-210.

From Yale University School of Medicine, New Haven, Connecticut. For current author addresses, see end of text.

Although most physicians have received no instruction in multivariable methods of statistical analysis, the methods now commonly appear in medical literature. The results of multivariable analyses are often expressed in statements such as, "When other risk factors are controlled, a decrease of 5 units in substance X reduced disease by 10%," or "After adjustment for age and stage of disease, treatment with procedure Y reduced mortality by 25%."

Our purpose in the current research was to note the frequency with which multivariable analyses now appear in general medical journals, to identify some common problems and desirable precautions in the analyses, and to determine how well the challenges are being met. The investigation also provided a framework for a brief review—intended for clinical readers—of commonly used multivariable statistical methods.

General Principles

Format of Multivariable Analysis

In the types of multivariable analyses discussed here, the mathematical expressions described in Appendix 1 are used to relate two or more independent variables to an outcome or dependent variable. In those expressions, a linear regression coefficient indicates the impact of each independent variable on the outcome in the context of (or "adjusting for") all other variables. The values of the regression coefficients are obtained as the best mathematical fit for the specified model, although the selected model and the multivariable analysis may or may not provide a good absolute fit for the data.

The four main multivariable methods in medical literature have many mathematical similarities but differ in the expression and format of the outcome expressed as the dependent variable:

1. In multiple linear regression, the outcome variable is a continuous quantity, such as blood pressure in millimeters of mercury or sodium concentration in milliequivalents per liter.

2. In multiple logistic regression, the observed outcome variable is usually a binary event, such as alive versus dead or case versus control. The event occurs at a fixed point in time, such as mortality 1 year after surgery.

3. In discriminant function analysis, the outcome variable is a category or group to which a subject belongs. For example, patients may be classified as having obstructive, restrictive, vascular, or "other" forms of pulmonary dysfunction. The analytic results are often converted to a "score" used to classify observations into one of the categorical groups. For only two categories (such as healthy or diseased), this form of multivariable analysis produces results similar to logistic regression.

4. In proportional hazards regression, which is also known as Cox regression (1), the outcome variable is the duration of time to the occurrence of a binary "failure" event during a follow-up period of observation. The most common event in such analyses is death, but other failures can also be modeled. For example, each person's final state can be classified as either dead at a specified time or as "censored" if lost to follow-up or alive at the end of the study period. The results of a Cox model may be considered as an instantaneous incidence rate for the failure event.

The technical details of all four methods have been described elsewhere (2-6), and the current review is limited to the way two of the methods—multiple logistic regression and proportional hazards analysis—are used in medical literature. These multivariable methods have become particularly popular because binary dependent variables, such as survival state, are frequently used in clinical research. (Although the term "multivariate" is often used interchangeably with "multivariable," the former does not apply to the most common medical situations, where a single outcome variable is studied.)

Purposes of Multivariable Analysis

Multivariable methods can be used for at least five major purposes.

Bivariate Confirmation

Before multivariable analyses are initiated, most investigators have previously done simple bivariate analyses in which each independent variable is evaluated, one at a time, for its association with the outcome variable. (Because of the one-at-a-time examination, the analyses are often called "univariate," although two variables are being associated.) After the important independent variables have been identified in the simple analyses, the multivariable examination is usually done to confirm that the independent variables retain their importance in the simultaneous context of the other variables. For example, the effect of smoking on cardiovascular mortality can be evaluated along with the concomitant effects of age, systolic blood pressure, serum cholesterol level, and serum triglyceride value. If the initial bivariate impact of smoking is changed in the multivariable context, specific relationships can be investigated further.

Multivariable Confirmation

A second purpose of the multivariable methods is to confirm the results of non-regression analyses such as adjustment by cross-stratification (7) or standardization (8). For example, data on survival after prostatectomy, independent of type of surgery, can be examined for subjects cross-stratified in categorical groups of age and severity of illness. After suitable composite strata are formed, long-term mortality can then be examined in each stratum for transurethral compared with open prostatectomy, and the results can be checked with a multivariable regression model (9).

Screening

With large administrative data bases, the total number of variables may make bivariate analyses too time-consuming to obtain and difficult to assimilate. In such situations, multivariable analysis can be used as an initial screening process. "Important" variables can be suggested by the screening, with thresholds of quantitative and statistical importance (significance) determined by the researcher; more detailed analyses can then be performed. Software programs typically have "default values" that can be modified for levels of statistical significance, but quantitative significance is usually ignored by the computer and must be recognized and specified by the analyst.

Creating Risk Scores

The multiple variables that seem important may be assigned simple rating scores and combined into a single risk score used for predicting outcomes of individual patients. The choice of these ratings can be aided by the regression coefficients found in the multivariable analysis. A well-known simple, single combination of arbitrary ratings is the Apgar score, created by Virginia Apgar, who assigned integer ratings of 0, 1, or 2 to each of five variables in the assessment of a newborn infant (10). Although Apgar's initial decisions were made using clinical judgment and were not related to infants' outcome, a multivariable analysis might have been used to suggest suitable "weights" for the ratings.

Quantifying Risk of Individual Variables

In many instances, particularly in studies of risk factors, the impact of an individual variable is quantified for its specific effect among the other independent variables. For example, proportional hazards analysis was used to determine that "a 19% reduction in coronary heart disease risk was associated with each decrement of 8% of total plasma cholesterol" (11), adjusting for age, systolic blood pressure, cigarette smoking, and other factors. In these situations, the regression coefficients found in the multivariable analysis are converted to expressions of relative risks or hazards (with Cox regression) or odds ratios (with multiple logistic regression), indicating the individual variable's effect on the outcome.

Assumptions and Limitations

In the first three purposes just cited, the multivariable analyses are used mainly to confirm results that have already been documented with other methods or to suggest important variables that will be evaluated later in greater detail. In the fourth purpose, the risk score will combine the different variables and will demonstrate the impact of the combinations rather than each variable alone. Although the assumptions and limitations of multivariable models are important, they may be considered as having a secondary role when the combined scores are formed.

In the fifth purpose, when individual variables are examined to determine the magnitude of their impact or risk, the estimates will vary with the structure of the mathematical model and the coding of variables. The assumptions and limitations of the multivariable methods then become especially important for ensuring accurate results and valid interpretations.

Methods

Selection of Articles

The BRS Colleague complete text computer data base (12) was used to identify the use of the four cited multivariable methods in articles from *The Lancet* and *The New England Journal of Medicine* from 1985 to 1989. The search was limited to these two journals because they include a variety of general clinical topics, have large circulations, and have complete texts (not just keywords) available in the data base.

The computer search was designed to focus on original and special articles while excluding editorials, reviews, letters, and so forth. The search process included the text but not the references of published articles. The main terms used in the search included "multiple linear," "discriminant function," "logistic," "proportional hazards," or "Cox." Each main term was paired with "regression," "analysis," "function," "model," and "method." The accuracy of the computer search was checked with a manual inspection of articles published in each journal during a 6-month period.

After the initial search confirmed that the use of logistic regression and proportional hazards analysis was particularly frequent, these two methods were selected for further review. For this purpose, a random, stratified sample of 60 articles (28-87) was selected to provide three articles for each of the two methods, for each of the two journals, for each of the 5 years 1985 through 1989. With information about each usage excerpted and recorded on standardized forms, the 60 articles were reviewed to document the purpose for which each multivariable method was used and to evaluate the application and reporting of results.

If regression coefficients, relative risks, or odds ratios were listed for individual variables, the multivariable method was classified as having risk quantification as a purpose, and mathematical details were checked. The other multivariable analytic purposes were noted and classified according to the foregoing taxonomy, but the articles were not evaluated further.

Management of Problems and Precautions

Various guidelines (2-6, 13) have been suggested to encourage the appropriate execution, interpretation, and reporting of multivariable methods. In logistic and Cox regression, inattention to the guidelines can cause unreliable results in estimates of risk when the impact of an individual variable is quantified from its regression coefficient. In examining the published results for logistic and Cox regression, we checked to see how the investigators had managed and reported the six problems and precautions that follow. We reviewed the relevant data when available or accepted even a brief statement by the authors that a potential problem was evaluated.

Overfitting

The risk estimates may be unreliable if the multivariable data contain too few outcome events relative to the number of independent variables. For example, consider a cohort study of 1000 persons in which 5 deaths occur during 6 months of follow-up. The factors associated with death are determined from only five observations, although numerous baseline characteristics might be included in a multivariable analysis of 6-month mortality. Because outcome events are sparse, the resulting regression coefficients for individual variables may not be trustworthy; they may represent spurious associations, or the effects may be estimated with low precision. (Because the analysis depends on the smaller number of the two complementary outcome events, the problem is not corrected by focusing on the survival in 995 persons.)

Consequently, a large number of outcome events is needed if many independent variables are included in the analysis. In general, the results of models having fewer than 10 outcome events per independent variable are thought to have questionable accuracy (13, 14), and the usual tests of statistical significance may be invalid. Large confidence intervals associated with individual risk estimates may indicate an overfitted model under these circumstances.

A counterpart problem, underfitting, is also due to a scarcity of outcome events. Underfitting occurs when the power for detecting important relationships is low, and important variables may be omitted from a model. For example, consider a longitudinal study of 200 persons in whom 2 develop lung cancer after 10 years of follow-up. In this situation, the association of cigarette smoking and lung cancer would probably be undetectable, although analysis with a larger number of outcome events might identify a relationship between the independent variable and the outcome. Thus, a study with a relative paucity of outcome events may have misleading results due to overfitting or underfitting.

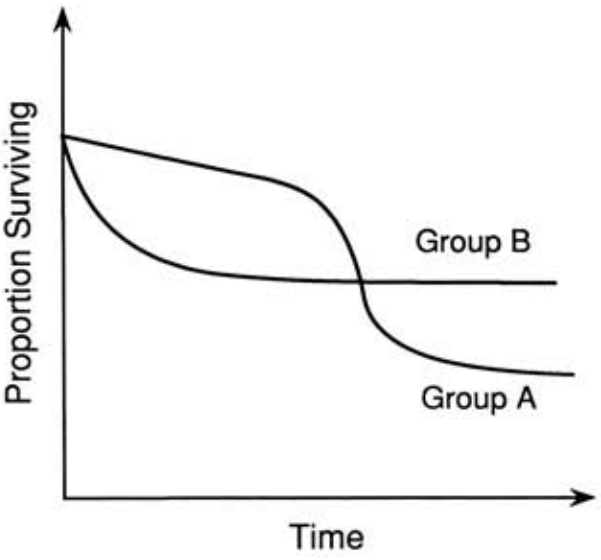

Figure 1. Nonproportional "hazards." Crossing of survival curves for groups A and B, violating the assumption of proportional hazards.

Nonconformity to a Linear Gradient

When a linear regression coefficient is estimated for a variable X, the implication is that regardless of the value of X, a unit change in X should always have the same effect on the outcome. If the independent variable is binary, then the regression coefficient represents only a single gradient as the variable "moves" from being absent to present. With ranked variables, however, several or many gradients may occur as the variable moves through its series of ordinal categories or continuous measurements. The value of the regression coefficient is calculated to be accurate as the average effect of X, but the result will be misleading if X has different effects in different zones. The implication may be particularly misleading if the average value of X does not occur in any of the zones.

For example, the impact of left ventricular ejection fraction on mortality is not linear: A decrease of 10% from 30% to 20% carries greater risk than a decrease from 50% to 40%. A single risk estimate may therefore be misleading unless conformity to the assumption of a linear gradient is evaluated. Various methods are available to check the assumption of a linear gradient, including assessing the impact of each variable separately in zones of the ranked data (2).

Violation of Proportional Hazards

Another potential problem occurs in the Cox regression method, in which the risk or "hazard" of an independent variable is assumed to be constantly proportional (the relative risk does not change with time). The problem can be illustrated by considering survival curves with a binary variable that identifies patients in groups A and B, representing two forms of treatment. If the hazard is proportional, the survival curve of one group will not cross the survival curve of the other group.

When crossover (as shown in Figure 1) or other nonproportional survival patterns occur, the corresponding risk estimates

may be inaccurate (3, 13, 14). For example, the effects of an early survival advantage followed by late excess mortality may cancel for group A and group B, so that the Cox method may indicate that group membership has no effect. Therefore, when the Cox method is used, the independent variables should be evaluated for adherence to the assumption of proportional hazards. If nonproportional hazards are found, the Cox method can be applied as an analysis stratified by the offending variable or by including time-dependent covariates (15).

No Report of Tests for Interactions

An interaction occurs between independent variables if the impact of one variable on the outcome event depends on the level of another variable. For example, consider a logistic or Cox regression model with two independent binary variables: smoking and use of oral contraceptives, each given possible values of 0 (no exposure) or 1 (exposure). If interactions are not considered, the regression coefficient for oral contraceptives represents the impact of oral contraceptive use on the outcome event for both levels of smoking. If oral contraceptive use and cigarette smoking have a significant interaction, however, the impact of oral contraceptives depends on whether exposure to smoking occurs. Without attention to interaction, the oral contraceptive coefficient would offer a misleading quantitative estimate of the impact of oral contraceptive use.

Interactions may be checked either because suspicions are raised by clinical judgment before the analysis is done or by a specific statistical examination whenever multivariable methods are applied. The regression methods do not automatically examine interactions, which can be evaluated only if appropriate interaction terms are explicitly included in the model. (A potential problem arises because of the increased opportunity for overfitting when interaction terms are added to a model.) Despite the absence of a universal rule dictating appropriate tests for interactions in all circumstances, readers of published results will not know whether any tests for interactions were conducted unless the testing is mentioned in the report.

Unspecified Coding of Variables

The apparent effect of an independent variable will depend on the corresponding units of measurement and coding for that variable. For example, the regression coefficient for the impact of age on long-term mortality will be different if age is coded in 1-year increments, in 10-year intervals (. . .; 50 to 59 years; 60 to 69 years; . . .), or dichotomously as < 65 versus ≥ 65 years. If the values of regression coefficients are cited without concomitant citation of the units of coding for independent variables, then readers will be unable to interpret the actual magnitude or effect of the risk estimates.

Independent variables expressed in ordinal or binary categories must receive arbitrary numerical codings for a multivariable analysis. For example, binary variables can be coded (2) as $-1/+1$ ("marginal method") or 0/1 ("partial method"). According to the selected code, the magnitude of the regression coefficient will vary by a factor of 2 in logistic and Cox regression. In addition, ordinal variables can be coded with integer values or with "dummy" variables (2) that compare all other categories to a single reference category. Because the particular measurement or coding scheme can have substantial effects on the numerical values and interpretation of the regression coefficients, readers should always be notified of how the coding was used in a multivariable analysis.

Unspecified Selection of Variables

The choice of independent variables included in a final multivariable analysis is not a simple task. Although candidate variables may be chosen from previous research results or from clinical experience, automated algorithms exist for selecting among variables thought to have possible prognostic value. Most software programs offer "forward" or "backward" selection of variables. These procedures include or delete variables one at a time until a specified threshold of statistical significance is met (16). For example, a forward model might include all variables for which the associated regression coefficient has $P < 0.05$. Yet another process investigates "all possible subsets" of candidate variables (16). An alternative "change-in-estimate" selection procedure has a forward direction but allows variables to be deleted after they are entered in a model if a subsequent variable improves overall prediction (16). The "principal components" method involves a reduction in the number of candidate variables before the modeling process begins (13).

Although intended to be used as tools in exploring data, the various automated procedures may not be recognized and fully understood by researchers. None of the automated multivariable methods uses clinical judgments or consider biological sensibility in the analysis. The fundamental issue for readers to understand is that the final model will depend on the chosen selection process and that variables may be retained or excluded merely by mathematical details of the analysis.

Additional Issues

Three other issues that are also important considerations in the application of multivariable analyses are not easily evaluated from published reports.

Collinearity of Variables

A problem occurs if independent variables have a high correlation with one another. For example, measurements of ventricular ejection fraction and ventricular contractility may contain redundant information regarding the risk for mortality (unless the relationship between the two variables is itself under study). The results of a multivariable analysis including both variables might identify one or the other as important but is not likely to include both variables if they are highly correlated. Furthermore, with a high correlation the quantitative risk estimates for each variable may be imprecise. Although software packages include tests to examine the data for collinearity, investigators have noted that " . . .it is possible for variables to pass these tests and have the program run but yield output that is clearly nonsense" (17). As a precaution, the most clinically relevant (among similar) variables can be chosen for inclusion in a multivariable analysis, or the principal components selection method can be used to choose among the variables.

Influential Observations

A small number of observations for individual subjects can have a substantial effect on the final analytic results if their corresponding values are distinctly different from all others. Such "outliers" produce a problem similar to the example of inadvertently including a child's age in the calculation of mean age for patients seen in a geriatric clinic. Unlike this simple situation, however, the problem of "influential" observations is often obscure in a multivariable context, where the relationship among numerous independent variables determines outlier status.

Although mathematical methods exist for dealing with influential observations (mostly by deleting observations from the data set), the optimal scientific approach to this situation cannot be specified in advance, particularly because outliers may accurately reflect an important biologic relationship rather than representing a mathematical aberration. The best precaution for this potential problem is to insist on high-quality data for the multivariable analysis (to avoid spurious outliers) or to analyze the data by other methods such as stratified analysis.

Validation of a Model

A multivariable analysis assigns coefficients to independent variables selected as important predictors of outcome. As with all statistical models, the results require validation to ensure protection against unrecognized problems and limitations. Common methods for validating models include 1) performing a test analysis on a subsample of patients followed by a subsequent validation analysis on the remaining patients; 2) repeating the analysis on an independent sample of patients; 3) using "jackknife" or "bootstrap" procedures (18) in which the same analysis is performed multiple times on a series of subsets from the same data set to investigate the stability of coefficients and predictive ability of the model.

A related issue is the mathematical fit of the final model.

Table 1. Multivariable Methods: Number of Appearances in The Lancet and The New England Journal of Medicine

Multivariable Method	Year				
	1985	1986	1987	1988	1989
Multiple logistic regression, *n*	32	23	43	40	56
Proportional hazards analysis, *n*	19	24	35	30	40
Multiple linear regression, *n*	10	21	19	17	22
Discriminant function analysis, *n*	2	7	4	5	2
Total, *n*	63	75	101	92	120
Number of original articles, *n*	661	661	661	644	660
Frequency of any of the four methods, %	10	11	15	14	18

Indexes of "goodness-of-fit" evaluate how effectively the calculated model fits the actual data for estimating the outcome variable. Although various indexes have been developed (2), a consensus is lacking even among statisticians about which index is most appropriate. (Details of the methods and discussions are beyond the scope of this review.) As a fundamental principle, including additional independent variables in a model will enhance mathematical goodness of fit but can cause problems such as overfitting of data or collinearity of variables.

Results

Frequency of Use

Table 1 shows results for the four multivariable methods in the two journals from 1985 to 1989. The number of annual citations has increased steadily, although the total number of pertinent articles has remained essentially constant. The frequency of the multivariable methods has thus increased from 10% to 18% over the 5-year period; and in 1989, one of the four methods appeared on average at least once per week in each journal.

The manual inspection of articles from July to December 1989 determined that the computer search had correctly identified "original articles," "special articles," "medical intelligence," "medical progress," and "case records" in *The New England Journal of Medicine* (n = 191); and "original articles," "preliminary communications," and "methods and devices" in *The Lancet* (n = 135). Eleven additional publications in *The Lancet,* however, were mistakenly identified as "original," including two articles on medicine and the law, a letter to the editor, a correction, and so forth. Thus, the computer search identified a proportionate excess of 11 of 326, or 3%, of the desired articles for the checked period in both journals. The error rate seemed too small to warrant corrections for the reported frequency counts.

The manual inspection was also used to evaluate the computer citations of the multivariable methods. In three articles, the search term appeared in the discussion (such as describing a logistic analysis done in another publication) rather than in the study methods. Several articles using major modifications of classical multivariable methods were not identified, but these techniques were not an intended subject of our investigation. Furthermore, in five articles the authors did not distinguish between simple linear regression and multiple linear regression when reporting the use of "linear regression." Thus, the results in Table 1 can be regarded as reasonably accurate frequencies of the cited multivariable methods in these two journals.

Because the logistic regression and proportional hazards methods were particularly common, they were the focus of further evaluation in the random sample of 60 articles.

Authors' Purpose in Using Multivariable Analysis

In 44 (73%) of the 60 publications using logistic and Cox regression, the multivariable methods were applied to quantify risk estimates reported as regression coefficients, odds ratios, or relative risks for individual variables. For example, in a study of prenatal X-ray exposure and childhood cancer in twins (31), the relative risk for cancer, adjusting for birth weight, was 2.4 for exposed compared with nonexposed children; and when type A behavior was related to outcome of coronary heart disease (79), the mortality for Type A persons, after adjustment for other risk variables, was 58% that of Type B persons.

The remaining 16 (27%) of the 60 articles used multivariable methods as follows: Thirteen studies confirmed the results of other forms of analysis (such as simple bivariate analysis [40] or a Mantel-Haenszel analysis [62]); one study screened data for important variables (to identify risk factors for gastric cancer after gastric surgery for benign disease [65]); one report created a risk score (to predict relapse among patients with testicular teratoma [72]); and one investigation checked for interactions only (in a report of tamoxifen therapy for breast cancer [58]).

Problems in Reporting and Application

The pertinent statistical "package" or program—such as SAS (19) or BMDP (20)—used to perform a multivariable analysis should be reported to the reader. Citing the information is analogous to a laboratory researcher indicating the particular experimental protocol used for physiologic measurements. Yet, the pertinent program was mentioned in only 17 (39%) of the 44 articles reviewed.

In addition to this general consideration, the six cited principles were evaluated in the 44 studies where logistic regression and proportional hazards analysis methods were applied for quantifying risk of individual variables. Potential problems involving collinear variables, influential observations, and model validation were not evaluated in the articles under review.

Overfitting

The criterion for overfitting data was violated in 19 (42%) of the 44 studies. For example, in an investigation (49) of coronary restenosis after dietary supplementation with n-3 fatty acids, 11 independent variables were included in a logistic model containing 29 outcome events for 85 patients. The ratio of 2.6 (29/11) events per independent variable is much smaller than the suggested ratio of 10, and the overfitting could affect the quantitative assertion that "therapy reduced the likelihood of restenosis by 77%." Of note, the 95% confidence interval for the effect of the intervention was quite wide (10% to 92% reduction in restenosis), consistent with overfitting.

Nonconformity to a Linear Gradient

The criterion for nonconformity to a linear gradient did not apply to 30 articles where the analyses used only binary independent variables. Of the 14 articles with ranked independent variables, however, 4 (29%) gave no indication of checking for nonconformity to a linear gradient. For example, when calcium intake (78) seemed to have a protective effect on incidence of hip fracture, the main result was reported as a relative risk = 0.6 "per 198 mg calcium/1000 kcal." (The value of 198 mg was chosen because it was the standard deviation of calcium intake.) The result implies that an increase in calcium intake of about 200 mg from any level lowers the risk for hip fracture by 40%, because a relative risk of 0.6 is a proportionate decrease of 0.4. A concomitant graphic display of the data, however, suggested otherwise: The difference in hip fracture rates was minimal for persons with "low" compared with "mid" calcium intake but was substantial for persons with "mid" compared with "high" calcium intake. The 40% risk reduction, representing an average risk among all patients, could therefore be due mainly to the very low risk of hip fracture in subjects with "high" calcium intake.

In the remaining 10 studies, ranked variables were appropriately divided and checked in several ordinal zones of data delineated by the investigators. Individual risk estimates were then calculated separately for each of the zones. In one study, for example, the risk for ulcerative keratitis was calculated for duration of contact lens use divided in zones of 1 day, 2 days to 1 week, more than 1 but not more than 2 weeks, and more than 2 weeks (57); in another study the risk for cataract formation was determined separately for ultraviolet radiation divided in quartiles of exposure (51).

Violation of Proportional Hazards

A check of the assumption of proportional hazards over time was not mentioned in 17 (81%) of the 21 studies using Cox regression. The proportionality criteria may have been violated in these instances, but quantitative risk estimates were nevertheless reported. The accuracy of the estimates is therefore uncertain.

No Report of Tests for Interactions

Tests for possible interactions between independent variables were not mentioned in 32 (73%) of the 44 articles. When this criterion was satisfied, the publications sometimes discussed interactions that were suspected before the analysis (for example, aspirin and sulfinpyrazone interacting on the risk of cardiac death among patients with unstable angina [63]). In other instances, the investigators evaluated pairwise interactions of all independent variables. Although testing for interactions may or may not have been important for the clinical and statistical context, the reader would at least be aware of the testing that was done. In the remaining articles, interactions may have been assessed during the research, but in the absence of a published statement, readers evaluating the results have no assurance that the assessments were even considered.

Unspecified Coding of Variables

The coding classification of pertinent independent variables was not reported in 37 (84%) of the 44 articles. Such omissions prevent the reader from interpreting the quantitative results. In the remaining 7 (16%) articles, the coding scheme was described in the methods section, in the table of results for the multivariable model, or in an appendix. The description occupied only a modest amount of space in the publication.

Unspecified Selection of Variables

The selection process for choosing independent variables was described in 38 (86%) of publications. This finding suggests that most investigators (or their statistical consultants) are aware of the impact of the selection mechanism on the results of multivariable analyses, and that the particular mechanism is considered important enough to mention.

Discussion

Multivariable methods have become increasingly popular forms of data analysis in medical research; and two of the methods—logistic regression and proportional hazards analysis—appeared in 15% (96 of 660) of articles published in two leading general medical journals during 1989. These two regression methods were frequently used to identify the effect of individual variables in a multivariable context and to offer a quantitative estimate of risk for the individual effect. Nevertheless, certain important precautions were often omitted or not reported when the methods were applied.

The six problems that were reviewed (and three more discussed) in this paper are listed in Table 2. The criteria represent our minimum standards for the conduct and reporting of research using multivariable analysis. We do not expect everyone to agree with these standards, but a consensus cannot be reached unless the issues are elaborated and investigated.

The problem of unstable risk estimates and inappropriate P values produced by overfitted data can be prevented by ensuring an adequate number of outcome events. Although inadequacy cannot be formally tested in a manner analogous to power calculations for type 2 error (21), a low (less than 10:1) ratio of outcome events to independent variables makes the risk estimates uncertain. In the earlier cited example in which a protective effect of dietary n-3 fatty acids (49) was based on an overfitted logistic regression model, the findings

Table 2. Problems and Issues in the Application and Reporting of Logistic Regression and Proportional Hazards Analysis

Problem or Issue	Description
Problem	
Overfitting of data	Fewer than 10 outcome events per independent variable in the model
Nonconformity to linear gradient	Nonconstant impact of variables in different zones of ranked data
Nonproportional risk	Violation of assumption of proportional hazard function over time (in the proportional hazards method)
No report of tests for interactions	Check not mentioned for interactions between independent variables
Unspecified coding of variables	Unknown classification or codings for independent variables
Unspecified selection of variables	Unknown method of selecting among candidate independent variables
Issue	
Collinear variables	Independent variables with high correlation to each other
Influential observations	"Outlier" observations that have a substantial effect on results
Validation of model	Separate method of confirming analytic results

were not confirmed in a subsequent, similar investigation (22). Although the inconsistent results were attributed to various elements of study design, limitations of the data analysis itself were not considered.

The problems of overfitted models have been reported (2,13, 23) in the statistical literature but are not widely recognized. The key issue in the overfitting is an ample number of outcome events, not just a large sample size. When numerous variables are included in an attempt to "control" or "adjust" the data, accuracy of results can be threatened by overfitting or by other mechanisms (24). The number of variables selected for analysis should therefore be parsimonious, based on clinical sensibility and suitable data quality.

In checking the problem of nonlinearity when ranked variables are used directly, the analyst can compare the observed and the multivariable model's predicted values for the outcome over the range of each variable. A single risk estimate is inappropriate if the pattern of "errors" is nonrandom. For example, if arterial carbon dioxide tension (PCO_2) is included as a predictor in a multivariable analysis of death from chronic obstructive lung disease, the corresponding linear regression coefficient will represent the average impact of PCO_2 on mortality. If the actual mortality substantially differs from predicted mortality for "high" values of PCO_2, then the analysis will incorrectly estimate the true risk for such patients.

In "nonlinear" circumstances the risks should be quantitatively estimated not as a single value but in zones or categories of the data. Although checking for a linear gradient is not a trivial exercise, a common method available in software packages involves visual inspection of appropriate data. Alternatively, the analyst can use other forms of multivariable analysis, such as cross-stratification (7), to evaluate whether the variables conform to a linear gradient.

In the papers under review, the problem of nonconformity to a linear gradient was frequently avoided by the strategy of using binary independent variables—a tactic found in 30 of the 44 pertinent articles. The true impact of continuous or ordinal variables, however, may be masked when two binary zones are created. For example, the "J-shape" relationship of serum cholesterol and mortality cannot be described by binary zones such as < 5.20 mmol/L versus ≥ 5.20 mmol/L (corresponding to < 200 mg/dL versus ≥ 200 mg/dL). Rather than using a dichotomous classification, continuous variables may be converted into an array of ordinal zones or transformed into "dummy" variables (2) appropriate for the clinical context of each analysis. Although the ideal number of zones cannot be specified in advance and often requires judgment, clinicians should be aware of this issue in multivariable modeling.

The problem of nonproportionality in Cox regression can be avoided if hazard functions are suitably checked and reported. Although criteria for identifying "severe" violations are lacking, a rigorous scientific analysis should include evaluating methodologic assumptions and reporting the results. Techniques such as checking for proportional hazards using logarithmic graphs (15) may not be familiar to all readers but, when described, would indicate that the proportional hazards assumption had been evaluated.

The interaction problem is illustrated by the association of asbestos exposure and cigarette smoking with lung cancer, initially thought (25) to interact: The risk for asbestos-exposed persons who also smoked cigarettes was substantially greater than the risk anticipated merely from combining the risks calculated individually for asbestos exposure and for cigarette smoking. Although subsequent data (26) did not confirm these results, such interactions represent another threat to the constancy implied by the reporting of regression coefficients. A variable whose impact is linear when acting alone may be nonlinear when acting jointly with other variables.

The fifth principle requires an explicit statement of the way the independent variables are analytically classified and coded. This statement can be easily incorporated in the text, tables, or appendix to allow the reader to interpret the quantitative results. Such disclosure is obviously crucial for interpreting the numerical magnitude of a cited risk factor.

Similarly, a statement indicating the method of selecting among candidate independent variables is desirable. Readers should be aware that some variables may have minimal impact on the outcome despite achieving "statistical significance," whereas other variables that fail to achieve the threshold of "$P < 0.05$" may still have a substantial effect on the outcome. (This distinction between quantitative and statistical significance occurs in

all forms of analysis [27] and is not unique to multivariable procedures.)

The remaining principles relating to collinear variables, influential observations, and model validation require raw data for evaluation. Readers of the published reports therefore cannot easily determine if a problem exists. Investigators who are aware of the principles, however, can publish descriptive statements to indicate that suitable precautions were taken during the analysis.

Because accurate and understandable results are required in communicating medical research, the findings of this review suggest a need for substantial improvements in reporting and perhaps in conducting multivariable analyses.

Appendix 1. Mathematical Expressions of Multivariable Analysis

Multivariable analysis relates independent variables $X_1, X_2, X_3, \ldots X_n$ to an outcome variable via a model expressed as a combination: $G = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \ldots + \epsilon$, where G is a function arranged in various mathematical forms (see below); b_j is a regression coefficient indicating the impact of each X_j variable on the outcome; and b_0 is an intercept term, which is usually included in the model. If a particular $b_j = 0$, then variable X_j has no impact on the outcome; a positive value of b_j indicates that higher values of X_j are associated with an increase in the outcome expressed as G; and negative values have the reverse effect. A random variable ϵ is an "error" term representing the increment by which any individual G value deviates from the calculated value of G.

The function G is arranged in different mathematical forms for each multivariable method.

1. Multiple linear regression: G = outcome variable
2. Multiple logistic regression: $G = \ln [P/(1-P)]$; where P is probability of the outcome event occurring and G is the "logit" or log odds of the outcome.
3. Discriminant function analysis: G = relative probability of each category.
4. Proportional hazards analysis: mortality or incidence rate as "hazard function" $H(t,x) = H_0(t)e^G$, where t is elapsed time after a starting point (zero time) and $H_0(t)$ is an underlying hazard when all X_j are zero.

The following comments can be added about the two main methods under review.

Logistic regression may be done as a "conditional" analysis for studies with small numbers of observations within strata, or as an "unconditional" analysis for unstratified studies or studies with large numbers of observations within strata. The method has also been adapted for modeling outcome variables having three or more ordinal categories.

Proportional hazard analysis is sometimes reported in terms of the hazard function, indicating the probability of a binary outcome event occurring at an instant in time, conditional on the subject surviving up to that instant. For pragmatic usage, survival is estimated as $S(t)^{\exp(G)}$, where S(t) is the summary survival curve for the group (possibly hypothetical) where G = 0.

Presented in part at the annual meeting of the American Federation for Clinical Research, Baltimore, Maryland, 6 May 1991.

Grant Support: Dr. Concato was a Postdoctoral Fellow in the Robert Wood Johnson Clinical Scholars Program, supported by the Department of Veterans Affairs, when this study was conducted.

Requests for Reprints: John Concato, MD, Medical Service/111, West Haven Veterans Affairs Medical Center, 950 Campbell Avenue, West Haven, CT 06516.

Current Author Addresses: Dr. Concato: Yale University School of Medicine, Department of Medicine and Medical Service/111, West Haven Veterans Affairs Medical Center, 950 Campbell Avenue, West Haven, CT 06516.

Dr. Feinstein: Yale University School of Medicine, Departments of Medicine and Epidemiology and Public Health, P.O. Box 3333, 333 Cedar Street, New Haven, CT 06510.

Dr. Holford: Yale University School of Medicine, Department of Epidemiology and Public Health, P.O. Box 3333, 60 College Street, New Haven, CT 06510.

References

[Note: The first 27 citations indicate sources of comment or other pertinent references. Citations 28 through 87 indicate the 60 articles that were selected for special review. They are cited by method and in chronological order for both journals.]

1. **Cox DR.** Regression methods and life tables (with discussion). J R Stat Soc B. 1972;34:187-220.
2. **Hosmer DW, Lemeshow S.** Applied Logistic Regression. New York: Wiley; 1989.
3. **Kalbfleisch JD, Prentice RL.** The Statistical Analysis of Failure Time Data. New York: Wiley; 1980.
4. **Lee ET.** Statistical Methods for Survival Data Analysis. Belmont, California: Lifetime Learning Publications; 1980.
5. **Kleinbaum DG, Kupper LL, Muller KE.** Applied Regression Analysis and Other Multivariable Methods. Boston: PWS-Kent Publishing Co.; 1988.
6. **Breslow NE, Day NE.** Statistical Methods in Cancer Research. v I (1980) and v II (1987). Lyon: International Agency for Research on Cancer.
7. **Feinstein AR.** Prognostic stratification. In: Clinical Biostatistics. St. Louis: CV Mosby; 1977:385-443.
8. **Armitage P, Berry G.** Statistical Methods in Medical Research. Boston: Blackwell; 1987:399-405.
9. **Concato J, Horwitz RI, Feinstein AR, Elmore JG, Schiff SF.** Problems of comorbidity in mortality after prostatectomy. JAMA. 1992; 267:1077-82.
10. **Apgar V.** A proposal for a new method of evaluation of the newborn infant. Curr Res Anesth. 1953;32:260-7.
11. Lipid Research Clinics Program Investigators. The lipid research clinics coronary primary prevention trial results. JAMA. 1984;251: 365-74.
12. BRS Colleague Users Manual. Latham, New York: BRS Information Technologies; 1988.
13. **Harrell FE Jr, Lee KL, Matchar DB, Reichert TA.** Regression models for prognostic prediction: advantages, problems, and suggested solutions. Cancer Treat Rep. 1985; 69:1071-7.
14. **Harrell F.** The PHGLM procedure. In: SUGI Supplemental Library User's Guide. Cary, North Carolina: SAS Institute Inc.; 1983:267-94.
15. **Christensen E.** Multivariate survival analysis using Cox's regression model. Hepatology. 1987;7:1346-58.
16. **Greenland S.** Modeling and variable selection in epidemiologic analysis. Am J Pub Health. 1989;79:340-9.
17. **Hosmer DW, Lemeshow S.** Applied Logistic Regression. New York: Wiley; 1989:131.
18. **Efron B, Gong G.** A leisurely look at the bootstrap, the jackknife, and cross-validation. American Statistician. 1983;37:36-48.
19. SAS Institute Inc. SAS Guide for Personal Computers. Version 6.03. Cary, North Carolina: SAS Institute Inc.; 1988.
20. **Dixon WJ.** BMDP Statistical Software. University of California Press; 1987.
21. **Freiman JA, Chalmers TC, Smith H Jr, Kuebler RR.** The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. N Engl J Med. 1978; 299:690-4.
22. **Reis GJ, Boucher TM, Sipperly ME, Silverman DI, McCabe CH, Baim DS, et al.** Randomised trial of fish oil for prevention of restenosis after coronary angioplasty. Lancet. 1989;2:177-81.
23. **Harrell FE Jr, Lee KL, Califf RM, Pryor DB, Rosati RA.** Regression modelling strategies for improved prognostic prediction. Stat Med. 1984;3:143-52.
24. **Walter SD, Feinstein AR, Wells CK.** A comparison of multivariable mathematical models in predicting survival—II. Statistical selection of prognostic variables. J Clin Epidemiol. 1990;43:349-59.
25. **Selikoff IJ, Hammond EC, Churg J.** Asbestos exposure, smoking, and neoplasia. JAMA. 1968;204:106-12.
26. **Rom WN; ed.** Environmental and Occupational Medicine. Boston: Little, Brown; 1983:157-82.

27. **Feinstein AR.** Clinical Epidemiology. Philadelphia: W.B.Saunders; 1985:396-400.

Logistic Regression

28. **Rahman M, Rahaman MM, Wojtyniak B, Aziz KM.** Impact of environmental sanitation and crowding on infant mortality in rural Bangladesh. Lancet. 1985;2:28-31.
29. **Ludlam CA, Tucker J, Steel CM, Tedder RS, Cheingsong-Popov R, Weiss RA, et al.** Human T-lymphotropic virus type III (HTLV-III) infection in seronegative haemophiliacs after transfusion of factor VIII. Lancet. 1985;2:233-6.
30. **Stadel BV, Rubin GL, Webster LA, Schlesselman JJ, Wingo PA.** Oral contraceptives and breast cancer in young women. Lancet. 1985;2:970-3.
31. **Harvey EB, Boice JD Jr, Honeyman M, Flannery JT.** Prenatal X-ray exposure and childhood cancer in twins. N Engl J Med. 1985;312: 541-5.
32. **Lesko SM, Rosenberg L, Kaufman DW, Helmrich SP, Miller DR, Strom B, et al.** Cigarette smoking and the risk of endometrial cancer. N Engl J Med. 1985;313:593-6.
33. **Hurwitz ES, Barrett MJ, Bregman D, Gunn WJ, Schonberger LB, Fairweather WR, et al.** Public Health Service study on Reye's syndrome and medications. N Engl J Med. 1985;313:849-57.
34. **Francis CW, Marder VJ, Evarts CM.** Lower risk of thromboembolic disease after total hip replacement with non-cemented than with cemented prostheses. Lancet. 1986;1:769-71.
35. **Knutsson A, Akerstedt T, Jonsson BG, Orth-Gomer K.** Increased risk of ischaemic heart disease in shift workers. Lancet. 1986;2:89-92.
36. **Witteman JC, Kok FJ, van Saase JL, Valkenburg HA.** Aortic calcification as a predictor of cardiovascular mortality. Lancet. 1986;2: 1120-2.
37. **Kreiss JK, Koech D, Plummer FA, Holmes KK, Lightfoote M, Piot P, et al.** AIDS virus infection in Nairobi prostitutes. N Engl J Med. 1986;314:414-8.
38. **Bowden RA, Sayers M, Flournoy N, Newton B, Banaji M, Thomas ED, et al.** Cytomegalovirus immune globulin and seronegative blood products to prevent primary cytomegalovirus infection after marrow transplantation. N Engl J Med.1986;314:1006-10.
39. **Freedman DS, Srinivasan SR, Shear CL, Franklin FA, Webber LS, Berenson GS.** The relation of apolipoproteins A-I and B in children to parental myocardial infarction. N Engl J Med. 1986;315:721-6.
40. **Cohen J, Moore RH, Al Hashimi S, Jones L, Apperley JF, Aber VR.** Antibody titres to a rough-mutant strain of *Escherichia coli* in patients undergoing allogeneic bone-marrow transplantation. Lancet. 1987;1:8-11.
41. **McManus C, Richards P.** Choice and ordering of medical school applications: cause for concern. Lancet. 1987;2:33-5.
42. **Krauss RM, Lindgren FT, Williams PT, Kelsey SF, Brensike J, Vranizan K, et al.** Intermediate-density lipoproteins and progression of coronary artery disease in hypercholesterolaemic men. Lancet. 1987;2:62-6.
43. **Mandel EM, Rockette HE, Bluestone CD, Paradise JL, Nozza RJ.** Efficacy of amoxicillin with and without decongestant-antihistamine for otitis media with effusion in children. N Engl J Med. 1987;316: 432-7.
44. **Louik C, Mitchell AA, Werler MM, Hanson JW, Shapiro S.** Maternal exposure to spermicides in relation to certain birth defects. N Engl J Med. 1987;317:474-8.
45. **Daling JR, Weiss NS, Hislop TG, Maden C, Coates RJ, Sherman KJ, et al.** Sexual practices, sexually transmitted diseases, and the incidence of anal cancer. N Engl J Med. 1987;317:973-7.
46. National Heart Foundation of Australia Coronary Thrombolysis Group. Coronary thrombolysis and myocardial salvage by tissue plasminogen activator given up to 4 hours after onset of myocardial infarction. Lancet. 1988;1:203-8.
47. **Marwick TH, Case C, Siskind V, Woodhouse SP.** Adverse effect of early high-dose adrenaline on outcome of ventricular fibrillation. Lancet. 1988;2:66-8.
48. **Yudkin JS, Forrest RD, Jackson CA.** Microalbuminuria as predictor of vascular disease in non-diabetic subjects. Lancet. 1988;2:530-3.
49. **Dehmer GJ, Popma JJ, van den Berg EK, Eichhorn EJ, Prewitt JB, Campbell WB, et al.** Reduction in the rate of early restenosis after coronary angioplasty by a diet supplemented with n-3 fatty acids. N Engl J Med. 1988;319:733-40.
50. **Ron E, Modan B, Boice JD Jr, Alfandary E, Stovall M, Chetrit A, et al.** Tumors of the brain and nervous system after radiotherapy in childhood. N Engl J Med. 1988;319:1033-9.
51. **Taylor HR, West SK, Rosenthal FS, Munoz B, Newland HS, Abbey H, et al.** Effect of ultraviolet radiation on cataract formation. N Engl J Med. 1988;319:1429-33.
52. **Price WH, Morris SW, Kitchin AH, Wenham PR, Burgon PR, Donald PM.** DNA restriction fragment length polymorphisms as markers of familial coronary heart disease. Lancet. 1989;1:1407-11.
53. **Webster A, Lee CA, Cook DG, Grundy JE, Emery VC, Kernoff PB, et al.** Cytomegalovirus infection and progression towards AIDS in haemophiliacs with human immunodeficiency virus infection. Lancet. 1989;2:63-6.
54. **Wahrendorf J, Chang-Claude J, Liang QS, Rei YG, Munoz N, Crespi M, et al.** Precursor lesions of oesophageal cancer in young people in a high-risk population in China. Lancet. 1989;2:1239-41.
55. **Sheinfeld J, Schaeffer AJ, Cordon-Cardo C, Rogatko A, Fair WR.** Association of the Lewis blood-group phenotype with recurrent urinary tract infections in women. N Engl J Med. 1989;320:773-7.
56. **Hayes EB, Matte TD, O'Brien TR, McKinley TW, Logsdon GS, Rose JB, et al.** Large community outbreak of cryptosporidiosis due to contamination of a filtered public water supply. N Engl J Med. 1989;320:1372-6.
57. **Schein OD, Glynn RJ, Poggio EC, Seddon JM, Kenyon KR, and the Microbial Keratitis Study Group.** The relative risk of ulcerative keratitis among users of daily-wear and extended-wear contact lenses. N Engl J Med. 1989;321:773-8.

Proportional Hazards Analysis

58. Controlled trial of tamoxifen as single adjuvant agent in management of early breast cancer. Analysis at six years by Nolvadex Adjuvant Trial Organisation. Lancet. 1985;1:836-40.
59. **Haybittle JL, Hayhoe FG, Easterling MJ, Jelliffe AM, Bennett MH, Vaughan Hudson G, et al.** Review of British National Lymphoma Investigation studies of Hodgkin's disease and development of prognostic index. Lancet. 1985;1:967-72.
60. **Harris KR, Digard N, Gosling DC, Tate DG, Campbell MJ, Gardner B, et al.** Azathioprine and cyclosporin: different tissue matching criteria needed? Lancet. 1985;2:802-5.
61. **Geresh BJ, Kronmal RA, Schaff HV, Frye RL, Ryan TJ, Mock MB, et al.** Comparison of coronary artery bypass surgery and medical therapy in patients 65 years of age or older. N Engl J Med. 1985; 313:217-24.
62. **The EC/IC Bypass Study Group.** Failure of extracranial-intracranial arterial bypass to reduce the risk of ischemic stroke. N Engl J Med. 1985;313:1191-200.
63. **Cairns JA, Gent M, Singer J, Finnie KJ, Froggatt GM, Holder DA, et al.** Aspirin, sulfinpyrazone, or both in unstable angina. N Engl J Med. 1985;313:1369-75.
64. **Holmberg LH, Tabar L, Adami HO, Bergstrom R.** Survival in breast cancer diagnosed between mammographic screening examinations. Lancet. 1986;2:27-30.
65. **Viste A, Bjornestad E, Opheim P, Skarstein A, Thunold J, Hartveit F, et al.** Risk of carcinoma following gastric operations for benign disease. Lancet. 1986;2:502-5.
66. **Amery A, Birkenhager W, Brixko R, Bulpitt C, Clement D, Deruyttere M, et al.** Efficacy of antihypertensive drug treatment according to age, sex, blood pressure, and previous cardiovascular disease in patients over the age of 60. Lancet. 1986;2:589-92.
67. **Brain Resuscitation Clinical Trial I Study Group.** Randomized clinical study of thiopental loading in comatose survivors of cardiac arrest. N Engl J Med. 1986;314:397-403.
68. **Storb R, Deeg HJ, Whitehead J, Applebaum F, Beatty P, Bensinger W, et al.** Methotrexate and cyclosporine compared with cyclosporine alone for prophylaxis of acute graft versus host disease after marrow transplantation for leukemia. N Engl J Med. 1986;314:729-35.
69. **Ettinger B, Tang A, Citron JT, Livermore B, Williams T.** Randomized trial of allopurinol in the prevention of calcium oxalate calculi. N Engl J Med. 1986;315:1386-9.
70. **Flegel KM, Shipley MJ, Rose G.** Risk of stroke in non-rheumatic atrial fibrillation. Lancet. 1987;1:526-9.
71. **Riou G, Barrois M, Le MG, George M, Le Doussal V, Haie C.** C-myc proto-oncogene expression and prognosis in early carcinoma of the uterine cervix. Lancet. 1987;1:761-3.
72. **Freedman LS, Parkinson MC, Jones WG, Oliver RT, Peckham MJ, Read G, et al.** Histopathology in the prediction of relapse of patients with stage I testicular teratoma treated by orchidectomy alone. Lancet. 1987;2:294-8.
73. **Philip T, Armitage JO, Spitzer G, Chauvin F, Jagannath S, Cahn JY, et al.** High-dose therapy and autologous bone marrow transplantation after failure of conventional chemotherapy in adults with intermediate-grade or high-grade non-Hodgkin's lymphoma. N Engl J Med. 1987;316:1493-8.
74. **Willett WC, Green A, Stampfer MJ, Speizer FE, Colditz GA, Rosner B, et al.** Relative and absolute excess risks of coronary heart disease among women who smoke cigarettes. N Engl J Med. 1987; 317:1303-9.
75. **Coates A, Gebski V, Bishop JF, Jeal PN, Woods RL, Snyder R, et al.** Improving the quality of life during chemotherapy for advanced breast cancer. N Engl J Med. 1987;317:1490-5.
76. **AIMS Trial Study Group.** Effect of intravenous APSAC on mortality after acute myocardial infarction: preliminary report of a placebo-controlled clinical trial. Lancet. 1988;1:545-9.
77. **Gill M, McCarthy M, Murrells T, Silcocks P.** Chemotherapy for the primary treatment of osteosarcoma: population effectiveness over 20 years. Lancet. 1988;1:689-92.
78. **Holbrook TL, Barrett-Connor E, Wingard DL.** Dietary calcium and

risk of hip fracture: 14-year prospective population study. Lancet. 1988;2:1046-9.
79. **Ragland DR, Brand RJ.** Type A behavior and mortality from coronary heart disease. N Engl J Med. 1988;318:65-9.
80. **Gelmers HJ, Gorter K, de Weerdt CJ, Wiezer HJ.** A controlled trial of nimodipine in acute ischemic stroke. N Engl J Med. 1988;318: 203-7.
81. **Ekelund LG, Haskell WL, Johnson JL, Whaley FS, Criqui MH, Sheps DS.** Physical fitness as a predictor of cardiovascular mortality in asymptomatic North American men. N Engl J Med. 1988;319: 1379-84.
82. International Bone Marrow Transplant Registry. Effect of methotrexate on relapse after bone-marrow transplantation for acute lymphoblastic leukaemia. Lancet. 1989;1:535-7.
83. **Barker DJ, Winter PD, Osmond C, Margetts B, Simmons SJ.** Weight in infancy and death from ischaemic heart disease. Lancet. 1989;2: 577-80.
84. **Spyratos F, Maudelonde T, Bouillet JP, Brunet M, Defrenne A, Andrieu C, et al.** Cathepsin D: an independent prognostic factor for metastasis of breast cancer. Lancet. 1989;2:1115-8.
85. **Fogelfeld L, Wiviott MB, Shore-Freedman E, Blend M, Bekerman C, Pinsky S, et al.** Recurrence of thyroid nodules after surgical removal in patients irradiated in childhood for benign conditions. N Engl J Med. 1989;320:835-40.
86. **Justice AC, Feinstein AR, Wells CK.** A new prognostic staging system for the acquired immunodeficiency syndrome. N Engl J Med. 1989;320:1388-93.
87. **Scott GB, Hutto C, Makuch RW, Mastrucci MT, O'Connor T, Mitchell CD, et al.** Survival in children with perinatally acquired human immunodeficiency virus type 1 infection. N Engl J Med. 1989;321:1791-6.

As to your method of work, I have a single bit of advice, which I give with the earnest conviction of its paramount influence in any success which may have attended my efforts in life—Take no thought for the morrow. Live neither in the past nor in the future, but let each day's work absorb your entire energies, and satisfy your widest ambition. This was the singular but very wise answer which Cromwell gave to Bellevire—"No one rises so high as he who knows not wither he is going," and there is much truth in it. The student who is worrying about the future, anxious over the examinations, doubting his fitness for the profession, is certain not to do so well as the man who cares for nothing but the matter in hand, and who knows not wither he is going!

William Osler, to his students
The Practical Cogitator
Dell, New York, 1962

Submissions from readers are welcomed. If the quotation is published, the sender's name will be acknowledged. Please include a complete citation, as done for any reference.—*The Editors*