51, 1965
4. Edwards JH: Familial predispositions in man. Br Med Bull **25:** 58, 1969
5. Morton NE, Yee S, Elston RC, Lew R: Discontinuity and quasi-continuity: alternative hypotheses of multifactorial inheritance. Clin Genet **1:** 81, 1970
6. Nora JJ, Nora AH: The evolution of specific genetic and environmental counseling in congenital heart disease. Circulation **57:** 205, 1978
7. Goldstein JL, Schrott HG, Hazzard WR, Bierman EL, Motulsky AG: Hyperlipidemia in coronary heart disease. II. Genetic levels analysis of lipid levels in 176 families a delineation of a new inherited disorder, combined hyperlipidemia. J Clin Invest **52:** 1544, 1973
8. Archer PG: The predictive value of clinical laboratory test results. Am J Clin Pathol **69:** 32, 1978
9. Pyorala K: Life habits — nutrition, physical activity, and non-smoking — in primary prevention of ischemic heart disease. Abstracts-1, Seventh World Congress of Cardiology, Tokyo, 1978, p 19

Clinical Judgment and Statistics

Lessons from a Simulated Randomized Trial in Coronary Artery Disease

KERRY L. LEE, PH.D., J. FREDERICK MCNEER, M.D., C. FRANK STARMER, PH.D., PHILIP J. HARRIS, M.B., D.PHIL., AND ROBERT A. ROSATI, M.D.

SUMMARY **A simulated randomized clinical trial in coronary artery disease was conducted to illustrate the need for clinical judgment and modern statistical methods in assessing therapeutic claims in studies of complex diseases. Clinicians should be aware of problems that occur when a patient sample is subdivided and treatment effects are assessed within multiple prognostic categories. In this example, 1073 consecutive, medically treated coronary artery disease patients from the Duke University data bank were randomized into two groups. The groups were reasonably comparable and, as expected, there was no overall difference in survival. In a subgroup of 397 patients characterized by three-vessel disease and an abnormal left ventricular contraction, however, survival of group 1 patients was significantly different from that of group 2 patients. Multivariable adjustment procedures revealed that the difference resulted from the combined effect of small imbalances in the distribution of several prognostic factors. Another subgroup was identified in which a significant survival difference was not explained by multivariable methods.**

These are not unlikely examples in trials of a complex disease. Clinicians must exercise careful judgment in attributing such results to an efficacious therapy, as they may be due to chance or to inadequate baseline comparability of the groups.

Downloaded from http://ahajournals.org by on February 7, 2021

CLINICAL JUDGMENT in chronic illness involves a knowledge of the natural history of the disease, the ability to assess the validity of therapeutic claims and a means of applying what is known to the individual patient. Much has been written about the natural history of angina pectoris.[1-10] The literature has emphasized the importance of multiple factors in determining outcome and of the heterogeneity of patients grouped together under the diagnosis of angina pectoris. This complexity makes the clinician's task of interpreting therapeutic claims particularly difficult. In this paper, a simulated randomized clinical trial in patients with coronary artery disease is presented to illustrate the effects that the heterogeneity of patients with angina pectoris may have on the results of clinical experiments. The role of statistics and clinical judgment in solving the problems encountered is explored. The randomized design is used because it represents the ideal method of treatment allocation in clinical experiments. The problems and the approaches to their solutions apply equally well to both randomized and nonrandomized clinical studies.

From the Department of Community and Family Medicine and the Division of Cardiology, Department of Medicine, Duke University Medical Center, Durham, North Carolina.

Supported by contract HRA-230-76-0300 from the Health Resources Administration, HEW, research grant HL-17670 from the NHLBI, training grant LM-07003 from the National Library of Medicine, and a grant from the Prudential Insurance Company of America.

Dr. Lee is the recipient of Research Career Development award LM-0042 from the National Library of Medicine.

Address for correspondence: Robert A. Rosati, M.D., Box 3337, Duke University Medical Center, Durham, North Carolina 27710.

Received January 17, 1979; revision accepted September 4, 1979.
Circulation 61, No. 3, 1980.

Methods

The subjects in this simulated trial consisted of the medically treated patients with angiographically proved coronary artery disease contained in the Duke University data bank. The data bank has been described previously.[11] Briefly, baseline historical, physical examination, laboratory, electrocardiographic, stress test, chest x-ray, angiographic data and follow-up information on 1073 medically treated

patients with significant coronary artery disease formed the basis of this study. Significant disease was defined as at least a 75% occlusion of one or more coronary arteries. This population represents all consecutive patients with coronary artery disease who were catheterized and treated medically at Duke University Medical Center between August 1969 and January 1977. The baseline characteristics were carefully collected and entered into the data bank during each patient's initial hospital evaluation.

These patients and their baseline and follow-up data were used to simulate a randomized clinical trial. The patients were randomized into two groups, designated "treatment" group 1 and "treatment" group 2. In this experiment, the patients received the same treatment, but the two groups would correspond to different therapeutic interventions in an actual clinical trial. The experimental design was similar to that used in the VA Cooperative Study of coronary artery bypass surgery.[12] The 1073 patients were stratified (blocked) according to age (under 50 years and 50 years and older) and the number of significantly diseased vessels (one, two or three). Randomization to group 1 or group 2 was performed separately within each of the six prognostic strata. In an actual trial, randomization would be performed at the time of patient entry into the study. This was mimicked in the simulated trial by allocating patients according to the date of cardiac catheterization. Well-described balancing methods were used to ensure relative chronologic balance between "treatments."[13] Random numbers were generated using an automated shift-register, random-number generator.[14]

Group 1 and group 2 were compared with respect to the distribution of baseline characteristics in order to determine whether equalization of baseline abnormalities had been achieved. The chi-square test was used in the case of discrete variables and the *t* test for two independent samples was used for continuous variables. These univariate methods compare the distribution of single characteristics between treatment groups.

Life-table survival rates[15] were calculated and compared using categorical data methods.[16] Multivariable methods, including the Cox survival model[17] and the Mantel-Haenszel (or log-rank) test for survivorship,[18, 19] were used 1) to identify baseline characteristics that were important predictors of survival, and 2) to perform survival comparisons that were adjusted or corrected for the effects of multiple baseline characteristics. Further details concerning the multivariable methods are provided in an appendix.

After comparing the overall survival in group 1 vs group 2 patients, selected subgroups of patients from both groups were examined. The subgroups were defined using combinations of prognostically important baseline characteristics such as the number of significantly diseased vessels and left ventricular contraction pattern (interpreted as normal or abnormal, where abnormal indicated the presence of one or more localized areas of asynergy). Baseline comparability, survival and the effects of the distribution of prognostically important baseline characteristics on survival were examined in the subgroups using the same methods detailed above.

Results

Selected baseline characteristics of the patients in this experiment are presented in table 1. Univariate statistical comparisons of the variables yielded no significant differences between groups 1 and 2 except in left ventricular contraction pattern.

As expected, yearly life-table survival rates of the randomized groups 1 and 2 were similar: 89.7% and 89.2% at 1 year, 76.9% and 79.1% at 3 years and 69.3% and 72.0% at 5 years. All multivariable methods indicated that the overall survival of group 1 and group 2 patients was consistent with the distribution of baseline characteristics and was not affected by the "treatment" (i.e., randomization). Survival rates for each of the strata (defined by number of diseased vessels and age) within which the patients were randomized are given in table 2.

The prognostic significance of the baseline variables in table 1 was assessed by the Cox model,[17] and the patient population was stratified with respect to two prognostically important descriptors: the number of diseased coronary vessels (one, two or three) and left ventricular contraction pattern (normal or abnormal). In one of the resulting six subgroups, a statistically

TABLE 1. *Distribution of Selected Baseline Characteristics*

	Prevalence (%)		
	Total (n = 1073)	Group 1 (n = 539)	Group 2 (n = 534)
Males	85	84	85
Age > 50 years	53	53	53
History of previous MI	51	49	53
History of CHF	14	14	14
Cardiomegaly on chest x-ray	20	18	22
Diagnostic Q waves on ECG	43	41	46
Resting ST-T-wave abnormalities	47	47	47
LVEDP > 18 mm Hg	15	14	15
AVO_2D > 5.5 vol %	19	19	18
Single-vessel disease	24	24	24
Three-vessel disease	51	51	51
Abnormal LV contraction	60	57	63
Significant mitral insufficiency	8	6	10
Left main stenosis ≥ 50%	16	17	15

Abbreviations: MI = myocardial infarction; CHF = congestive heart failure; LVEDP = left ventricular end-diastolic pressure; AVO_2D = arteriovenous oxygen difference; LV = left ventricular.

Downloaded from http://ahajournals.org by on February 7, 2021

TABLE 2. *Life-table Survival Data*

Stratum*				Survival (%)		
Diseased vessels	Age (years)	Group	n	1 year	3 year	5 year
1	< 50	1	70	94	92	84
		2	70	97	95	92
1	≥ 50	1	59	98	93	84
		2	57	94	94	88
2	< 50	1	64	98	93	90
		2	64	95	91	91
2	≥ 50	1	73	94	81	71
			73	90	78	75
3	< 50	1	95	83	75	66
		2	93	85	79	70
3	≥ 50	1	178	83	59	48
		2	177	84	69	60

*These are the strata within which the patients were randomized.

significant difference in survival was found between group 1 and group 2 patients ($p < 0.025$). This subgroup was composed of 397 patients: 194 group 1 and 203 group 2 patients with three-vessel disease and an abnormal left ventricular contraction pattern. In figure 1, the long-term survival of group 1 patients in this subgroup is markedly lower at 3 years and beyond than that of group 2 patients.

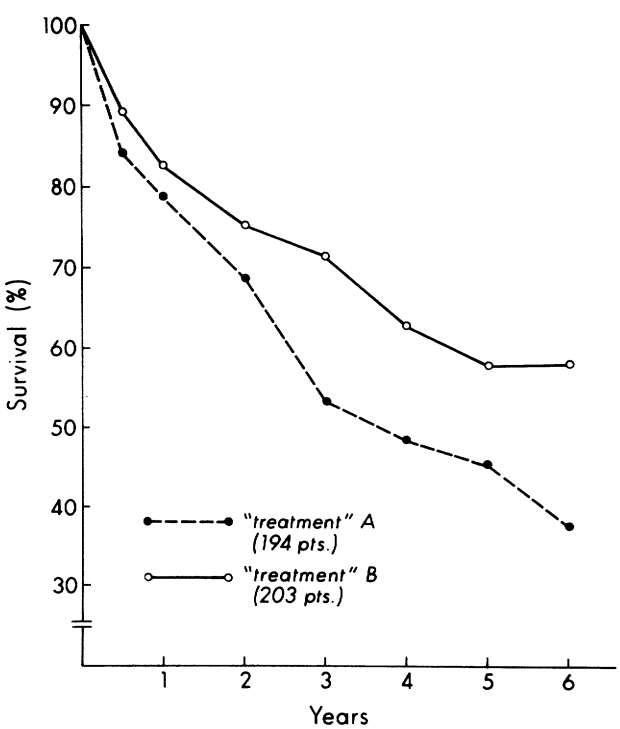

FIGURE 1. *Comparison of long-term survival in a subgroup composed of 194 group 1 patients ("treatment" A) and 203 group 2 patients ("treatment" B) with three-vessel disease and an abnormal left ventricular contraction pattern.*

The prevalence of prognostically important baseline abnormalities in this subgroup is shown in table 3 for each group. The percentage distribution suggests that group 1 patients are slightly sicker (the prevalence of each abnormal condition is generally higher), even though univariate statistical comparisons of these variables fail to reveal significant baseline differences between groups.

Multivariable methods were used to determine if the distribution of prognostically important baseline characteristics rather than the "treatment" might have caused the survival differences in this subgroup of 397 patients. Based on the survival-modeling method of Cox,[17] the prognostic significance of the factors in table 3 is shown in table 4. The entry in column 1 for each variable is a chi-square statistic whose magnitude indicates the relative prognostic importance of each characteristic considered individually. On a univariate basis, the "treatment" (randomization) appeared to be a significant factor in the survival of this subgroup ($\chi^2 = 5.4$; $p < 0.025$). When the variables were considered jointly (column 2, table 4), the treatment effect became nonsignificant ($\chi^2 = 2.4$; p = NS). Similar results were obtained with the Mantel-Haenszel test for survivorship data.[18, 19] That is, the apparent "treatment" difference was explained by the combined effect of small inequalities among several prognostically important baseline characteristics.

Additional stratification based on a third prognostic variable, history of congestive heart failure (yes or no), was also done. In the subgroup consisting of 143 group 1 and 155 group 2 patients with three-vessel disease, an abnormal left ventricular contraction pattern and no history of congestive heart failure, the significance of the two-group "treatment" comparison was even more dramatic ($\chi^2 = 10.0$; $p < 0.01$). Three-year survival rates were 60% and 80% in groups 1 and 2, respectively.* Again, the prevalence of other prognostically important baseline abnormalities in this subgroup appeared slightly higher in group 1 than in group 2 patients. In this case, however, the significance of the observed difference in survival between the two groups persisted after adjusting for multiple prognostic characteristics. (After adjusting for the variables in table 3, the chi-square for "treatment" was 9.3, $p < 0.01$.)

Discussion

This simulated randomized clinical trial was performed to illustrate problems that may arise in assess-

*The Cox and Mantel-Haenszel statistics take into account the overall structure of the survival curves being compared, not just their values at one time point. The pattern of deaths in this particular subset explains the more significant "treatment" comparison. Twice as many deaths (22 vs 11) occurred during the first year in group 1 compared with group 2; this dramatic early difference was maintained throughout the follow-up period. By contrast, in the 397 patients in the larger subgroup, the pattern of deaths did not cause a marked separation in the survival curves until the third year (see fig. 1).

Downloaded from http://ahajournals.org by on February 7, 2021

TABLE 3. *Comparison of Baseline Characteristics: Subgroup with Three-vessel Disease and an Abnormally Contracting Ventricle*

	Prevalence (%)	
	Group 1 (n = 194)	Group 2 (n = 203)
History of CHF	26	24
Cardiomegaly on chest x-ray	33	33
Resting ST-T-wave abnormalities	67	58
Significant mitral insufficiency	13	18
$AVO_2D > 5.5$ vol %	31	24
LV diffusely abnormal contraction	32	24
Left main stenosis $\geq 50\%$	20	19

Abbreviations: CHF = congestive heart failure; AVO_2D = arteriovenous oxygen difference; LV = left ventricular.

ing therapeutic effects when seeking information relevant to specific patients from a study that contains patients with various prognoses. The problems of subgrouping patients according to prognostic descriptors and performing treatment comparisons within some or all of the subgroups are well known to biostatisticians, and to many clinicians on a theoretical basis, but have not received adequate practical attention in real clinical trials.

Many randomized trials in chronic disease are prospectively designed to answer a single, usually rather broad question. At completion, such trials are frequently best known for the answers to questions that were asked retrospectively in particular patient subgroups. For example, in the University Group Diabetes Program, a prospective study was undertaken to determine whether various treatments would affect the development of vascular complications in a selected group of diabetic patients.[20] The controversial claim from this study that in certain patients tolbutamide increased the risk of cardiovascular deaths[21, 22] is well known. The Veterans Administration Cooperative Study of Surgery for Coronary Arterial Occlusive Disease was prospectively designed to investigate whether coronary bypass surgery affected survival in patients with coronary artery disease.[12] Its claim that surgery improved survival in the subgroup of patients with left main coronary artery disease has greatly influenced treatment decisions.

TABLE 4. *Comparison of Individual and Joint Prognostic Significance of Baseline Variables**

	Individually	Jointly
Treatment	5.4	2.4
History of CHF	36.1	3.8
Cardiomegaly on chest x-ray	15.7	0.3
Resting ST-T-wave abnormalities	8.4	2.2
Mitral insufficiency	25.1	2.7
Arteriovenous oxygen difference	52.0	11.1
LV diffusely abnormal contraction	17.3	0.3
Left main stenosis	6.9	5.7

*Subgroup with three-vessel disease and an abnormally contracting ventricle (397 patients). Numerical entries are chi-square statistics with one degree of freedom. For values above 3.84, $p < 0.05$.

Abbreviations: CHF = congestive heart failure; LV = left ventricular.

Considering the tremendous investment of resources that goes into the conduct of a clinical trial, most investigators are not content with an answer to only one broad question. Furthermore, in many important studies of chronic disease, the state of knowledge at the initiation of the investigation (e.g., regarding prognostic factors) limits rigid prespecification of precisely which subgroups will be examined. Additional questions are justified because of the clinical variation in the patients and the relevance of subgroups to individual patient treatment.

In the overall sample in this simulated trial, the random allocation of patients achieved a reasonably balanced distribution of baseline risk factors. The overall survival of group 1 was not different from the survival of group 2 at the 0.05 level of significance. This was the expected result because the "treatment" was randomization. A significant overall difference in survival in such an experiment would rarely occur by chance (i.e., if testing at the 0.05 level, five differences in 100 such experiments are expected).

What happens when additional questions, perhaps unanticipated in the initial study design, are asked of the data? In this simulated clinical trial, stratifying the patients on the basis of two prognostically important variables — the number of diseased vessels and left ventricular contraction pattern — produced a subgroup of 397 patients in which group 1 patients had significantly different long-term survival from group 2 patients. With the addition of a third stratifying variable — history of congestive heart failure — treatment differences were even more dramatic in the 298 patients with three-vessel disease, abnormal ventricular contraction and no history of congestive heart failure.

Such differences are likely to arise.[23, 24] For example, if 10 independent tests are performed (e.g., comparisons in 10 nonoverlapping subgroups), each at the 0.05 level of significance, there is a 40% chance that one or more significant results will be observed even if the treatment is ineffective.* If 20 independent comparisons are made, the chances increase to 64%. These chance differences are most likely to appear in high-risk patients (e.g., patients with left main coronary artery disease, three-vessel disease or ventricular impairment). In such groups, death is a frequent occurrence, and chance maldistributions have a greater likelihood of being statistically significant. The important point is that as patients are subgrouped and treatment comparisons are performed in multiple classes of patients, not all observed differences in survival can be attributed to efficacious treatment.

*If k independent comparisons are performed, each at the 0.05 significance level, the probability of one or more significant results, even when there are no treatment differences, equals $1 - (0.95)^k$.

Downloaded from http://ahajournals.org by on February 7, 2021

There are at least three strategies for dealing with this problem. One approach is to ignore the results obtained in subgroups and use only the overall result. This seems acceptable, however, only if the interest is oriented toward population or public health concerns and an answer that applies "on the average" is satisfactory. For the doctor caring for a patient, this solution is clinically unsatisfactory. He or she needs information about prognosis in a subgroup of patients with clinical manifestations similar to those of the patient under treatment. The alternative approach of performing trials in many subgroups of patients is impractical and the possibility that multiple comparisons would yield one or more significant results by chance still remains.

Careful statistical interpretation is helpful. On the simplest level, one can apply a rule derived from a relationship that statisticians refer to as the Bonferroni inequality.[25] If the investigator can specify in advance the number of subgroups to be examined, a guideline for accepting a particular comparison as significant is whether the p value is less than 0.05 divided by the number of subgroups.[26]* For example, in the simulated trial, the p value would have to be 0.05 divided by 18 (6 + 12 subgroups) or approximately 0.003. The comparison in the first subgroup does not meet this criterion for significance. The p value in the subgroup with three-vessel disease, abnormal ventricular contraction and no history of congestive heart failure does meet this criterion. Thus, while this guideline is helpful, its use does not completely protect the study from an incorrect conclusion. In addition, in a complex disease where there is considerable variation in the patients, the number of comparisons will probably be large. The required p value may be nearly impossible to reach, particularly because comparisons of different outcomes adds to the multiplicity in the same way as examining additional subgroups.[26]

A relatively common preliminary step in assessing therapeutic effects has been to compare the distribution of prognostic baseline variables between treatment groups to assure that the groups are similar. If each variable appears equally distributed between groups (that is, differences are not statistically significant), the treatment groups are claimed to be comparable. As our experiment indicates, univariate comparisons alone do not necessarily reveal baseline differences between treatment groups that can affect the assessment of treatment effects. With modern multivariable analysis techniques, the significance of observed differences in outcome may possibly be explained by the combined effects of several statistically insignificant differences between groups in the important prognostic variables. Multivariable statistical techniques such as the Cox model[17] and the Mantel-Haenszel test for survivorship[18, 19] permit survival comparisons that are adjusted for the effects of multiple baseline characteristics. This is mandatory in a complex setting such as coronary artery disease. In addition, these methods include all patients in the analysis regardless of the length of follow-up, and use information on when patients died rather than simply the number of deaths. As Peto et al. pointed out, this allows a more informative, sensitive assessment of the value of each treatment.[23, 24]

From the results of this simulated trial, one might attempt to explain the differences in outcome between patients from groups 1 and 2 in the second subgroup by postulating differences in other baseline characteristics. Indeed, one may legitimately ask whether a more quantitative measure of ventricular function (e.g., ejection fraction) or information about the actual medical treatment received may have explained the difference in survival. Recognizing that random patient allocation does not guarantee equal distribution of all risk factors, particularly in subgroups selected for examination after the study was designed and data were collected, there are undoubtedly factors that might explain many such differences in outcome. Despite sophisticated statistical techniques, however, when many subgroups are analyzed, there will be differences in outcome that cannot be explained by baseline differences of known variables. In such cases there are four possible explanations: 1) the patients were inadequately characterized; 2) the patients were adequately characterized but the statistical methods were inadequate to properly relate baseline differences to outcome; 3) the treatment was effective; or 4) the difference was due to chance.

In assessing therapeutic claims, the doctor must consider the adequacy of the experimental design and conduct of the study, the adequacy of the analysis and presentation of results, the strength of the conclusion and how the findings relate to his clinical experience and to clinical knowledge. The essential components of experimental design have been reviewed.[27] We and others have discussed modern multivariable analysis techniques.[23, 24] Regarding strength of the conclusion, one will normally have more confidence in results significant at the 0.0001 level than in those that achieve only the 0.05 level. The consistency of therapeutic effects over time (e.g., as observed by subdividing the follow-up period and analyzing the data within more than one time interval) will also strengthen the conclusion.[24]

If satisfied with the adequacy of the design and analysis, the clinician must weigh results of a clinical trial with what his training and experience could reasonably lead him to expect. For example, the clinician might expect that treatment would affect more than one manifestation of the target disease. In coronary artery disease, a therapy that reduces the frequency of anginal pain and death might be expected to reduce the frequency of nonfatal myocardial infarction as well. A therapy that increases survival in left main coronary artery disease might also be expected to improve survival, though perhaps less dramatically,

*If k is the number of subgroups to be examined, requiring the p value of each comparison to be less than α/k will assure an overall level of significance for the k tests that is no greater than α, regardless of whether the tests are independent.

Downloaded from http://ahajournals.org by on February 7, 2021

in three-vessel disease. Such expectations are based on clinical experience and biologic principles. Finally, results of different studies that are designed and analyzed adequately should be consistent. Frequently, new interpretations, especially if based on the analysis of subgroups, should be treated as tentative until confirmed by other studies.

The problems faced by the clinician in assessing therapeutic claims is well illustrated by studies of the efficacy of aortocoronary bypass surgery. Many studies use medical controls whose comparability to the surgical series has not been or cannot be determined.[28-30] Because differences in survival can arise so easily in a simulated randomized trial, one should not readily accept the results of a study with unmatched controls whose baseline data are insufficient for adequate statistical analysis. In the interpretation of three studies that do have fairly reasonable controls, careful assessment is still necessary. The VA Cooperative Study is a randomized trial.[31] The Duke[7] and the Seattle[32] groups have studied concurrent operated and nonoperated series that were not randomized. These three studies have reached some consistent conclusions. Their reports indicate that aortocoronary bypass surgery did not prolong 4–5-year survival when medically and surgically treated patients were compared overall. All reported a subgroup in which surgery did appear to improve survival. Each subgroup contained relatively high-risk patients: left main coronary artery disease,[33] three-vessel disease with abnormal ventricular contraction but normal arteriovenous oxygen difference (subgroup F),[7] and two-vessel disease with moderately reduced ejection fraction.[32] None of the studies confirmed the differences found in others' subgroups. All claimed significance with only moderately low p values (0.02, 0.028 and 0.03). Multivariable methods were not applied to these subgroups to adjust for possible baseline differences.

In the study at Duke, continued observations of outcome in the patients with three-vessel disease, abnormal ventricular contraction and normal arteriovenous oxygen differences failed to confirm the original findings.[34] For the left main coronary artery disease subgroup of the VA Cooperative Study,[33] the prevalence of several prognostic factors in table 5 suggests that the medically treated group was slightly sicker. None of the baseline differences taken individually achieve statistical significance in these small sample sizes. If there had been 200 patients on each treatment, and the proportions in table 5 remained the same, the medically treated group would have been statistically different (at greater risk) from the surgically treated group with respect to the proportions of patients having an ejection fraction less than 50%, right coronary involvement and severe left ventricular dysfunction. Further discussions of the VA study and randomized clinical trials appear in a recently published symposium debate.[35-37]

The clinician must attempt to decide rationally if the differences reported in these and other subgroups are chance occurrences or whether they are actually due to aortocoronary bypass surgery. Frequently, this is not a straightforward judgment. As our simulated study makes clear, caution is necessary, even in randomized clinical trials. In today's studies of therapeutic effects in a complex chronic disease, investigators generally accept the importance of complete, carefully collected baseline data, a comparable control group, and physiologic rationale for treatment effects. Other points such as the likelihood of chance differences achieving statistical significance when multiple subgroups are analyzed and reanalyzed, or the influence on outcome comparisons of the combined effect of several relatively small differences in baseline prognostic variables may not be so obvious. These points are often ignored or not mentioned in the presentation of results of clinical research. Clinical studies must continue to add to our understanding of diseases such as coronary artery disease and to provide information relative to the care of individual patients. We hope that the information presented here will enable clinicians to make a more careful and informed evaluation and interpretation of these studies.

TABLE 5. *Distribution of Selected Prognostic Factors (VA Coronary Arterial Disease Study: Left Main Subgroup)*

		Prevalence (%)	
		Medicine (n = 53)	Surgery (n = 60)
Left main stenosis ≥ 70%		66	57
No. vessels involved with	One	15	10
significant disease in	Two	23	35
addition to left main	Three	62	53
Right coronary involvement		90	78
Abnormal LV size		24.5	26.7
Abnormal LV contractility		45.3	51.7
Ejection fraction < 50%		28.3	16.7
Severe LV dysfunction (score ≥ 5)		28	18

Abbreviation: LV = left ventricular.

Acknowledgment

The authors thank J. Scott Wells for his assistance with the computer programming and numerical computations required in this study and Judie Armstrong for her assistance in preparing the manuscript.

References

1. Richards DW, Bland EF, White PD: A completed twenty-five year follow-up study of 456 patients with angina pectoris. J Chronic Dis **4:** 423, 1956
2. Weinblatt E, Shapiro S, Frank CW, Sager RV: Prognosis of men after first myocardial infarction: mortality and first recurrence in relation to selected parameters. Am J Pub Health **58:** 1329, 1968
3. Weinblatt E, Shapiro S, Frank CW: Prognosis of women with newly diagnosed coronary heart disease — a comparison with course of disease among men. Am J Pub Health **63:** 577, 1973
4. Kannel WB, Feinleib M: Natural history of angina pectoris in the Framingham Study. Am J Cardiol **29:** 154, 1972
5. Friesinger GC, Page EE, Ross RS: Prognostic significance of coronary arteriography. Trans Assoc Am Physicians **83:** 78, 1970
6. Humphries JO, Kuller L, Ross RS, Friesinger GC, Page EE:

Downloaded from http://ahajournals.org by on February 7, 2021

Natural history of ischemic heart disease in relation to arteriographic findings: a twelve-year study of 224 patients. Circulation **49:** 489, 1974
7. McNeer JF, Starmer CF, Bartel AG, Behar VS, Kong Y, Peter RH, Rosati RA: The nature of treatment selection in coronary artery disease. Experience with medical and surgical treatment of a chronic disease. Circulation **49:** 606, 1974
8. Oberman A, Jones WB, Riley CP, Reeves TJ, Sheffield LT, Turner ME: Natural history of coronary artery disease. Bull NY Acad Med **48:** 1109, 1972
9. Burggraff GW, Parker JO: Prognosis in coronary artery disease. Angiographic, hemodynamic and clinical factors. Circulation **51:** 146, 1975
10. McNeer JF, Margolis JR, Lee KL, Kisslo JA, Peter RH, Kong Y, Behar VS, Wallace AG, McCants CB, Rosati RA: The role of the exercise test in the evaluation of patients for ischemic heart disease. Circulation **57:** 64, 1978
11. Rosati RA, McNeer JF, Starmer CF, Mittler BS, Morris JJ, Wallace AG: A new information system for medical practice. Arch Intern Med **135:** 1017, 1975
12. Detre KM, Hultgren HN, Takaro T: Veterans Administration cooperative study of surgery for coronary arterial disease. III. Methods and baseline characteristics, including experience with medical treatment. Am J Cardiol **40:** 212, 1977
13. Zelen M: The randomization and stratification of patients to clinical trials. J Chronic Dis **27:** 365, 1974
14. Tausworthe RC: Random numbers generated by linear recurrence modulo two. Math Comput **19:** 201, 1965
15. Cutler SJ, Ederer F: Maximum utilization of the life table method in analyzing survival. J Chronic Dis **8:** 699, 1958
16. Koch GG, Johnson WD, Tolley HD: A linear models approach to the analysis of survival and extent of disease in multidimensional contingency tables. J Am Stat Assoc **67:** 783, 1972
17. Cox DR: Regression models and life tables (with discussion). J R Stat Soc (Series B) **34:** 187, 1972
18. Mantel N: Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemother Rep **50:** 163, 1966
19. Breslow NE: Analysis of survival data under the proportional hazards model. Int Stat Rev **43:** 45, 1975
20. University Group Diabetes Program: A study of the effects of hypoglycemic agents on vascular complications in patients with adult-onset diabetes. I. Design, methods and baseline results. Diabetes **19** (suppl 2): 747, 1970
21. University Group Diabetes Program: A study of the effects of hypoglycemic agents on vascular complications in patients with adult-onset diabetes. II. Mortality results. Diabetes **19** (suppl 2): 787, 1970
22. Gilbert JP, Meier P, Rumke CL, Saracci R, Zelen M, White C: Report of the committee for the assessment of biometric aspects of controlled trials of hypoglycemic agents. JAMA **231:** 583, 1975
23. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Mantel N, McPherson K, Peto J, Smith PG: Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. Br J Cancer **34:** 585, 1976
24. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Mantel N, McPherson K, Peto J, Smith PG: Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. Br J Cancer **35:** 1, 1977
25. David HA: Order Statistics. New York, John Wiley & Sons, 1970
26. Tukey JW: Some thoughts on clinical trials, especially problems of multiplicity. Science **198:** 679, 1977
27. Feinstein AR: Clinical Biostatistics. St. Louis, CV Mosby Co, 1977
28. Sheldon WC, Rincon G, Effler DB, Proudfit WL, Sones FM Jr: Vein graft surgery for coronary artery disease. Survival and angiographic results in 1,000 patients. Circulation **47** (suppl 3): 184, 1973
29. Spencer FC, Isom OW, Glassman E, Boyd AD, Engleman RM, Reed GE, Pasternack BS, Dembrow JM: The long-term influence of coronary bypass grafts on myocardial infarction and survival. Ann Surg **180:** 439, 1974
30. Kaiser GC, Barner HB, Tyras DH, Codd JE, Mudd JG, Willman VL: Myocardial revascularization: a rebuttal of the cooperative study. Ann Surg **188:** 331, 1978
31. Murphy ML, Hultgren HN, Detre K, Thomsen J, Takaro T: Treatment of chronic stable angina. A preliminary report of survival data of the randomized Veterans Administration Cooperative Study. N Engl J Med **297:** 621, 1977
32. Hammermeister KE, DeRouen TA, Murray JA, Dodge HT: Effect of aortocoronary saphenous vein bypass grafting on death and sudden death. Comparison of nonrandomized medically and surgically treated cohorts with comparable coronary artery disease and left ventricular function. Am J Cardiol **39:** 925, 1977
33. Takaro T, Hultgren HN, Lipton MJ, Detre, KM: The Veterans Administration cooperative randomized study of surgery for coronary arterial occlusive disease. II. Subgroup with significant left main lesions. Circulation **54** (suppl 3): III-107, 1976
34. Rosati RA, Lee KL, Mittler BS, McNeer JF, Behar VS, Margolis JR: Does coronary surgery prolong life in comparison with medical management? Postgrad Med J **52:** 749, 1976
35. Chalmers TC, Smith H, Ambroz A, Reitman D, Schroeder B: In defense of the VA randomized control trial of coronary artery surgery. Clin Res **26:** 230, 1978
36. Proudfit WL: Criticisms of the VA randomized study of coronary bypass surgery. Clin Res **26:** 236, 1978
37. Feinstein AR: The scientific and clinical tribulations of randomized clinical trials. Clin Res **26:** 241, 1978

Appendix

The Cox model,[17] an analog of the multiple linear regression model, was developed for the analysis of survival data. Unlike methods that consider survival at isolated time points, the Cox model takes into account the overall structure of the survival curves under study. It can accommodate varying lengths of patient follow-up, and it uses information on the time to death (i.e., the order in which the deaths occur) rather than simply enumerating the number of deaths. With this model, relationships can be studied between an outcome (response) variable and multiple patient baseline characteristics. The outcome variable of primary interest in most clinical follow-up studies is patient survival time; however, the time to end points other than death, such as the first myocardial infarction, can also be studied. The baseline patient descriptors may be either continuous or discrete measurements. Each baseline descriptor can be considered individually to determine whether it has a significant univariate relationship with survival. More important, however, multiple baseline characteristics can be analyzed jointly to determine which descriptors contribute independent prognostic information after adjustment for the effects of other variables. In an analysis to determine whether therapy is significantly related to survival, a treatment variable is included as one of the baseline characteristics. Therapeutic effects can thus be assessed without adjustment and with adjustment for the effects of other prognostic variables. The latter is of critical importance when there are multiple prognostic factors, some of which may be unequally distributed between treatment groups. Another advantage of the Cox model is that, in contrast to standard multiple regression procedures, it can accommodate "censored" survival times. Such observations arise because at the time of analysis, many patients are still alive, and the time to their deaths is known only to be greater than the current length of their follow-up. These patients are said to have a "censored" survival time.

A mathematical expression for the model is given by the equation

$$\log_e \left(\frac{\lambda_i(t)}{\lambda_o(t)} \right) = \beta_1 X_{i1} + \ldots + \beta_p X_{ip} \qquad \text{(A1)}$$

where $\lambda_i(t)$ is the hazard function at time t for the ith patient, $\lambda_o(t)$ is an arbitrary (unspecified) underlying hazard function, $X_{i1}, \ldots, X_{ip}$ are characteristics of the ith patient and the $\beta_1, \ldots \beta_p$ are regression coefficients estimated from the data. The hazard function is

Downloaded from http://ahajournals.org by on February 7, 2021

similar to a death rate. Mathematically, it is the probability that a subject will die within an arbitrarily small interval of time (t, t + Δt), given that the subject has survived to time t. An assumption of the Cox model is that the hazard functions for any two subjects are proportional over time. From equation (A1) a likelihood function has been derived, and in any particular application, the regression parameters are estimated by maximizing the value of the likelihood function. To assess whether a particular characteristic is a significant independent predictor of survival, a test of significance is performed on the corresponding regression coefficient to determine if it is significantly different from zero. When many variables must be analyzed, the Cox model can be fitted in a stepwise fashion so that the first variable included is the single variable most important in explaining the survival pattern of the patients under study; the second is the variable which, when added to the first, gives the best pair of prognostic variables, and so on. Documentation of an efficient computer program that we use for performing these calculations will be furnished upon request.

The Mantel-Haenszel test, also applicable with variable-length follow-up and censored observations, is based on combining information from a set of contingency tables (one table for each time a failure occurs). A clear expository explanation with clinical examples is provided in part two of the excellent two-part series of papers in the *British Journal of Cancer*.[23, 24]

Spectral Analysis of Arterial Bruits (Phonoangiography): Experimental Validation

ARNOLD MILLER, M.B.CH.B., ROBERT S. LEES, M.D., J. PHILIP KISTLER, M.D. AND WILLIAM M. ABBOTT, M.D.

SUMMARY Turbulent flow in arteries produces sound recognized at the skin surface as a bruit. Spectral analysis of such bruits (phonoangiography) is the basis for a simple, noninvasive method of quantifying arterial stenosis. In human studies of carotid stenoses, the spectral break frequency of the bruit (f_o) (frequency beyond which bruit amplitude drops sharply) was directly related to the angiographic residual lumen diameter (d), i.e., $d = U/f_o$, where U is flow velocity. In the clinical situation, flow velocity remains relatively constant because of cerebrovascular autoregulation. In order to test the effects of flow velocity on bruit frequency, we have correlated, under controlled conditions, stenosis anatomy, blood flow, and the sounds originating from an abdominal aortic stenosis produced in adult mongrel dogs by external application of a 5-mm wide Teflon band. Aortic flow was measured in arbitrary units with an electromagnetic flowmeter and varied by stepwise constriction of bilateral femoral arteriovenous fistulas. Bruits were recorded on tape and analyzed by computer.

The relationship between flow through the stenosis and break frequency of the bruit was linear ($r = 0.89$) in 10 dogs. Where d was altered in three other dogs, the relationship between flow and break frequency remained linear for each different d.

The data suggest that the relationship between break frequency, flow velocity and residual lumen diameter holds over a wide range of values of each of those variables.

Downloaded from http://ahajournals.org by on February 7, 2021

ARTERIAL STENOSIS, when severe, produces turbulent flow distal to the stenotic site.[1] This turbulence in turn produces pressure fluctuations that may be recorded as a bruit, either on the vessel wall or on the skin surface. The spectra of the turbulent pressure fluctuations have been shown to be similar to those of in vitro pipe turbulence.[1] We have devised a method of arterial bruit analysis based on the analysis of turbulent flow that allows estimation of the extent of arterial stenosis at the site of origin of the bruit.[2-5] This method is based on the algorithm $d = U/f_o$, where d is the residual lumen diameter at the stenosis, U is linear flow velocity and f_o is the characteristic break frequency of the bruit, the frequency beyond which amplitude falls off with increasing frequency.[3]

The algorithm, although applied with significant accuracy to evaluate the degree of carotid artery stenosis in humans, has not been tested under controlled conditions in experimental animals.

We present here studies in the dog of the relationships among blood flow velocity, extent of stenosis and the characteristics of the sound produced.

From the Noninvasive Diagnostic Laboratory, the Vascular Laboratory and the Departments of Surgery, Medicine and Neurology, Massachusetts General Hospital, Boston, Massachusetts, and the Arteriosclerosis Center, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Supported by grants from the Ambrose Monell and Charles E. Culpeper Foundation and by USPHS grant HL-06664.

Address for correspondence: Robert S. Lees, M.D., Arteriosclerosis Center, 40 Ames Street, Cambridge, Massachusetts 02142.

Received June 4, 1979; revision accepted September 24, 1979.
Circulation 61, No. 3, 1980.

Methods

Theoretical Background

Initial studies in our laboratory suggested a complex relationship between residual lumen diameter at a stenosis, flow velocity and the power spectrum of the sound produced,[1, 2] but further theoretical analysis[3] and a sizeable clinical experience[3-5] confirmed that the simple relationship $f_o d = SU$ accurately describes the production of sound by a vascular stenosis, where f_o is