METHODS

Computer-Assisted Diagnosis in the Noninvasive Evaluation of Patients With Suspected Coronary Artery Disease

GEORGE A. DIAMOND, MD, HOWARD M. STANILOFF, MD, FACC,

JAMES S. FORRESTER, MD, FACC, BRAD H. POLLOCK, MPH, H. J. C. SWAN, MD, PhD, FACC

Los Angeles, California

A microcomputer program called CADENZA, which employs Bayes' theorem to analyze and report the results of various clinical descriptors and noninvasive tests relative to the diagnosis of coronary artery disease, was evaluated in 1,097 consecutive patients without previous myocardial infarction. With this program, each patient was characterized by a probability for coronary artery disease, based on Framingham risk factor analysis, symptom characterization, electrocardiographic stress testing, cardiokymography, cardiac fluoroscopy, thallium perfusion scintigraphy and technetium equilibrium-gated blood pool scintigraphy. A total of 11,808 probability estimates derived from various combinations of the available observations were analyzed: 2,180 in 170 patients undergoing coronary angiography and 9,628 in 969 patients who completed a 1 year follow-up for coronary events.

The predicted probability of disease correlated linearly with observed angiographic prevalence in the 170 patients who subsequently had coronary angiography (prevalence = $[0.001 \pm 0.011] + [0.966 \pm 0.019] \times$ probability). The difference between probability and prevalence averaged 3.1%, and the magnitude of this correlation was not affected by the type or amount of data analyzed. The prevalence of multivessel disease in these patients increased as a monotonic function of disease probability. Below a probability of 25%, single vessel disease was slightly more common than multivessel disease. Above a probability of 75%, multivessel disease predominated. In the 969 patients followed up for 1 year from the date of testing, the incidence of cardiac death and nonfatal infarction increased as a cubic function of disease probability (from approximately 0 to 8% per year for each). Above a probability of 90%, however, the standard deviation for predicting these events was wide.

These data indicate that Bayes' theorem in general—and CADENZA in particular—is an accurate, clinically applicable means for quantifying the prevalence of angiographic coronary artery disease, the risk of multivessel disease and the incidence of morbid coronary events in the year after testing.

CADENZA, an acronym for "computer-assisted diagnosis and evaluation of coronary artery disease," is a microcomputer program that analyzes and reports the results of various important clinical descriptors and noninvasive tests relative to the diagnosis of coronary artery disease (1–4). The program employs a data base of more than 60,000 patients from the medical literature to calculate the probability of coronary artery disease according to Bayes' theorem (1,2). This report describes our experience with CADENZA in a large cohort of patients typical of those referred for noninvasive diagnostic testing and assesses its applicability for predicting the angiographic presence and severity of disease as well as the subsequent incidence of morbid coronary events.

Methods

Study patients. The study population comprises 1,097 consecutive patients (mean age ± standard deviation 56 ± 11 years) without previous myocardial infarction or coronary bypass surgery who were evaluated by noninvasive testing for suspected coronary artery disease in the Cedars-Sinai Medical Center Cardiac Stress Laboratories between January 1, 1979 and November 15, 1980. The majority of these patients were referred for testing because of symptoms or findings that their physicians considered consistent with possible myocardial ischemia. Asymptomatic patients were a heterogeneous group, including those referred for physical fitness

evaluation, arrhythmias, rest electrocardiographic abnormalities or a previously "abnormal" stress test, or any combination thereof. Baseline characteristics of the study population are summarized in Table 1.

Chest pain classification. All patients completed an explicit questionnaire concerning the presence of symptoms consistent with myocardial ischemia at the time of testing and were classified into one of four groups based on three historic criteria relative to location, precipitation and relief of discomfort (2,4,5). Patients were considered to have *typical angina* if they complained of substernal discomfort which was precipitated by physical exertion and relieved within 10 minutes by rest or nitroglycerin. Patients were considered to have *atypical angina* if their discomfort was either not substernal, not precipitated by exertion or not relieved by rest or nitroglycerin. Patients were considered to have *nonanginal discomfort* if more than one of these three defining characteristics were absent. Any patient who denied discomfort above the level of the diaphragm was considered to be *asymptomatic*.

Diagnostic testing. Each patient was evaluated for coronary risk factors in accordance with the definitions employed in the Framingham study (6), and underwent a symptom-limited, maximal treadmill exercise study according to the Bruce protocol. Cardiac medications were withheld for at least 12 (and usually 48 to 72) hours before testing. If more than one exercise test was performed on any patient, only the first was employed for data analysis. Every patient subsequently or simultaneously underwent at least one additional diagnostic test, including cardiac fluoroscopy for detection of coronary artery calcification (7), cardiokymography (8), thallium perfusion scintigraphy (9) and technetium-gated blood pool scintigraphy (10). The details of each of these procedures, as performed in our laboratory, have been published previously (2,9,10). All tests were obtained at the specific request of the referring physician, as part of his clinical evaluation, and the findings were transmitted to him by way of the laboratory's standard clinical reports.

A total of 170 patients (15% of the study population) were subsequently referred for diagnostic coronary angiography during the study period. This decision was made by the referring physician—presumably on the basis of the total clinical evaluation—with the informed consent of the patient. These patients were classified from the official catheterization laboratory clinical angiographic report according to the number of major coronary vessels with at least 50% diameter narrowing.

Follow-up for coronary events. Patients were telephoned by a trained interviewer during the month of the first anniversary of their initial stress test. The follow-up interview consisted of a 5 minute questionnaire designed to determine if the patient had died of cardiac (8 events), or noncardiac (5 events) causes, had had a nonfatal myocardial infarction (7 events) or had undergone coronary artery bypass surgery (47 events) during the year after testing. If such an event was reported, it was then confirmed by obtaining the appropriate official record (hospital chart or death certificate, or both). Over the period of this study, only 3% of patients were lost to follow-up. Those who underwent coronary angiography during the follow-up period were analyzed separately, because the primary physician's knowledge of coronary anatomy might have influenced subsequent outcome.

Probability of coronary artery disease. In addition to the clinical history and risk factor assessment, each patient underwent from 2 to 5 noninvasive diagnostic tests (average 3.3 per patient), providing from 4 to 32 different test combinations per patient. Thus, a total of 9,680 test combinations were available for analysis in the 974 patients with annual follow-up, and 2,180 conbinations in the 170 patients referred for coronary angiography. The probability of having coronary artery disease after testing—conventionally termed the "posterior" coronary artery disease probability (3)—was calculated for each of these combinations by analysis of 20 specific observations summarized in Table 2. Four of these observations were binary, nine were compartmentalized into three to six grades and seven were assessed as continuous variables. The probability determinations employed an Apple II (Apple Computer, Cupertino, California) microcomputer and a software program called CADENZA (Cardiokinetics, Seattle, Washington) (11). This program uses previously published empiric data (1–5) which are analyzed by Bayes' theorem (12). The computational algorithms incorporated in this program are enumerated in the first three sections of the Appendix. A less formal presentation of Bayesian probability analysis is summarized in previous reports from our laboratory (1–3).

Statistical analysis. Statistical analysis of these data is summarized in the Appendix and will be discussed as the results are presented. Statistical significance was assessed by use of Jeffreys' K statistic (13), an estimator of likelihood that was converted into an "exact" posterior probability (Appendix), denoted by the Greek letter ψ. This conversion assumes the null hypothesis in question to be represented by a uniform beta distribution (14). The ψ statistic is properly interpreted as the maximal informational probability that the hypothesis complementary to the null hypothesis is true, given the empiric observations. In practical terms, the ψ statistic expresses the probability that the experimental hypothesis is true. Its value ranges from 0, where the hypothesis is certain to be false, to 1, where the hypothesis is certain to be true.

Table 1. Study Population

	Follow-up (no = 974)		Angiography (no. = 170)	
	no.	%	no.	%
Sex				
Men	683	70	123	72
Women	291	30	47	28
Symptoms				
Asymptomatic	278	29	41	24
Nonanginal pain	294	30	36	21
Atypical angina	226	23	43	25
Typical angina	176	18	50	29
Test Procedures				
Risk factor analysis	974	100	170	100
ECG stress test	974	100	170	100
Cardiac fluoroscopy	649	67	82	48
Cardiokymography	425	44	93	55
Thallium scintigraphy	710	73	115	68
Technetium scintigraphy	413	42	102	60

ECG = electrocardiogram.

Table 2. Variables Analyzed by CADENZA

Variable	Measurement Interval	Conditional Variable
History		
Age (yr)	Continuous	—
Sex	Male, female	—
Chest discomfort	AS, NA, AA, TA	Age, sex
Systolic BP (mm Hg)	Continuous	Sex
Cholesterol (mg/dl)	Continuous	Age, sex
Currently smoking	Yes, no	Sex
Glucose intolerance	Yes, no	Sex
Rest ECG	Normal, abnormal	Sex
ECG Stress Test		
Duration of exercise (min)	Continuous	—
Magnitude of ST depression (mm)	<0.5, 0.5, 1.0, 1.5, 2.0, >2.5	Sex, rest ECG
Slope of ST segment	Upsloping, horizontal, downsloping	Rest ECG
R wave amplitude change (mm)	Continuous	—
Fluoroscopy		
No. of calcified vessels	0, 1, 2, 3	Age
Cardiokymography		
Rest pattern	I, II, III	—
Postexercise pattern	I, II, III	Rest pattern
Thallium Scintigraphy		
Type of defect	None, fixed, reversible	—
Magnitude of defect	Mild, moderate, severe	Type of defect
Pulmonary uptake	None, mild/moderate, moderate/severe	—
Technetium Scintigraphy		
Rest ejection fraction (%)	Continuous	—
Peak exercise ejection fraction (%)	Continuous	—

AA = atypical angina; AS = asymptomatic; BP = blood pressure; ECG = electrocardiogram, I = inward systolic motion, II = mid-systolic outward motion, III = holosystolic outward motion; NA = nonanginal discomfort, TA = typical angina

Results

Probability of coronary artery disease and disease prevalence. We first evaluated the ability of coronary artery disease probability to predict the prevalence of anatomic disease in the 170 patients who were referred subsequently for angiography. The probability of disease based on age and sex did not differ significantly from angiographic prevalence within each symptom class. This finding is similar to that reported previously by us (2,5) and by the Coronary Artery Surgery Study (4), and emphasizes the importance of the clinical history as a potent diagnostic test. Figure 1 illustrates the probability of disease based on age, sex, symptom class and Framingham risk factors. In each symptom class, the probability of coronary artery disease was slightly, but consistently, higher in the 124 patients with disease than in the 46 patients without disease ($\psi = 0.878$), indicating that the risk factors developed within the Framingham study were modest discriminators for coronary artery disease independent of symptom classification (15).

All 170 catheterized patients underwent electrocardiographic stress testing. Additionally, 82 patients had cardiac fluoroscopy, 93 had cardiokymography, 115 had thallium scintigraphy and 102 had technetium scintigraphy. Table 3 summarizes the probability of disease according to the number of diseased vessels found at coronary angiography. These

data were assessed in three ways: 1) based on age, sex, symptom class and risk factors, but prior to diagnostic testing; 2) based on all available data prior to catheterization; and 3) based on every possible combination of the tests performed on each patient. In each case, the probability of

Figure 1. Probability for angiographic coronary artery disease (CAD), based on (thus, "posterior" to) age, sex, symptom class and Framingham risk factors. AA = atypical angina; AS = asymptomatic; NA = nonanginal discomfort; TA = typical angina.
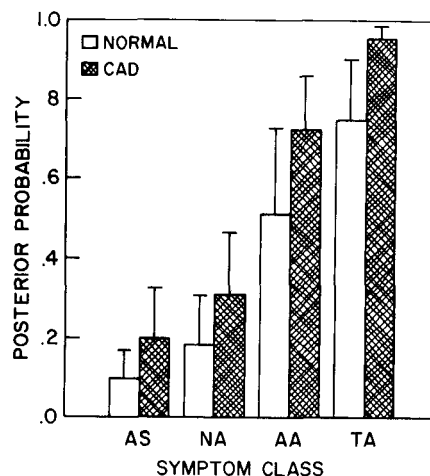
Table 3. Coronary Artery Disease Probability and Angiography

|	No. of Diseased Vessels				
	0	1	2	3	1 + 2 + 3
Patients (no.)	46	21	46	57	124
Estimates Before Testing					
Mean probability	0.291	0 595	0.623	0.660	0.635
Standard deviation	0.259	0.342	0.334	0.327	0.332
Estimates Before Angiography					
Mean probability	0.253	0.745	0.772	0.843	0.800
Standard deviation	0.322	0.387	0.321	0.284	0.315
All Estimates					
Test combinations	500	316	640	724	1680
Mean probability	0.304	0.557	0 730	0.746	0.704
Standard deviation	0.321	0 377	0.323	0.331	0.322

disease tended to increase in proportion to the number of diseased vessels, but the standard deviations were large.

Figure 2 (left panel) illustrates the frequency distributions of all 2,180 raw probability estimates according to the number of diseased vessels. The mean probability for disease increased from 30% for the 500 estimates in the normal group to 56% for the 316 estimates in patients with single vessel disease ($\psi = 0.999$), and to 75% for the 724 estimates in patients with triple vessel disease ($\psi = 0.999$). There was a great deal of overlap in the distribution of these data,

however, especially between two vessel and three vessel disease, which were not significantly different from each other ($\psi = 0.018$). Thus, 8% (40 of 500) of the probability estimates in the 46 normal patients were in excess of 90%, while 9.7% (163 of 1,680) of the probability estimates in the 124 patients with angiographic disease were under 10%.

These raw distributions are highly asymmetric. For statistical analysis, therefore, these data are appropriately modeled by being fit to the beta probability distribution (Appendix, Beta Frequency Distribution). Figure 2 (right

Figure 2. Posterior probability distribution. **Left panel**: Raw distribution of posterior probability based on age, sex, symptom class, Framingham risk factors and all combinations of noninvasive test results (2,180 estimates in 170 patients). There is a trend toward a greater probability as the number of diseased vessels increases. **Right panel**: Beta distribution of posterior probability derived from the means and standard deviations of the data in the **left panel**. The fitted distributions correlate closely with the raw data (average difference = 3.5%). Analogous distributions based on specific combinations of test results were similar.

NORMAL

SINGLE VESSEL

DOUBLE VESSEL

TRIPLE VESSEL

NUMBER OF ESTIMATES

POSTERIOR PROBABILITY
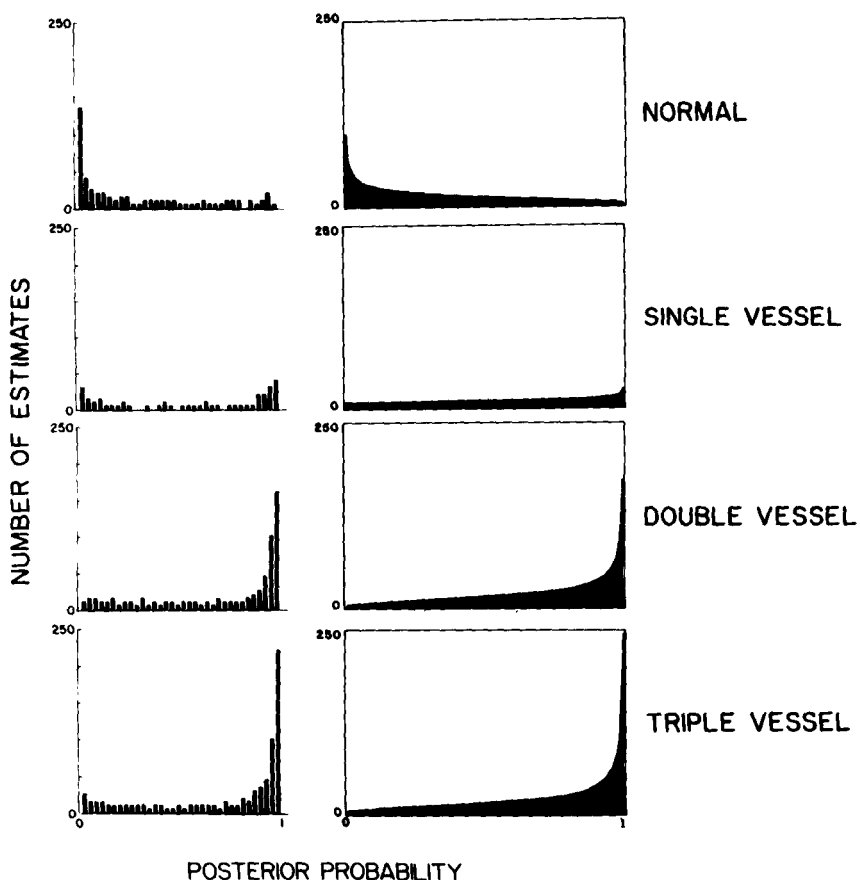
panel) illustrates the beta distributions for the data in the left panel of the figure. The difference between the raw distributions and the corresponding beta distributions averaged only 3.5%. On this basis, the beta distributions were employed in all subsequent data analysis.

Figure 3 illustrates the relation between probability of coronary artery disease and the observed prevalence of disease at angiography, based on the beta distributions defined by the data in Table 3 and illustrated in Figure 2. The average difference between the observed prevalence of disease and that predicted by probability of coronary artery disease was 3.4% for estimates based on age, sex, symptoms and risk factors ($\psi = 0.039$). For estimates based on all available tests before angiography, the difference between observed and predicted prevalence averaged 3.0% ($\psi = 0.038$), and for all combinations of test results the difference was 3.1% ($\psi = 0.013$). For each set of data, the slope of the correlation between probability and prevalence was near unity and the y intercept was near zero by chi-square minimization (Fig. 3). These correlations indicate that calculated probability of coronary artery disease accurately reflected the actual angiographic disease prevalence, regardless of the *amount* of data analyzed.

To determine if the magnitude of the correlation between probability and prevalence varied as a function of the specific tests employed, different combinations of tests were analyzed separately. Figure 4 illustrates the correlation between probability of coronary artery disease and angiographic prevalence of disease for eight different combinations, including the electrocardiographic stress test alone and the four two-test combinations that employed this test. In each case, the previously noted linear relation between probability and prevalence was maintained, indicating that the accuracy of a given probability estimate was generally independent, not only of the amount, but also of the *type* of data analyzed.

Probability of coronary artery disease and extent of disease. Figure 5 illustrates the relation between coronary artery disease posterior probability and the relative extent of angiographic disease. These curves were calculated from the beta distributions for single, double and triple vessel disease illustrated in Figure 2, and are normalized for total disease prevalence. Thus, at any posterior probability, the sum of the three relative prevalences is fixed at 100%. As in Figure 2, the configuration of the graphs representing double vessel disease and triple vessel disease was nearly identical, supporting the practical relevance of the widely employed term "multivessel disease." Accordingly, Figure 6 expresses the prevalence of single vessel disease and multivessel disease (the sum of double and triple vessel disease) as a function of posterior probability. These curves illustrate that below a probability of 25% (p_1), when disease was present, single vessel disease was slightly more prevalent than multivessel disease, while above a probability of 75%
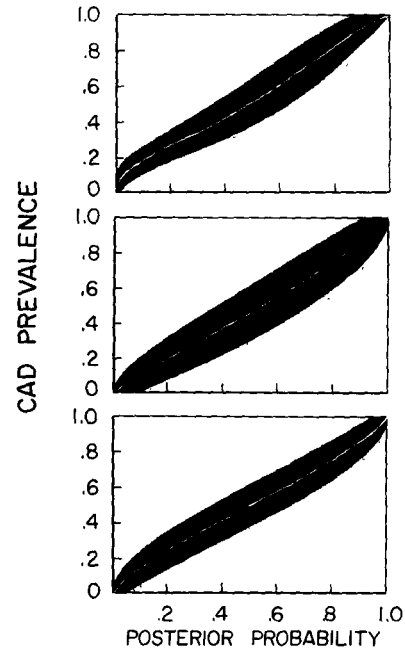

Figure 3. Angiographic prevalence of coronary artery disease (CAD) (y axis) as a function of posterior probability (x axis). The **top panel** employs 170 probability estimates derived from consideration of age, sex, symptom class and Framingham risk factors. The **middle panel** is for 170 estimates that employ these data plus the results of all tests performed. The **bottom panel** is for all 2,180 combinations of data. The **white line** represents the mean of each correlation and the **black band** is 1 standard deviation on either side of the mean. In each case, the correlation between prevalence and probability was closely approximated by a linear function approaching identity:

$$\text{Top: } y = 0.074 \pm 0.010 + (0.932 \pm 0.011)\, x,$$
$$\text{Middle: } y = 0.001 \pm 0.011 + (0.966 \pm 0.019)\, x,$$
$$\text{Bottom: } y = 0.039 \pm 0.011 + (0.917 \pm 0.018)\, x.$$

When the mean probability and its standard deviation were replaced by the "exact" posterior probability (see Appendix, The Exact Posterior Probability), the standard error of the estimate for its correlation with angiographic prevalence decreased by an average of 59%, justifying its use as a summary statistic.

Figure 4. Correlation between posterior probability and angiographic prevalence of coronary artery disease (CAD) for various combinations of tests. The two most extreme combinations are labeled. ECG = electrocardiographic stress test; TC = technetium blood pool scintigraphy; TL = thallium perfusion scintigraphy. The standard deviations for the correlations all overlap each other and, therefore, are not illustrated.
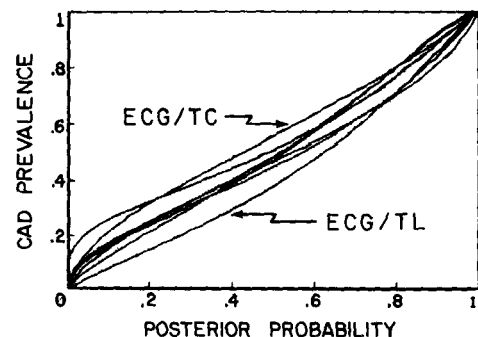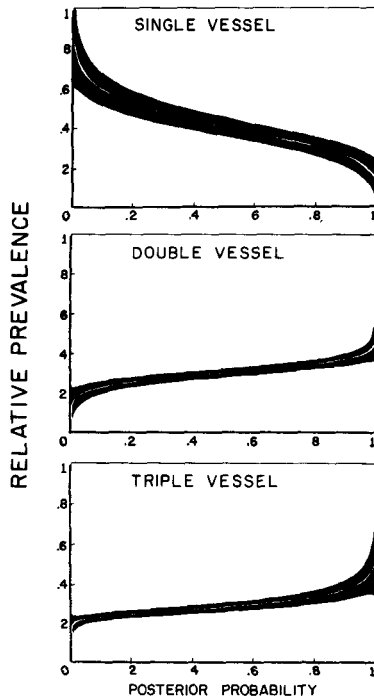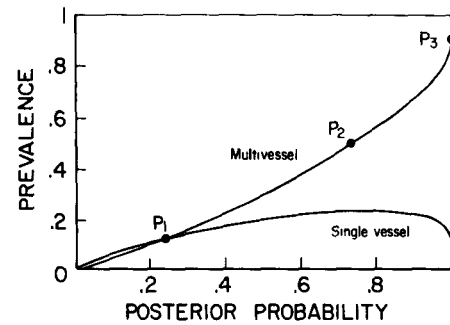
Figure 5. Posterior probability and relative prevalence of single, double and triple vessel disease. The **black bands** are the standard deviation of the mean correlations. Relative prevalence of single vessel disease decreases, while those for double and triple vessel disease increase in parallel as a function of probability.

Figure 6. Prevalence of single vessel and multivessel disease as a function of posterior probability. At a probability of 25% (p_1), it is most likely that disease is not present. In those with disease, however, the prevalence of single vessel disease just equals that for multivessel disease. At a probability of 75% (p_2), the prevalence of multivessel disease just equals that for nonmultivessel disease. At a probability of 100% (p_3), multivessel disease accounts for 89% of all disease.

(p_2), multivessel disease predominated. At a probability of 100% (p_3), multivessel disease accounted for 89% of all angiographic disease. The significance level for these observations varied as a function of posterior probability. Nevertheless, these data indicate that disease probability also acted as a quantitative measure of anatomic severity.

Probability of coronary artery disease and future coronary events. Table 4 summarizes the results of probability analysis in the 969 patients followed up for 1 year after the date of testing, excluding 5 patients who had a documented noncardiac death. In 47 patients, follow-up was interrupted by referral for coronary artery bypass surgery. There were 15 (1.6%) morbid events (7 nonfatal infarctions and 8 cardiac deaths) in the 922 patients who did not undergo coronary angiography or bypass surgery during the follow-up period. A total of 9,628 estimates of probability were analyzed: 8,900 in the 907 patients without morbid events, 592 in the 47 surgical patients and 136 in the 15 patients with morbid events. Posterior probability was significantly higher in each of the three event groups compared with the nonevent group (ψ = 0.999 for all). Figure 7 illustrates the annual coronary event rate as a function of posterior probability. The event rates for myocardial infarction and for cardiac death were similar in magnitude (ψ = 0.103), and each increased as a cubic function of probability.

The curves in Figure 7 were derived from only those patients who were not referred to surgery. This analysis,

then, should be expected to underestimate the rate of morbid events in a diseased stress test population for two reasons. First, it is likely that not all of the patients without events had coronary disease because their probability of coronary artery disease averaged only 49%; second, the surgically treated patients, presumably with a higher risk, were excluded from analysis. In order to better estimate *the event rate for a diseased cohort before surgical referral*, the data were normalized by assuming that the event rate for patients with coronary artery disease referred to surgery is the same as for those not so referred at a given level of posterior probability (Table 4). The total event rate was thereby predicted to be 3.1% (1.7% for infarction and 1.4% for cardiac death). The relation between posterior probability and normalized annual event rate is illustrated in Figure 8. These graphs allow the posterior probability to serve as an esti-

Table 4. One Year Follow-up for Coronary Events

Class	No. of Patients	No. of Estimates	CAD Probability	Standard Deviation
Observed (patients)				
No events	907		0.486	0.403
Bypass surgery	47		0.898	0.251
Myocardial infarction	7		0.874	0.308
Cardiac death	8		0.795	0.333
Observed (estimates)				
No events		8900	0.527	0.381
Bypass surgery		592	0.858	0.252
Myocardial infarction		72	0 816	0.282
Cardiac death		64	0.746	0.301
Predicted (estimates)				
No events		5250*	0.547	0.375
Myocardial infarction		92¶	0.825	0.276
Cardiac death		76†	0.763	0.294

*Includes 4690 estimates predicted from posterior probability to have disease but no event, and 560 surgical estimates predicted from Figure 7 not to have an event. (8900 × 0.527) + (592 − 20 − 12) = 5250 ¶Includes 20 surgical estimates predicted from Figure 7 to have infarction †Includes 12 surgical estimates predicted from Figure 7 to have a cardiac death

CAD = coronary artery disease

mator of risk for a morbid coronary event in patients with coronary artery disease over the year following testing.

Discussion

This study documents the clinical applicability of a microcomputer program that employs empiric data and Bayesian analysis for the noninvasive diagnosis of coronary artery disease in three broad areas: for *diagnosis*, by the correlation of posterior probability with angiographic prevalence; for *evaluation*, by the relation of posterior probability to the anatomic severity of disease and for *prognosis*, by the relation of posterior probability to the annual incidence of morbid coronary events.

Prevalence of disease. The probability for coronary artery disease correlated almost linearly with the observed angiographic prevalence of disease, and the magnitude of this correlation was not dependent on the type and amount of data analyzed. Thus, a given value of probability based only on the patient's age, sex, symptom classification and Framingham risk factors was just as accurate as one based also on various combinations of noninvasive test procedures.

The first impression given by such a result is that additional testing is of no value. On reflection, however, it becomes apparent that posterior probability should always correlate well with actual disease prevalence whether 1 or even 10 tests are employed, provided that the sensitivity and specificity of each test are accurately known and that the observations are independent of one another. For instance, an individual patient's probability for having coronary artery disease may change from 10 to 50% on the basis of test results, but the actual prevalence of disease in a group of patients with a 10% probability should be 10% and, in a group of patients with a 50% probability, the actual prevalence should be 50%. Additional testing, then, moves the individual patient to a different population subset with a different disease prevalence and a different probability which represents that prevalence.

Previous reports on CADENZA. This correlation of posterior probability with angiographic prevalence has been reported in various ways both by us (2) and by others. Chaitman et al. (5) demonstrated the relation of a symptom classification very similar to ours with angiographic prevalence of disease in 8,192 patients evaluated as part of the Coronary Artery Surgery Study (CASS); probability estimates obtained from CADENZA correlated accurately with their empiric observations (4).

Greenberg et al. (16) evaluated 113 patients similar to ours by history and electrocardiographic stress testing. They analyzed these data both by multivariate analysis (17) and by CADENZA. The difference between probability and prevalence in these patients averaged 4.1% for multivariate analysis and 8.1% for CADENZA. The patient by patient correlation between the two methods was linear and highly significant. These findings are all the more impressive because the design of this study favored the multivariate format, in that the discriminant coefficients were derived internally (personal communication).

Dans (personal communication) employed CADENZA in a retrospective analysis of patients referred for angiography at the Johns Hopkins Hospital. Of 43 catheterized patients with a disease probability of less than 10% (average 2.8%), only 1 (2.3%) had angiographic coronary artery disease. Patients with a high probability were not analyzed in this study.

Wong et al. (18) prospectively evaluated 253 patients similar to those in our study, using history, stress testing, nuclear scintigraphy and CADENZA. They found that posterior probability was predictive of angiographic prevalence in the 68 catheterized patients. In addition, they reported that the probability estimates correlated significantly with the judgment of experienced cardiologists.

Hlatky et al. (19) also confirmed the accuracy of CADENZA in predicting angiographic disease. They reported an accuracy of 90% based on age, sex, symptoms, electrocardiographic stress testing and thallium scintigraphy in 51 patients. In their study, the computer predictions were slightly, but significantly, more accurate than the clinical judgment of 91 experienced cardiologists (19,20).

Possible limitations in clinical application. A number of other investigators have verified the underlying applicability of Bayes' theorem for diagnosis of coronary artery disease (21–28). Yet, despite this impressive body of data,

Figure 7. Relation between coronary events at 1 year and posterior probability after excluding patients referred for coronary artery bypass surgery. The incidence of cardiac death and that of nonfatal myocardial infarction both increase similarly as a cubic function of probability.
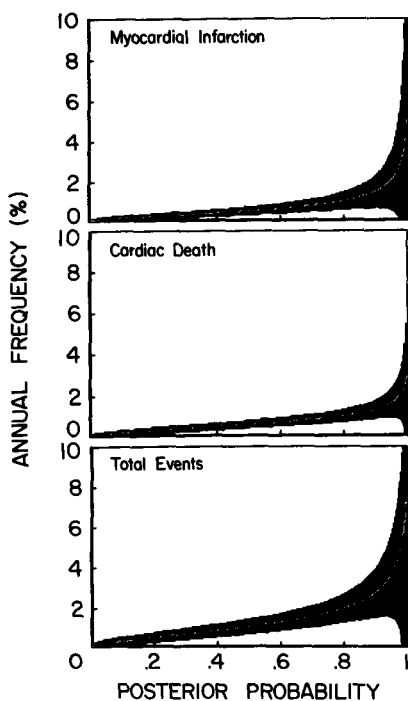
Table 5. Regression Functions for Continuous Variables

Characteristic	Probability (p)	Observation (j)	Regression Function F(j)
Symptom class	p(D+)	Age (yr)	$1/1 + e^{-(aj^2 + bj + c)}$
R wave amplitude	$p(T_j\|D_i)$	mm	$1/ae^{(i-b)j^2/c}$
Coronary calcification	$p(T_j\|D_i)$	Age (yr)	$1/1 + e^{-(aj^2 + bj + c)}$
Ejection fraction	$p(T_j\|D_i)$	Percent	$aj^b(1-j)^c$
Exercise time	$p(T_j\|D_i)$	Minutes	$1/ae^{(i-b)j^2/c}$

The symbols a, b and c are constants and e is the base to the natural logarithms Other symbols are defined in the Appendix.

probability analysis is sparsely employed as a formal clinical tool. There are at least two reasons for its limited acceptance. First, its recondite nature conflicts with the requirement of practicality. The ready availability of microcomputers and their growing acceptance by clinicians are, however, slowly overcoming this objection (24,29–32).

The second impediment to the acceptance of probability analysis relates to a fundamental theoretical objection. Unlike conventional clinical judgment, the use of Bayes' theorem rigorously and explicitly assumes that the individual observations analyzed are all stochastically independent. Although it is very unlikely that this condition is fulfilled often, there are at least three indirect lines of evidence to support the appropriateness of this assumption on a practical level. First, Bayes' theorem has been shown to be relatively robust in terms of dependence unless the number of such variables is large (33). In our study, anywhere from 7 to 20 variables per patient were analyzed (average = 14), but only a small fraction of these would be expected to be highly dependent. Charuzi et al. (34), for instance, reported no significant dependence among several of the major variables analyzed in this study. Second, as noted earlier, methods that do not assume independence (multivariate analysis, for example) result in probabilities similar to those determined from Bayes' theorem (16). Third, the linear correlation between probability and prevalence reported in this study is of itself support for the relative independence of the variables analyzed. If, for example, the degree of dependence were major, any correlation that incorporated all available data should have been significantly less accurate than one that employed the fewest variables. As shown in Figure 3, however, the accuracy of these correlations was almost identical. In fact, the observed probability and standard deviation for the normal groups and patients with disease (Table 3) were very similar to those predicted on the ideal assumption of *total* independence (the predicted mean for normal subjects was 33%; the predicted mean for patients with disease was 67% and the standard deviation for each was 27%) (35).

We believe, therefore, that the broad clinical experience reported here and from other institutions adequately supports the accuracy and practical utility of empiric Bayesian analysis. These observations further imply that the published

estimates, on which the method is based, "travel" well from laboratory to laboratory (36) despite local differences in methodology. Thus, posterior probability might be an appropriate characteristic for antecedently defining subsets of patients in multicenter trials and epidemiologic studies where angiographic diagnosis is not feasible (Rozanski A, Diamond GA, Berman DS, et al., unpublished observations).

Severity of disease. Posterior probability also served as a predictor of the extent of angiographic disease, because the prevalence of multivessel disease increased as a monotonic curvilinear function of probability to a high of 90% when disease probability was 100%. This result is not unexpected. It is well known that the sensitivity of noninvasive tests is highest in patients with more severe disease. Thus, both multivessel disease and higher grade stenosis have been

Figure 8. Predicted relation between coronary events at 1 year and posterior probability before surgical referral. The curves are shifted upward in comparison with Figure 7, reflecting an increase in the predicted event rate (see text and Table 4).
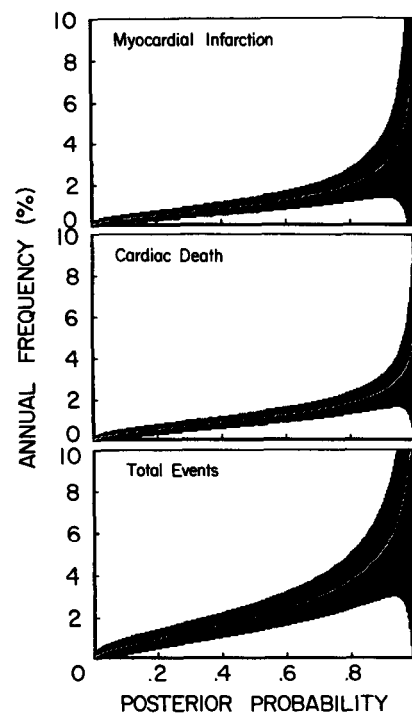
shown to be more frequently associated with abnormal test responses than are lesser degrees of disease severity (22,34). Because posterior probability also increases in proportion to the number and magnitude of various test abnormalities, it thereby is an index of the anatomic extent of disease. This inference could serve as an aid in the appropriate selection of patients for coronary angiography.

Future coronary events. Because the angiographic severity (that is, the magnitude and extent) of disease is an important determinant of prognosis (37), it is also not surprising that posterior probability correlated with the incidence of both nonfatal myocardial infarction and cardiac death in the year after testing. By using the empiric relation between posterior probability and coronary events to predict the number of events that were "lost" to follow-up in the surgical patients and then adding these "events" to those observed, the normalized morbid event rate for a medically treated diseased population was predicted to be $3.1 \pm 1.0\%$ in the year after testing. This value is consistent with recent reports of similar groups of medically treated patients without previous myocardial infarction (37–42). The almost equal incidence of nonfatal infarction and cardiac death in our study is in agreement with the observations of Harris et al. (37), who followed up a similarly sized cohort over a longer period of time.

When posterior probability was high (for example, >90%), the probability of an event in the year after testing increased substantially, but the standard deviation for this prediction was large. This observation has an important implication that is easily overlooked by clinicians. The data imply that when noninvasive testing results in a low probability of disease, we should be justifiably confident that such patients have a good short-term prognosis. When posterior probability is greater than 90%, however, we should be very unsure of a given patient's outlook. In this context, a statement to the effect that a patient is at "high risk" can be made with only a limited degree of confidence, and carries an emotional impact out of proportion to our knowledge of its accuracy and precision (5,43).

Clinical relevance of the study design. Several aspects of our study design deserve emphasis. First, we sought to determine the accuracy of probability analysis in an environment that mirrored, as closely as possible, the realities of clinical practice. Patients were excluded from entry into this study, therefore, only if a diagnosis of coronary artery disease was already *absolutely* established or excluded by a documented prior myocardial infarction or coronary angiography. No attempt was made to influence the referring physician in his choice of diagnostic tests or in his decision to recommend coronary angiography or bypass surgery, beyond the customary communication that took place as part of routine clinical care. Consequently, all tests were not performed on all patients and the resultant data base was

unavoidably incomplete. Also, every test result we analyzed was that which was formally reported to the referring physician. No separate "research readings" were performed. Under these circumstances, the probability for disease averaged 52% for all 1,097 patients referred for stress testing; probability increased to 65% for the 170 patients referred for angiography and to 90% for the 47 patients referred for surgery. Whether or not this pattern of referral is representative of other laboratories (or is even appropriate) was not addressed by our analysis. Lastly, the interpretation of each test—from the history to the coronary angiogram—required a variable amount of subjectivity. This, however, also mirrors clinical reality, and is in one sense desirable, because the empiric data on which CADENZA is based were similarly obtained (43).

In summary, a practical, readily available microcomputer program that integrates a large base of published data has been employed successfully in the diagnosis and evaluation of coronary artery disease. Its continued development and its current role as an aid to clinical management (44) appear warranted.

APPENDIX

Bayes' Theorem

If we have a complete system of n mutually exclusive states, D_1, D_2, . . . D_n, then the probability of the ith state is denoted $p(D_i)$, where $\sum_{i=1}^{n} p(D_i) = 1$. If any j of k independent observations, T_j, is associated with each of these n states, then the conditional probability of each of the k associations, denoted $p(T_j \mid D_i)$, represents the probability that the observation T_j will occur given the state D_i. The inverse probability, the conditional probability of D_i given T_{jk}, is defined by Bayes' theorem:

$$p(D_i \mid T_{jk}) = \frac{p(D_i) \prod_{jk} p(T_{jk} \mid D_i)}{\sum_i p(D_i) \prod_{jk} p(T_{jk} \mid D_i)}.$$

For the diagnosis of coronary artery disease, the set of D_i includes only $D+$ (disease) and $D-$ (no disease). In this case, $p(T_{jk} \mid D+)$ is the sensitivity or true positive rate, and $p(T_{jk} \mid D-)$ is the false positive rate $(1 - \text{specificity})$ of the kth observation, T_j. The inverse probability calculated by Bayes' Theorem, given the prior probability for disease, $p(D+)$, and the set of k observations, T_j, is $p(D+ \mid T_{jk})$, the posterior probability for coronary artery disease.

Derivation of Input Probabilities

The discrete probabilities $p(D+)$, $p(T_{jk} \mid D+)$ and $p(T_{jk} \mid D-)$ were obtained directly from reported empiric observations (1–4):

$$p = \frac{\text{Number of individuals in a population, N, with characteristic j}}{\text{Total number of individuals in the population, N}}.$$

In certain cases, these empiric probabilities were converted into continuous variables (3) by regression analysis of a test characteristic relative to a specific observation, j, such that $p = F(j)$. Table 5 summarizes these conversions.

Calculation of Variances

The variances of empiric probabilities were determined from the equation:

$$\sigma_p^2 = pq/N = n(N - n)/N^3,$$

where N = total population, n = number of individuals in N with the characteristic, j, $p = n/N$, and $q = 1 - p = (N - n)/N$. The variances of continuous probabilites were calculated from the first order terms of a Taylor series expansion of the function $F(j) = p$ such that:

$$\sigma_p^2 = \left(\frac{dp}{dj}\right)^2 \cdot \sigma_j^2.$$

When those independent and uncorrelated errors are introduced into Bayes' equation for calculation of $p(D+ \mid T_j)$, the resultant posterior variance (σ^2) is given by a partial differential equation which sums the individual errors in quadrature:

$$\sigma^2 = \left(\frac{\partial L}{\partial A}\right)^2 \cdot \sigma_A^2 + \left(\frac{\partial L}{\partial B}\right)^2 \cdot \sigma_B^2 + \left(\frac{\partial L}{\partial P}\right)^2 \cdot \sigma_P^2,$$

where $L = p(D+ \mid T_j)$, $A = p(T_j \mid D+)$, $B = p(T_j \mid D-)$, $P = p(D+)$. We evaluate the partial derivatives by substitution into Bayes' equation:

$$\frac{\partial L}{\partial A} = \frac{\prod_i p(D_i \mid T_j)}{p(T_j \mid D+)},$$

$$\frac{\partial L}{\partial B} = \frac{\prod_i p(D_i \mid T_j)}{p(T_j \mid D-)},$$

$$\frac{\partial L}{\partial P} = \frac{\prod_i p(D_i \mid T_i)}{\prod_i p(D_i)},$$

$$\sigma^2 = \prod_i^2 p(D_i \mid T_j) \cdot \left[\left(\frac{\sigma_A}{p(T_j \mid D+)}\right)^2 \right.$$
$$\left. + \left(\frac{\sigma_B}{p(T_j \mid D-)}\right)^2 + \left(\frac{\sigma_P}{\prod_i p(D_i)}\right)^2\right]$$

Beta Frequency Distribution

The beta function is defined by:

$$B(n,r) = \frac{\Gamma(n)}{\Gamma(r)\,\Gamma(n - r)} \cdot p^{(r-1)} \cdot q^{(n-r-1)}.$$

This function describes the frequency distribution of a continuous variable, such as posterior probability, which varies from 0 to 1 in terms of two parameters, n and r:

$$n = \prod_i p(D_i \mid T_j)/\sigma^2,$$

$$r = p(D+ \mid T_j) \cdot n.$$

The symbol Γ represents the gamma function:

$$\Gamma(y) = \int_0^\infty e^{-x} x^{y-1}\, dx = \int_0^1 \left(\ln \frac{1}{x}\right)^{y-1} dx.$$

The uniform beta distribution defines the classic uninformative Bayesian prior (14). Its parameters n and r may be derived by recognizing that the restriction of uniformity requires that for all values of p, $B(n,r) = 1$, and its first derivative, $B'(n,r) = 0$:

$$B'(n,r) = \left[\frac{r - 1}{p} - \frac{n - r - 1}{q}\right] B(n,r) = 0,$$

$$\frac{r - 1}{p} - \frac{n - r - 1}{q} = 0.$$

When $p = 0$, then $q = 1$ and $r - 1 = 0$. When $p = 1$, then $q = 0$ and $-n + r + 1 = 0$. Thus, $r = 1$ and $n = 2$.

The uniform prior distribution is therefore defined by the specific beta distribution, $B(2,1)$, with mean $\bar{p} = r/n = 1/2$ and variance $\sigma^2 = \bar{p}\bar{q}/n = 1/8$. The parameters of a posterior beta distribution which derive from the uniform prior are $n = 1/2\bar{p}\bar{q}$ and $r = 1/2\bar{q}$. These parameters were employed in calculating the test statistic, ψ (see Appendix, The Exact Posterior Probability). If one has two sets of N_i posterior probabilities, $p(D+ \mid T_j)$ and their individual associated variances, σ_p^2, for each of two exhaustive states, D_i, we can define the mean posterior probability of each set as:

$$\mu_i = \sum_i p(D+ \mid T_j, D_i)/N_i.$$

The variance of this mean is denoted $\sigma_{\mu_i}^2$, and the average variance of the individual probabilities is:

$$\bar{\sigma}_{p_i}^2 = \sum_{N_i} \sigma_{p_{N_i}}^2 /N_i.$$

We can then express each set as a beta distribution, $B(n_i, r_i)$, and define a new function that represents the normalized probability (or prevalence) of the state $D+$, given a specific observed posterior probability:

$$p\left(D+ \mid p(D+ \mid T_j)\right) = \frac{B(n_1, r_1)}{B(n_1, r_1) + B(n_2, r_2)} = F(P).$$

This function has a variance defined by:

$$\sigma^2_{F(P)} = [F'(P)]^2 \cdot \bar{\sigma}_P^2,$$

where $F'(P)$ is the first derivative of $F(P)$:

$$F'(P) = \left[\frac{r_2 - r_1}{P} - \frac{n_1 - r_1 - n_2 + r_2}{1 - P}\right] \cdot F(P) \cdot [1 - F(P)]$$

and $\bar{\sigma}_P^2$ is the total variance of P:

$$\bar{\sigma}_P^2 = \frac{P(1 - P)}{2} \sum_i \bar{\sigma}_{P_i}^2 \Big/ [\mu_i(1 - \mu_i)].$$

The Exact Posterior Probability

When disease probability is expressed as a density function such as B(n,r), its information content (45) is defined by the entropy of each possible value of posterior probability weighted for its beta distributed frequency of occurrence (46). We thereby define the entropy distribution S(n,r) as:

$$S(n,r) = -B(n,r) \sum_i p(D_i \mid T_j)\log_2 p(D_i \mid T_j).$$

This function evaluates to zero at the extremes of posterior probability, and increases to a single maximum at some $0 < \psi < 1$ which satisfies the condition that the derivative of the entropy function is zero:

$$S'(n,r) = 0 = \left[(n - r) - \frac{1 - \psi}{\psi}(r - 1)\right] \ln(1 - \psi)$$
$$- \left[r - \frac{\psi}{1 - \psi}(n - r - 1)\right] \ln \psi.$$

This "exact" posterior probability is uniquely defined by the mean and standard deviation of the posterior probability distribution which define the parameters n and r. It is almost identical to the median of the distribution and represents the *maximal informational* probability for a given patient (43). In physics, entropy is the sole determinant of thermodynamic equilibrium at constant energy (47). It is convenient, therefore, to view the exact posterior probability, ψ, as the informational "equilibrium" of the posterior probability distribution. The exact posterior probability, therefore, may be employed as a meaningful summary statistic of the entire posterior probability distribution. When ψ is calculated from a likelihood estimator such as K (13), it represents the posterior probability of the hypothesis complementary to the null hypothesis (see Methods). This statistic, therefore, is more relevant to hypothesis testing than the conventional p value which is equivalent to a false positive rate (48–50).

References

1. Diamond GA, Forrester JS. Analysis of probability as an aid in the clinical diagnosis of coronary artery disease. N Engl J Med 1979;300:1350–8.

2. Diamond GA, Forrester JS, Hirsch M, et al. Application of conditional probability analysis to the clinical diagnosis of coronary artery disease. J Clin Invest 1980;65A:1210–21.

3. Diamond GA, Forrester JS. Improved interpretation of a continuous variable in diagnostic testing: probabilistic analysis of scintigraphic rest and exercise left ventricular ejection fractions for coronary disease detection. Am Heart J 1981;102:189–95.

4 Diamond GA, Forrester JS. Probability of CAD Circulation 1982;65:641–2.

5. Chaitman BR, Bourassa MG, Davis K, et al. Angiographic prevalence of high-risk coronary artery disease in patient subsets (CASS). Circulation 1981;64:360–7.

6. Kannel WB, McGee D, Gordon T. A general cardiovascular risk profile: the Framingham study. Am J Cardiol 1976;38:46–52.

7   Aldrich RF, Brensike JF, Battaglini JW, et al Coronary calcifications in the detection of coronary artery disease. Circulation 1979;59:1113–24.

8. Silverberg RA, Diamond GA, Vas R, Tzivoni D, Swan HJC, Forrester JS. Noninvasive diagnosis of coronary artery disease: the cardiokymographic stress test. Circulation 1980;61:579–89.

9. Rozanski A, Berman DS, Gray R, et al. Use of thallium-201 redistribution scintigraphy in the preoperative differentiation of reversible and nonreversible myocardial asynergy. Circulation 1981;64:936–44.

10 Rozanski A, Berman D, Gray R, et al. Preoperative prediction of reversible myocardial asynergy by postexercise radionuclide ventriculography. N Engl J Med 1982;307:212–6.

11. Diamond GA. (in collaboration with Forrester JS) CADENZA: computer-assisted diagnosis and evaluation of coronary artery disease. 1979, Cardiokinetics, Seattle, Washington (software and documentation magnetic tape).

12. Ledley RS, Lusted LB. Reasoning foundations of medical diagnosis. Science 1959;130:9–21.

13. Jeffreys H. Theory of Probability. 3rd ed. London: Oxford University Press, 1961:245–331, 432–41.

14. Novic MR, Jackson RH. Statistical Methods for Educational and Psychological Research. New York: McGraw-Hill, 1974:1–150

15. Salel AF, Fong A, Zelis R, et al. Accuracy of coronary profile. correlation of risk factors with arteriographically documented severity of atherosclerosis. N Engl J Med 1977;296:1447–50.

16. Greenberg P, Diamond G, Ellestad MH A comparison of the multivariate analysis and CADENZA systems for determination of the probability of coronary artery disease. In Exercise Stress Testing: Today and Tomorrow. Presented by the American College of Cardiology and the Division of Cardiology, Memorial Hospital Medical Center of Long Beach, 1982:34–5.

17. Ellestad MH, Savitz S, Bergdall D, Teske J. The false positive stress test: multivariate analysis of 215 subjects with hemodynamic angiographic and clinical data. Am J Cardiol 1977;40:681–5.

18. Wong DF, Tibits P, O'Donnell J, et al. Computer-assisted Bayesian analysis in the diagnosis of coronary artery disease (abstr). J Nucl Med 1982;23:P83.

19. Hlatky M, Botvinick E, Brundage B. Cardiologists versus Bayes rule (abstr). Circulation 1981;64(suppl IV):IV-78

20. Hlatky M, Botvinick E. Brundage B. Diagnostic accuracy of cardiologists compared with probability using Bayes' rule. Am J Cardiol 1982;49:1927–31.

21. Melon JA, Piret LJ, Vanbutsele RJM, et al. Diagnostic value of exercise electrocardiography and thallium myocardial scintigraphy in patients without previous myocardial infarction. a Bayesian approach. Circulation 1981;63.1019–24

22 Weiner DA, Ryan TJ, McCabe CH, et al. Correlations between history and exercise ECG with coronary arteriography N Engl J Med 1979;301:230–5.

23. Rifkin RD, Hood WB. Bayesian analysis of electrocardiographic exercise testing. N Engl J Med 1977;297:681–6.

24. Santinga JT, Flora J, Maple R, Brymer JF, Pitt B The determination of the post-test likelihood for coronary disease using Bayes theorem. J Electrocardiol 1982;15:61–8.

25. Christopher TD, Konstantinow G, Jones RH. Bayesian analysis applied to radionuclide angiographic diagnosis of coronary artery disease (abstr). Circulation 1981;64(suppl IV):IV-262.

26. Epstein SE. Implications of probability analysis on the strategy used for noninvasive detection of coronary artery disease. Role of single or combined uses of exercise electrocardiographic testing, radionuclide cineangiography and myocardial perfusion imaging. Am J Cardiol 1980;46:491–9.

27. Wagner HN. Bayes' theorem: an idea whose time has come? Am J Cardiol 1982;49:875–8.

28. Patterson RE, Horowitz SF, Eng C, et al. Can exercise electrocar-

diography and thallium-201 myocardial imaging exclude the diagnosis of coronary artery disease? Bayesian analysis of the clinical limits of exclusion and indications for coronary angiography Am J Cardiol 1982;49:1127–35.

29. Wagner HN Jr. Principles of Nuclear Medicine. The Diagnostic Process. Philadelphia: WB Saunders, 1968:15–22.

30. Diamond GA. Computer diagnosis: revolution or revelation? Internat J Cardiol 1982;2:219–20.

31. Blois MS. Clinical judgment and computers N Engl J Med 1980;303:192–7.

32. Goldman L, Weinberg M, Weisberg M, et al. A computer-derived protocol to aid in the diagnosis of emergency room patients with acute chest pain N Eng J Med 1982;307.588–96.

33. Fryback DG. Bayes' theorem and conditional nonindependence of data in medical diagnosis. Comp Biomed Res 1978;11:423–34.

34. Charuzi Y, Diamond GA, Pichler M, et al. Analysis of multiple noninvasive test procedures for the diagnosis of coronary artery disease. Clin Cardiol 1981;4:64–74.

35. Diamond GA. An improbable criterion of normality (letter). Circulation 1982;66:681.

36. de Dombal FT, Staniland JR, Clamp SE. Geographical variation in disease presentation Does it constitute a problem and can information science help? Med Decision Making 1981;1:59–69.

37. Harris PJ, Lee KL, Harrell FE, Behar VS, Rosati RA. Outcome in medical treated coronary artery disease. Ischemic events: non-fatal infarction and death. Circulation 1980;62.718–25.

38. Harris PJ, Harrell FE, Lee KL, Behar VS, Rosati RA. Survival in medically treated coronary artery disease. Circulation 1979;60:1259–69.

39. Vliestra RE, Assad-Morell JL, Frye RL, et al. Survival predictors in coronary artery disease. Medical and surgical comparisons. Mayo Clin Proc 1977;52:85–90.

40. 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 1977;297:621–7.

41. Gross H, Vaid AK, Cohen MV. Prognosis in patients rejected for coronary revascularization surgery. Am J Med 1978;64:9–19.

42. European Coronary Surgery Study Group: coronary-artery bypass surgery in stable angina pectoris. Survival at two years. Lancet 1979.1:889–93

43. Diamond GA, Forrester JS. Metadiagnosis: an epistemological model of clinical judgment. Am J Med (in press).

44. Wagner HN Jr. The future of nuclear cardiology. Am J Cardiol 1982;49:1355–61.

45. Shannon EC, Weaver W The Mathematical Theory of Communication. Chicago: University of Illinois Press, 1949:36–91.

46. Diamond GA, Hirsch M, Forrester JS, et al. Application of information theory to clinical diagnostic testing: the electrocardiographic stress test. Circulation 1981,63:915–21.

47. Moore WJ. Physical Chemistry, 3rd ed. Englewood Cliffs, New Jersey: Prentice-Hall, 1962:75–85, 245–50.

48. Peto R, Pike MC, Armitage P, et al Design and analysis of randomized clinical trials requiring prolonged observation of each patient. Br J Cancer 1976;34:585–612.

49. Edwards W, Lindman H, Savage LJ. Bayesian statistical inference for psychological research. Psychol Rev 1963;70:193–242.

50. Diamond GA, Forrester JS. Clinical trials and statistical verdicts: probable grounds for appeal. Ann Intern Med (in press).