A New Approach for Evaluating Risk Factors in Coronary Artery Disease: A Study of Lipid Concentrations and Severity of Disease in 1847 Males

DAVID W. SCOTT, PH.D., G. ANTHONY GORRY, PH.D., RAYMOND G. HOFFMAN, PH.D., JOSEPH J. BARBORIAK, SC.D., AND ANTONIO M. GOTTO, M.D., D. PHIL.

SUMMARY Using recently developed statistical techniques, we estimated joint probability density functions for plasma cholesterol and plasma triglyceride concentrations in 12 subgroups of 1847 males with angiographically demonstrated coronary artery disease. The subjects were grouped by patterns of age, history of hypertension, history of smoking and severity of disease. From the estimated joint probability density functions, we determined the way in which the risk of severe vs moderate coronary artery disease varied with the joint variation of plasma lipids for given combinations of the other risk factors. The patterns of risk that we found are evidence of the complicated way in which the relationship between plasma lipids and the severity of coronary artery disease is affected by other risk factors.

THE RISK of coronary artery disease associated with a particular characteristic such as elevated plasma cholesterol depends on the extent to which the characteristic is disproportionally associated with those who have the disease. When that characteristic varies substantially among both normal and diseased populations, as does plasma cholesterol, a thorough analysis of its role as a risk factor rests on the estimation of the probability density functions that embody the statistical variations. For any level of the risk factor in question, we can use estimates of these probability density functions to calculate an index of the risk of coronary artery disease. This index is the so-called likelihood ratio, the ratio of the probabilities of the particular level of the risk factor in diseased and nondiseased patients, respectively. Specifically, the revised odds of disease after observation of the risk factor equal the product of the likelihood ratio and the odds of the disease before the observation of the risk factor. Because the likelihood ratio is calculated from estimates of probability density functions, the methods of obtaining those estimates are very important.

The risk of coronary artery disease implied by a particular risk factor may be influenced by its interaction with other risk factors. For example, in a recent study of males with chest pain, the risk of coronary artery disease implied by elevated plasma triglyceride varied markedly with variations in the coexisting plasma cholesterol level. A key step in that study was the estimation of joint probability density functions, which represent the way in which plasma cholesterol and plasma triglyceride concentrations vary conjointly in diseased and nondiseased patients. From our estimates of these functions, we could assess the way in which the likelihood ratio, and hence the risk of coronary artery disease, varied with the joint variation of plasma lipid concentrations.

To estimate joint probability density functions adequately, new statistical methods may be necessary. These methods permit us to deal with multimodal probability density functions which, in contrast to the commonly assumed Gaussian density function, have more than one mode or local maximum. In this paper, we shall use such methods to study risk factors for the severity of coronary artery disease. Our approach rests on the statistical estimation of the conjoint variation of risk factors among a group of 1847 patients with angiographically demonstrated coronary artery disease. In the patients we studied, the Gaussian model, which underlies most analyses of risk factors, hides some interesting and potentially important patterns of risk factors that our techniques reveal.

Methods

Patient Selection

Data were collected from patients referred for a diagnostic arteriographic examination in two Milwaukee area hospitals and who volunteered to participate in the Milwaukee Cardiovascular Data Registry. The criteria for the referral were unstable angina pectoris, moderate-to-severe stable angina, previous myocardial infarction and recurrent chest pain of unknown etiology. The coronary arteriography was carried out by the Sones and Judkins technique. The angiograms were reviewed independently and by consensus by a cardiologist and a radiologist and were recorded accord-
ing to a uniform protocol. The extent of coronary artery occlusion was graded, as suggested by Rowe et al., with the modification that the scale was inverted and the occlusion score ranged from 0 (no occlusion) to 300. Fasting blood samples, collected after an overnight fast shortly before the arteriographic procedure, were analyzed for total plasma cholesterol and plasma triglyceride levels using the automated procedures with quality control monitored by the Lipid Standardization Program, Center for Disease Control, Atlanta, Georgia. Information on history of smoking, expressed in a five-point scale, history of hypertension, height and weight of the patients, was also obtained. We chose to analyze male patients only because the female patients represent only 14% of the population, so the groups would have been too small for proper statistical evaluation.

Estimation of Probability Density Functions

The 1847 males with occlusion scores of 1–300 were classified as "moderately diseased" if the severity of their disease was 1–174, and as "severely diseased" if it was 175–300. To study the relationship of plasma lipid levels to severity of disease in the presence of other risk factors, we also classified each of the 1847 patients into one of eight groups by age, history of hypertension and smoking level. The eight groups corresponded to the eight possible combinations of age (above or below the sample median of 52 years), history of hypertension (present or absent), and smoking level (above or below sample median) (table 1). The number of young males with a history of hypertension was too small to provide for an adequate estimate of the required density functions. For each of the other six groups, two bivariate (joint) probability density functions for lipid levels were estimated, one for moderately diseased patients and the other for severely diseased patients (table 2).

Probability density functions were estimated using nonparametric techniques suggested by Rosenblatt and developed by Parzen and Cacoullos. We have discussed these techniques previously. The procedure is analogous to the construction of a histogram, in this case in three dimensions. However, the procedure attains greater accuracy than a histogram by adjusting the cell boundaries for each datum rather than using the fixed cell boundaries. The procedure of Duin was used to estimate smoothing parameters for one bimodal group, and those values were extrapolated to the other cases.

We shall represent estimates of bivariate density functions by contour maps in which the surface corresponding to the probability density function is shown as seen from above with contours of equal probability suggesting its topography. These contours are formed by the intersection of 10 horizontal planes and the density surface. The highest plane intersects the density surface at one point, namely, the highest mode. The distance between the planes that determine adjacent contours for a given density function is one-ninth of the maximum height of that density function. In this report we use the highest eight contours to depict each density. Whenever a mode falls between two adjacent intersecting planes, it will be indicated by a point within the contour of the lower intersecting plane.

Calculation of Likelihood Ratios

From the ratio of the two joint probability density functions for plasma lipid concentrations in the moderately and severely diseased groups, we obtained a picture of the way in which the likelihood ratio (LR) varies with the conjoint variation of plasma lipids. The likelihood ratio function can also be depicted as a surface in three dimensions, with two dimensions corresponding to the respective plasma lipid concentrations and the third (height) to the likelihood ratio. On such a surface, peaks (LR > 1) represent relatively high risk, and valleys (LR < 1), relatively low risk. As is the case with density functions, it is convenient to represent these surfaces for likelihood ratios by contours. The most important contour corresponds to LR = 1, separating the cholesterol-triglyceride plane into regions of elevated risk (LR > 1) and decreased risk (LR < 1). Other contour levels are given for powers of the square root of 2 (LR = \ldots , 0.25, 0.35, 0.50, 0.71, 1.0, 1.4, 2.0, 2.8, 4.0, \ldots ).

If, for example, we consider a patient in whom the odds of severe disease a priori are 1 to 5, we can use the likelihood ratio to calculate the odds of severe disease given his plasma lipid concentrations. If the corresponding likelihood ratio is 2.5, then the revised odds are 2.5 × 1/5 = 1/2 or 1 to 2. A second patient selected from the same population (same a priori odds of severe disease) with plasma lipid levels that yield a likelihood ratio of 0.5 has revised odds of severe dis-

### Table 1. Profile of the 1847 Male Patient Characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Range</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Occlusion score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-50</td>
<td>108</td>
<td></td>
</tr>
<tr>
<td>51-100</td>
<td>273</td>
<td></td>
</tr>
<tr>
<td>101-150</td>
<td>357</td>
<td></td>
</tr>
<tr>
<td>151-200</td>
<td>474</td>
<td></td>
</tr>
<tr>
<td>201-250</td>
<td>438</td>
<td></td>
</tr>
<tr>
<td>251-300</td>
<td>197</td>
<td></td>
</tr>
<tr>
<td>B. Age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-24</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>25-44</td>
<td>227</td>
<td></td>
</tr>
<tr>
<td>45-64</td>
<td>1517</td>
<td></td>
</tr>
<tr>
<td>65-84</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td>C. Hypertension history</td>
<td>No</td>
<td>1173</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>674</td>
</tr>
<tr>
<td>D. Smoking level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>234</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>176</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>320</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>650</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>467</td>
<td></td>
</tr>
</tbody>
</table>
ease of $0.5 \times 1/5 = 1/10$, or 1 to 10. The relative risk for the two patients, which originally was the same, differs by a factor of 5 $(2.5/0.5)$ after their lipid values are known.

### Results

In table 2 we present the means and modes for all eight groups. We present the results of the likelihood ratio analysis for only six of eight groups defined by age, history of hypertension and history of smoking. (Again, too few young hypertensive males were available for adequate analysis.)

In the contour maps of the likelihood ratio for the six groups, only regions in which we have high confidence in the estimated density functions are shown. In such regions, the height of each density function exceeds one-ninth of its maximum height. Only modes in the upper four contour levels are shown. The means are shown by small crosses. The order of presentation of the groups reflects our general assessment of the importance of plasma lipids as risk factors.

#### Group 4: Older Males Without Histories of Hypertension or Smoking (fig. 1)

Each probability density function has one primary mode. The other modes are much lower and therefore are less important. The mean and mode in each density differ substantially, indicating non-Gaussian behavior. However, the relative positions of the means and modes in the two densities are similar, with no triglyceride differences and cholesterol differences of 14 mg/dl in the means and 24 mg/dl in the modes.

Using a Gaussian model for this group would lead to predictions of elevated risk for patients with cholesterol above 243 mg/dl, with little effect on that risk from elevated triglyceride. Indeed, the contours of the likelihood ratio show the region of elevated risk for high cholesterol. However, the few moderately diseased patients corresponding to the minor mode at (252, 270) gives rise to a region of relatively low risk in spite of the elevated cholesterol. Within the regions of decreased risk (LR < 1) are two distinct valleys where the likelihood ratio falls below 0.35. These are a result of the non-Gaussian features of the density functions for these patients.

#### Group 1: Older Males with Histories of Hypertension and Smoking (fig. 2)

Our analysis of this group shows the power of the present method. Although the mean vectors for the two densities are almost the same, the bimodal nature of the density for severely diseased males is strikingly non-Gaussian. The two modes for the severely diseased group differ by 60 mg/dl in cholesterol, and this difference has a marked impact on the likelihood ratio contours. Surprisingly, the highest risk (LR > 5) occurs for a relatively low level of cholesterol (210 mg/dl) and a moderately elevated triglyceride level (200 mg/dl). However, patients with the same triglyceride level but with a cholesterol level of 240 mg/dl are at low risk (LR < 0.35). Alternately, patients with the same cholesterol of 210 mg/dl but
FIGURE 1. Group 4 — older males with no history of hypertension and low smoking. Joint density functions of plasma cholesterol and plasma triglyceride for patients with severe coronary disease and patients with moderate coronary artery disease. The likelihood ratio is also represented by contours of equal likelihood for powers of the square root of 2. The region where the likelihood ratio exceeds one (elevated risk) are shaded. The contours of the likelihood ratio are shown only for regions in which sufficient data were available for the estimation of the density functions. The contours are labeled:

```
Contour       Likelihood ratio
-------------  ------------
-xxxx-         4.0
--xx-          2.8
-----x-        2.0
------        1.4
-------        1.0
--------       0.71
--------       0.50
---------      0.35
----------     0.25
```

with triglycerides of 115 mg/dl have a similar reduction in risk. The two modes in the density for severely diseased males represent increases in either cholesterol or triglyceride from the one primary mode for the moderately diseased males.

Group 8: Younger Males Without Histories of Hypertension or Smoking (fig. 3)

In this group our analysis reveals another interesting manifestation of multimodal density functions.

The single peak in the density function for the severely diseased patients is indistinguishable from a large peak in that for the moderately diseased group. The other primary mode for the latter group at (194, 95) is clearly distinct from the single mode for severely diseased patients, lower by 49 mg/dl in cholesterol and 32 mg/dl in triglyceride. The contours of the likelihood ratio function show that the levels of cholesterol and triglyceride, which correspond to reduced risk, form a rather narrow band from the

FIGURE 2. Group 1 — older males with a history of hypertension and heavy smoking. Joint density functions of plasma cholesterol and plasma triglyceride for patients with severe coronary disease and patients with moderate coronary artery disease. The likelihood ratios are represented by contours of equal likelihood.
lower left to the upper right. High risk (LR > 3.0) occurs for either elevated triglyceride or elevated cholesterol. The presence of a third small mode in the density for moderately diseased patients gives rise to a rather unexpected region of reduced risk for elevated lipids in the neighborhood of 300 mg/dl for cholesterol and 250 mg/dl for triglyceride.

Group 7: Younger Males with a History of Smoking but not Hypertension (fig. 4)

Both densities have two primary modes and a seemingly less important third mode. Portions of the density function for the severely diseased males are shifted toward higher triglyceride and higher cholesterol levels, implying elevated risk (LR > 1.0) for those values. The region of reduced risk is approximately a rectangle, with cholesterol ranging from 200–285 mg/dl and triglycerides less than 200 mg/dl. The apparent interaction of smoking and lipids in the younger, nonhypertensive males with severe disease seems to have a marked impact on lipid variations, as evidenced by the greater dispersion of the density contours. As a result, the two density functions overlap more and provide less information about risk.

Group 3: Older Males with a History of Smoking but not Hypertension (fig. 5)

This is the largest group (370 males) and so the likelihood ratio contours have the greatest statistical accuracy. There is a central band of moderately elevated risk with a rather small region of high risk around. Surprisingly, the regions of lowest risk correspond to elevated triglyceride concentrations. The lower mode in severely diseased patients leads to a region of elevated risk for 250 mg/dl cholesterol and 100 mg/dl triglyceride.

Group 2: Older Males with a History of Hypertension but not Smoking (fig. 6)

The density functions shown in figure 6 reveal an unusual situation. The leftmost mode in the severely diseased group is nearly 40 mg/dl lower in cholesterol than the single mode in the moderately diseased males. This is in sharp contrast to the relative
locations of the means, which are at higher values for more severe disease. Thus, cholesterol of 180–205 mg/dl and triglycerides greater than 100 mg/dl imply elevated risk, while higher cholesterol implies decreased risk, as seen in the contours of the likelihood ratio.

Discussion

A number of studies have addressed the relationship of specific risk factors to the probability of developing coronary heart disease. Generally, evidence of clinical coronary heart disease, such as myocardial infarction or cardiovascular death, has been correlated with risk factors at the point of entry into a longitudinal study. Several epidemiologic studies have used multivariate analysis to determine the independent contribution of various risk factors on the occurrence of coronary heart disease end points during the period of the study. These include the Framingham study, the study of Keys et al. carried out in seven countries, and the study of Medalie et al. In all of these studies, the risk associated with the initial cholesterol concentration was similar, a reduction of 9% being equivalent to a decrease of approximately 20% in the incidence of nonfatal myocardial infarction. Our study differs from an epidemiologic study in that we analyzed the relationship between the angiographically demonstrated degree of coronary narrowing and constellations of risk factors present at the time of that determination. A previous study has correlated risk factors with the severity of coronary artery disease, but the methods we use enable us to identify subgroups of patients in whom one risk factor may be of overriding importance or in whom there is an interaction between one or more risk factors.

Our approach involves the estimation of probability density functions from patient data and the calculation of likelihood ratios from those density functions. The smoothing of the estimates of the density functions is an important aspect of this approach, analogous to the selection of the width of the cells for a histogram. Too much smoothing eliminates features of the density function, while too little makes the estimate very sensitive to the vagaries of only moderately large samples. We approach optimal smoothing for the available data, through the use of the method suggested by Duin.

Figure 5. Group 3 — older males with no history of hypertension but heavy smoking. Joint density functions of plasma cholesterol and plasma triglyceride for patients with severe coronary disease and patients with moderate coronary artery disease. The likelihood ratios are represented by contours of equal likelihood.

Figure 6. Group 2 — older males with a history of hypertension but low smoking. Joint density functions of plasma cholesterol and plasma triglyceride for patients with severe coronary disease and patients with moderate coronary artery disease. The likelihood ratios are represented by contours of equal likelihood.
A second, less well understood problem is the construction of confidence intervals for the contours of the likelihood ratio. We have made some progress on this problem, but much remains to be done. Therefore, we emphasize only the gross features of the contour maps for the likelihood ratio. Even at this level, however, the subcategorization of patients with documented coronary artery obstruction into the risk factor groups could give insight into the pathogenesis and treatment of the disease.

The value of the categorization of patients is suggested by a recent prospective double-blind intervention trial in the primary prevention of ischemic heart disease using clofibrate.20 The study showed the importance of identifying groups of subjects with risk factors who may be at particular risk or who may especially benefit from a given form of treatment. In this trial, men younger than age 45 years received a greater benefit from treatment in terms of reduced incidence of ischemic heart disease than did older men, independent of other risk factors. Similarly, there was a greater decrease in nonfatal myocardial infarction in the treated group if the individual smoked or had a systolic pressure over 135 mm Hg. The subgroup of patients who would have benefited most from the clofibrate treatment were those of mean age 46.5 years, mean plasma cholesterol of 258 mg/dl, and systolic blood pressure greater than 135 mm Hg. This subgroup had a 34% reduction in major ischemic heart disease events, compared with a 23% reduction in the group that did not have these characteristics. Of some 15,000 subjects entered into the cooperative clofibrate trial, only about 6% fell into this subgroup. This is an important point because side effects have been observed with clofibrate as with all the other lipid-lowering drugs, and in fact some excess in total mortality was noted.

In our study, we expected the variation of the risk implied by plasma lipid levels to be clearest when other risk factors of smoking and hypertension were absent, that is, in groups 4 and 8 (figs. 1 and 3). In general, for patients in these groups, the risk is related primarily to the cholesterol level. Patients with cholesterol below 240 mg/dl cluster into a low-risk group, while those above 240 mg/dl are in a high-risk group. However, triglyceride levels contribute to the relative amount of risk in patients with cholesterol values above 240 mg/dl. This effect may be related to the inverse relationship between concentrations of triglyceride and HDL cholesterol. The inverse correlation between HDL and triglycerides was found in 70% or more of the subjects by Nikkila et al.21 As shown in Figure 3 for the younger patients (group 8), high risk exists for either elevated cholesterol or elevated triglyceride. The narrow band of low risk as the cholesterol and triglyceride values increase cannot be explained.

In the older males in group 1 (fig. 2), who have hypertension and who are smokers, a very interesting probability density function appears in those who are severely diseased. The density is clearly non-Gaussian. In the likelihood ratio, there are two widely separated regions of elevated risk. In fact, the highest risk occurs in a group with a relatively low level of cholesterol and a moderate elevation of triglyceride. We cannot explain why the patients with cholesterol at 240 mg/dl are at low risk. We may be seeing two groups of subjects, those at the upper end of cholesterol and triglyceride levels who are at quite high risk, and a second group who are at high risk despite low levels of cholesterol and triglyceride due to a long history of heavy cigarette smoking and hypertension, with a possible contribution of low levels of HDL. This is speculative, but the likelihood ratio function suggests these categories. However, we know from previous studies that only about 50% of coronary heart disease can be explained by known risk factors. There are obviously individual variations in the susceptibility to ischemic heart disease.

The likelihood ratio function shown in figure 4 represents young males of group 7 who are smokers but are nonhypertensive. In both the moderately and severely diseased males, there are two major modes in the density functions. It is of interest that patients with the more severe disease have a portion of their density function shifted to higher levels of cholesterol and triglyceride. Again, there is a group with relatively low values of cholesterol and triglyceride but which is at high risk, in this case, apparently primarily due to the overriding influence of smoking. Certainly there are individual variations in susceptibility to smoking in the development of ischemic heart disease, lung cancer and emphysema. Thus, subjects in one group who have cholesterol values from 200–285 mg/dl have not yet developed severe coronary heart disease, although they smoke. They probably have an inherited resistance to the effects of smoking and cholesterol on the arterial wall, at least until they reach an older age. This conclusion is strengthened by examining figure 5, which represents group 3, older males who smoke but who are not hypertensive. There is a central band of moderate elevation of risk with normal cholesterol and triglyceride values in whom the major risk factor would appear to be cigarette smoking. Again, there is a group of high-risk patients who are smokers with low levels of lipids and a group whose risk is increased by hypercholesterolemia and hypertriglyceridemia.

Figure 6 is a list of the values for group 2, older males who are hypertensive but are nonsmokers. A group of patients, perhaps with long-standing hypertension, develops atherosclerosis independent of cholesterol and triglyceride values. In some other studies, including the clofibrate trial, patients would be expected to derive more benefit from cholesterol reduction if they were hypertensive than if they were hypercholesterolemic and normotensive. Cholesterol tends to rise with age. This factor may contribute to the levels of cholesterol in this group. There is no known clinical, biologic or pathologic reason why increasing cholesterol in hypertensive patients should decrease risk. Such a conclusion would be inconsistent with all that is known from previous studies. We believe it is more likely that these patients have long-standing atherosclerosis, primarily because of
hypertension, and may also have hypertensive heart
disease.

In summary, our analysis demonstrates that triglyceride and cholesterol levels give information about risk of coronary artery disease severity. However, we have shown that complex interactions exist not only among the two lipid risk factors but also with the three other risk factors. We believe our method for analyzing multiple risk factors complements existing parametric approaches by providing evidence of the nature and complexity of several risk factors and their interactions. The patterns for the various combinations of these risk factors may suggest new facets in the complex relationship of lipoprotein metabolism and atherogenesis.

References

A new approach for evaluating risk factors in coronary artery disease: a study of lipid concentrations and severity of disease in 1847 males.
D W Scott, G A Gorry, R G Hoffman, J J Barboriak and A M Gotto

Circulation. 1980;62:477-484
doi: 10.1161/01.CIR.62.3.477

Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1980 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/62/3/477

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org//subscriptions/