PAPERS AND ORIGINALS

Who is at risk of a coronary?

T KHOSLA, R G NEWCOMBE, H CAMPBELL

British Medical Journal, 1977, 1, 341-344

Summary

The Joint Working Party of the Royal College of Physicians and the British Cardiac Society have recently given advice to medical practitioners about preventing coronary heart disease. We have formulated an easy method by which the medical practitioner may apply a rule to identify individuals who are at moderate to high risk of developing the disease. The simple arithmetic requires finding the product of four numbers and can be done with a pocket calculator costing less than £10. The four numbers correspond to known values of the four risk variables—age, systolic blood pressure, serum cholesterol levels, and smoking—for a given individual.

Tables of relative risk components for men aged 40 to 59 years for coronary heart disease defined by hard criteria (fatal or definite myocardial infarction) have been drawn up. The overall effect of the four risk variables is multiplicative so that some men aged 40 are at far greater risk than some of 55 year olds.

Introduction

About 40% of all deaths that occurred among men aged 45-54 years in England and Wales in 1973 were due to coronary heart disease (CHD). The mortality rate depends on (a) the attack rate and (b) the case fatality rate. CHD may occur suddenly, and the overall case fatality rate is about 40%. In a series of 998 fatalities in Belfast about a quarter of the deaths occurred within 15 minutes of the onset, and about 40% occurred within two hours.¹

Efficient ambulance services, intensive care units, and medical care at home may each play an important part in reducing the

case fatality rate among those surviving the initial onslaught. But the attack rate itself can be reduced only by appropriate preventive measures in groups of people who are judged, from known values of predisease characteristics, to run moderate to high risks of developing the disease.

The Joint Working Party of the Royal College of Physicians and the British Cardiac Society² recently published recommendations that practitioners can use to reduce the risk of coronary heart disease in selected patients. The report was aimed at a wide range of health workers (hospital and community physicians, nurses, paramedical staff, and community health workers). One of the major problems in the preventive approach is to identify which people are at high risk of developing CHD, because many precursors^{2,3} are implicated.

In 1972 Keys *et al*⁴ calculated the probability that an individual might develop CHD with the use of a multiple logistic equation. Five variables—age, systolic blood pressure, cholesterol levels, smoking, and body mass index (weight/height²)—were entered as predisease predictors in the subsequent development of CHD episodes within five years. Their follow-up study on 11 132 men aged 40-59 years included 2404 full-time employees of USA railways and 8728 men in geographically specified areas in Finland, Italy, Greece, the Netherlands, and Yugoslavia.

The risk prediction based on the European data gave good results on the American data and vice versa. They suggested that a doctor in the United States could easily calculate the probability of a patient developing CHD from known values of predisease characteristics with a desk computer. A leader in the *British Medical Journal* commented⁵: "Another possibility, which they do not suggest, is that tables could be constructed which would permit probabilities to be read off." The tables would increase the usefulness of the prediction equation in both Europe and the USA by identifying the high-risk individuals and thereafter showing that the incidence of CHD can be reduced by programmes which control blood pressure, serum cholesterol levels, and smoking as recommended by the Joint Working Party.² We have therefore consolidated the distributions of the identified risk factors of blood pressure, cholesterol levels, and smoking habits by age, along with the tables on the calculated probability of developing the disease in five years. We present here a simplified method, based on the relative risk factors, of identifying individuals at high risk who would benefit most from detailed counselling from their medical practitioners.

Department of Medical Statistics, Welsh National School of Medicine, Heath Park, Cardiff, CF4 4XN

T KHOSLA, MSC, PHD, senior lecturer in medical statistics

R G NEWCOMBE, MA, FSS, scientific officer

H CAMPBELL, FRCM, FSS, professor of medical statistics

Mathematical formulation

In the study by Keys *et al*⁴ age, systolic blood pressure, cholesterol levels, and smoking contributed significantly to risk. In their multiple logistic equation the probability of a given individual developing CHD in five years is estimated as:

$$P = (1 + \exp - (\alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4))^{-1}$$

where α is a parametric constant; X_1 , X_2 , X_3 , and X_4 are measurements of the predisease characteristics age, systolic blood pressure, cholesterol levels, and number of cigarettes smoked; and β_1 , β_2 , etc are the coefficients of the corresponding variables fitted by the method of Truett *et al*.⁶ Theoretically the limits of P are 0, no chance of CHD, and 1, absolute certainty of CHD.

In practice there is an upper limit of expected risk, corresponding to each of the risk factors taking its most unfavourable value. Thus in table IV case 4 has a relative risk of 90, which when multiplied by the baseline risk of 0.0041 gives 0.369, closely agreeing with the highest risk case of Keys *et al* of 0.377 (Keys *et al*; table IV).

In our paper we deal largely with the concept of relative risk (RR) defined as:

$$RR = \frac{\text{probability of developing CHD for a given individual}}{\text{probability of developing CHD for a standard person}} = \frac{P_i}{P_s}$$

where P_i and P_s are estimated from the logistic model of Keys *et al*.

We have used an approximate formula (see Appendix) to calculate the overall risk at specified levels of the risk variables because innumerable cases can be defined with different combinations and levels of risk characteristics of age, systolic blood pressure, cholesterol level, and number of cigarettes smoked. For systolic blood pressure and cholesterol five classes are taken with boundaries close to the

TABLE I—Percentiles of serum cholesterol levels (mmol/l) in men (USA 1960-2)

Percentiles	Age (years)		
	35-44	45-54	55-64
10	4.46	4.66	4.51
20	4.95	5.08	5.08
30	5.26	5.34	5.44
40	5.54	5.59	5.78
50	5.80	5.91	6.01
60	6.06	6.16	6.27
70	6.35	6.53	6.55
80	6.76	6.92	6.89
90	7.41	7.43	7.36
95	8.03	7.82	8.08
99	9.32	8.55*	9.32

*Irregularity can be attributed to sampling variation.

Conversion: SI to traditional units—Cholesterol: 1 mmol/l $\approx$ 38.6 mg/100 ml.

TABLE II—Percentiles of systolic blood pressure (mm Hg) in men (USA 1960-2)

Percentiles	Age (years)		
	35-44	45-54	55-64
10	110	113	115
20	114	119	122
30	119	124	127
40	123	129	133
50	127	133	139
60	130	137	144
70	135	142	150
80	141	149	162
90	152	159	169
95	161	168	180
99	180	190	205

TABLE III—Relative risk factors associated with predictive variables

Age (years)	Hard CHD		Cholesterol		Hard CHD		Systolic blood pressure (mm Hg)	Hard CHD		Smoking (No of cigarettes/day)	Hard CHD	
	USA	Europe	(mmol/l)	(mg/100 ml)	USA	Europe		USA	Europe		USA	Europe
40-44	1.00	1.00	<4.66 (4.40)	<180 (170)*	0.74	0.71	<115 (110)*	0.73	0.77	None	1.00	1.00
45-49	1.36	1.47	5.93 (5.31)	229 (205)	1.00	1.00	129 (123)	1.00	1.00	<5	1.25	1.11
50-54	1.84	2.2	7.49 (6.73)	289 (260)	1.61	1.72	159 (145)	1.70	1.54	5-9	1.56	1.22
55-59	2.4	3.2	7.74 (7.64)	299 (295)	2.2	2.4	169 (165)	2.80	2.3	10-19	1.95	1.36
	2.5		≥7.77 (8.55)	≥300 (330)	3.0	3.4	≥170 (190)	5.0	3.7	20-29	2.4	1.50
										≥30	3.0	1.66

*Figures in parentheses represent mid-range values for calculating RR factors.

10th, 50th, 90th, and 95th percentiles derived from the distributions in the American random sample of men aged 45-54 years^{7, 8} (tables I and II). For these two variables the category of individuals from 10th to 50th percentile was chosen as the base for relative risks; this baseline (or standard) provides good stability because of large numbers.

Our approximate formula simplifies the computation in the practical applications, so that it will be necessary only to multiply together four numbers found in the tables. The results of the approximate formula and the exact values as computed by the model of Keys *et al* agree well in the range of relative risk (RR) values considered.

Results

Table III gives the relative risk for hard CHD (fatal or definite myocardial infarction) associated with the predictive factors in American and European men aged 40-59 years. The four factors derived from the table are to be multiplied together for a given male with known characteristics:

$$\text{Combined RR} = \text{RR (age)} \times \text{RR (blood pressure)} \times \text{RR (cholesterol level)} \times \text{RR (smoking)}.$$

Calculating the combined RR is a simple task with a pocket calculator costing less than £10.

For example, table IV gives the results of some hypothetical cases compared with the baseline of RR=1 (age 40-44, non-smoker, and normal blood pressure and cholesterol levels) for case 1. The RR in case 4 is 90 (age 55-59, heavy smoker, and blood pressure and cholesterol levels at the 99th percentile). In other words case 4's risk is 90 times that of case 1. Men aged 40-44 with abnormal values of the risk variables carry a greater relative risk (case 2 RR 20.2) than men of 55-59 years with normal or below normal values of the risk factors (case 3 RR 2.5).

The relative contributions of the risk predictors are depicted in the figure, which compares the trend in the calculated risk per 1000 men (USA hard CHD) by age in eight groups (table V) categorised by low-high levels of the predictive variables: systolic pressure—110 v 160 mm Hg; cholesterol—4.66 v 7.51 mmol/l (180 v 290 mg/100 ml); smoking—none v ≥20 cigarettes/day. The computed probabilities in table V are based on the exact values derived from the logistic model of Keys *et al*. The low-high levels approximately represent the cut-off points at the 10th and 90th percentiles of the systolic blood pressure and the serum cholesterol level distributions (tables I and II).

The results in table V and the figure show that among the eight selected groups, the risk per 1000 ranges from 2.3 (non-smokers aged 40 with cholesterol 4.66 mmol/l (180 mg/100 ml) and systolic pressure 110 mm Hg: group A1), to 161 (heavy smokers aged 60 with cholesterol 7.51 mmol/l (290 mg/100 ml) and systolic pressure 160 mm Hg: group D2). In table V the ratio of the two extreme comparisons 161/2.3 gives a relative risk of 70, which is quite close to the RR value calculated by the approximate method $36.5/0.49=75$; here the individual RR factors have been found from table III by linear interpolation and where necessary extrapolation. It is not, however, intended that the clinician should interpolate. Table III can be extended to include more groupings of the risk variables. Here we are demonstrating the goodness of the approximate method. Higher cholesterol levels and blood pressures than those in group D2 would indicate a probability even higher than 161. For each of the eight groups in the figure the trends with age are linear and parallel on a semi-logarithmic scale. The gap of constant width between two lines represents the ratio of probabilities or relative risks between the two groups, which the model assumes to remain constant with age. For example, the approximately constant and equal ratios (A2:A1, B2:B1, etc) represent the relative risk of heavy smokers. Similarly B2:A1,

TABLE IV—Relative risk of developing hard CHD in five years (hypothetical cases). Figures in parentheses are RR values from table III for USA

Case No	Age (years)	Systolic blood pressure (mm Hg)	Serum cholesterol (mmol/l)	Smoking (No of cigarettes/day)	Relative risk	
					USA	Europe†
1 (Baseline)	40-44 (1)	120 (1)	5.18 (1)	None (1)	1*	1*
2	40-44 (1)	165 (2.8)	8.03 (3.0)	25 (2.4)	20.2	11.7
3	55-59 (2.5)	120 (1)	5.18 (1)	None (1)	2.5	3.2
4	55-59 (2.5)	190 (5.0)	8.81 (3.0)	≥20 (2.4)	90	60.4

*For USA data RR = 1 represents a probability of 0.0041 and for the Europeans 0.0031.
†RR for Europeans is calculated similarly.

TABLE V—Risk per 1000 men of hard CHD (USA experience) at specified levels of low-high values of risk variables

Group	Systolic blood pressure (mm Hg)	Serum cholesterol (mmol/l)	Smoking (No of cigarettes/day)	Age (years)		
				40	50	60
A1	Low (110)	Low (4.66)	0	2.3	4.3	8.1
A2	Low	Low	≥20	5.8	11.0	20.6
B1	Low	High (7.51)	0	6.2	11.6	21.7
B2	Low	High	≥20	15.7	29.3	54.0
C1	High (160)	Low	0	7.6	14.2	26.6
C2	High	Low	≥20	19.3	36.0	65.8
D1	High	High	0	20.4	37.8	69.2
D2	High	High	≥20	50.8	92.0	160.7

Relative contributions of risk predictors. Groups are defined in table V (USA data).

D2:C1, etc represent the relative risk of high serum cholesterol levels. The interpretation of any particular ratio—for example, D2:A1—in the figure is made with reference to the identity of the comparison groups in table V. The contributions of smoking, cholesterol levels, and systolic blood pressure are not dissimilar, but a high systolic blood pressure seems to contribute most to high risk. The components of relative risk for the selected comparisons are: smoking (heavy *v* non-smokers) 2.5; cholesterol (7.51 *v* 4.14 mmol/l (290 *v* 160 mg/100 ml)) 2.6; systolic blood pressure (160 *v* 110 mm Hg) 3.2; age (60 *v* 40 years) 3.4.

The horizontal lines in the figure would give contours of constant risk at different ages. Smokers incur a relative risk of 2.5, which is equivalent to that of an extra 15 years of age, so that non-smokers aged 60 (group A1) are at similar risk to smokers aged 45 (group A2).

Thus there is an increase in risk with advancing age, and the other risk factors can be regarded as each equivalent to a certain number of extra years of age. The effect of these other factors can be interpreted as an acceleration of the aging processes throughout adult life. Thus there are groups of men (B2, C2, D1, D2) who at age 40 experience far greater risks than group A1 when aged 60 (table V and figure). The former groups are men for whom two or three factors are unfavourable.

Discussion

The report of the Royal College of Physicians³ calculated the age-standardised relative risk of coronary heart disease from each of several predisease characteristics within groups of non-smokers and smokers. Each characteristic was dichotomised into broad groups (low *v* high levels; absence *v* presence; etc). The risk factors included in the report were systolic blood pressure, blood cholesterol concentrations, physical activity, social mobility, and behaviour type. The relative risk from obesity was not reported.

The multivariate study of Keys *et al*⁴ included five variables: age, systolic blood pressure, serum cholesterol, smoking, and body mass index. Of those five, body mass index, defined as a function weight/height², did not contribute significantly to the overall risk in the presence of other variables in either the European populations or the American population.

This apparent insignificance of obesity (or large body mass index) as a risk factor is surprising, but in a multivariate analysis the interpretation of a non-significant factor in the presence of other significant characteristics needs caution. Body mass index is correlated with both age and blood pressure. The finding that excess weight by itself does not increase significantly the risk of coronary heart disease may imply that if there is an increased risk it is manifested through the identified risks of age and systolic blood pressure. From an insurance study⁹ reported in 1959 and pertaining largely to the period before the second world war the evidence is overwhelming that excess weight is a risk factor in overall mortality and in coronary heart disease. For this reason extra premiums are added to the insurance policies of those who are overweight.

An additional complex interrelationship has emerged between obesity and smoking. Our large-scale survey of steel workers in South Wales¹⁰ has shown a considerable difference between the body weight of smokers and of non-smokers. Men aged over 40 years who had never smoked were on average 5.9 kg heavier than smokers. Even so, smokers were about 6.8 kg heavier than the weight considered desirable by the Metropolitan Life Insurance Company, while non-smokers were 12.7 kg heavier. Since the war obesity has almost become a “norm” in the affluent parts of the world. But there is a danger in underestimating the risk of obesity because of its inverse relation with smoking.

Smoking is associated with many diseases, such as chronic bronchitis, lung cancer, reduced pulmonary function, and peripheral vascular disease. These, together with coronary artery disease, account for almost half of all deaths to men aged 45-64. The manifestation of a known risk such as obesity is hidden because groups of people who are overweight—that is, have a large body mass index—will contain a larger proportion of non-smokers than will be included in groups of people who are of average or below average weight. Detailed analysis of body mass index within groups of smokers and non-smokers would identify the extra risk of death from CHD that can be attributed to excess weight in both smokers and non-smokers, but there have been few such studies.

There is evidence that on average men in England and Wales have got heavier recently.¹¹ The sedentary way of life has induced obesity in the general population. Men aged 30-40 years working in Port Talbot in 1965 were about 4.5 kg heavier

than men of the same age and height working in Birmingham in 1960, and in their turn Birmingham factory workers were more than 4.5 kg heavier than the national samples of factory workers measured in 1943 and in 1935.

Reduced physical activity has been incriminated by several studies as an important risk factor, but physical activity was omitted from the model of Keys *et al*, apparently because they found it impossible to measure reliably in their prospective study. In this context the comments of the leader in the *BMJ*⁵ are relevant:

"The evidence that physical inertia at work is important has been growing through the years, and to this have been added studies which indicate that physical inertia during leisure hours is also important. It does not require elaborate epidemiological research to show that there is a close interrelationship between physical inertia, body bulk, and consumption of food, so that these secondary risk factors must act through the three important primary risk factors—age, blood pressure, and serum cholesterol—the three which apparently can effectively be used to determine who will get a coronary attack."

The aetiology of coronary heart disease is multifactorial, and a multidimensional approach is needed in the preventive programmes. Apparently healthy people (but with high cholesterol levels or a heavy smoking rate, or both) leading physically very active lives may experience fatal CHD.^{12 13}

Type A behaviour, the aggressive, assertive type, is an important determinant of coronary heart disease³ and studies have shown that the relative risk compared with that of type B, phlegmatic and passive, may be about 2.5. Some combinations of the four identified risks in table III may, however, give a multiplicative relative risk of hard coronary heart disease of the order of 90. It is not at all clear whether temperament and personal behaviour are additional risks or whether they are risks that have already manifested themselves through the basic risk factors. Further studies are required to throw light on this.

The relative risk attached to the four identified variables in table III and their application (table IV) on specific subjects would be of value to the doctor and for screening purposes in identifying the high-risk individuals. A doctor can calculate the overall relative risk for a patient from the four factors concerned but he will also have the patient's medical history before him including weight, personality type, physical activity, family history, and he should be able to use his knowledge of these factors in counselling a patient on how to reduce the relative risk of coronary heart disease.

Prospective studies are needed to substantiate the extent of benefit provided by intervention programmes in groups of people who run moderate to high risks of developing the CHD. In the meantime some positive action is necessary because the "price of inaction"¹² is high.

We thank Mr John Zlosnik and Mr Peter Samuel for their help with the computer program; and Dr Iain Chalmers for his constructive comments and suggestions.

References

- 1 McNeill, R H, and Pemberton, J, *British Medical Journal*, 1968, 4, 139.

- 2 Joint Working Party of the Royal College of Physicians of London and the British Cardiac Society, *Journal of the Royal College of Physicians of London*, 1976, 10, 213.
- 3 Royal College of Physicians, *Smoking and Health Now*. London, Pitman, 1971.
- 4 Keys, A, *et al*, *Circulation*, 1972, 45, 815.
- 5 *British Medical Journal*, 1973, 2, 375.
- 6 Truett, J, *et al*, *Journal of Chronic Diseases*, 1967, 20, 511.
- 7 Moore, F E, and Gordon, T, *Serum Cholesterol Levels for Adults, United States 1960-62*. Washington, National Centre for Health Statistics, 1967.
- 8 Gordon, T, *Blood Pressure of Adults by Age and Sex, United States 1960-62*. Washington, National Centre for Health Statistics, 1964.
- 9 *Build and Blood Pressure Study*. Chicago, Society of Actuaries, 1959.
- 10 Khosla, T, and Lowe, C R, *British Medical Journal*, 1971, 4, 10.
- 11 Khosla, T, and Lowe, C R, *Lancet*, 1968, 1, 742.
- 12 Opie, L H, *Lancet*, 1975, 1, 263.
- 13 Khosla, T, *Lancet*, 1975, 1, 395.

Appendix

Keys *et al* have used the multiple logistic model:

$$P = \{1 + \exp -(\alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4)\}^{-1}$$

where P denotes the probability that an individual will develop CHD during the next five years, X_1 , X_2 , X_3 , and X_4 are the predictive variables, and α , β_1 , β_2 , β_3 , and β_4 are parameters fitted in the regression.

The relative risk (RR) associated with an increase "h" in factor X_1 , while X_2 , X_3 , and X_4 are all held constant, is about $\exp \beta_1 h$. We make this approximation to simplify computation so that it will be necessary only to multiply together four numbers given in table III. This approximation will be good if the range of P over which we are working is low, as then the logistic model is practically multiplicative, being essentially:

$$P = \exp (\alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4)$$

In general, however, the extra complexity of the logistic model is required to correct the behaviour of P as it becomes larger. The relative risk associated with an increase "h" in factor X_1 , holding the other factors constant, can be shown to be:

$$RR = \frac{C + \exp \frac{1}{2} \beta_1 h}{C + \exp -\frac{1}{2} \beta_1 h}$$

Here C denotes $A/(1-A)$, where A is the incidence of CHD at the fixed values of X_2 , X_3 , and X_4 , and with X_1 midway between the values being compared. In practice A can be taken as the incidence in the whole series. It is this formula that has been used to calculate the factors quoted in table III.

The figures in our table III are based on those in table II of Keys *et al*. Body mass index has not been included as a screening variable, since its contribution to the regression was not statistically significant. The effect of deleting body mass index from the regression can be assessed by comparing the footnotes to table IV of Keys *et al* with row 2 of their table II. The effect on the discrepancy of factors from unity is less than 7%.

Nothing can be said about the effect of deleting body mass index on European hard data as Keys *et al* do not present an appropriate table analogous to their table IV. Certainly the effect of deleting body mass index will cause changes well within the confidence limits of the factors.

The results of the approximate formula and the exact values as computed by the model of Keys *et al* are in good agreement in the range of relative risk (RR) values considered.

(Accepted 2 December 1976)