

## Use of risk factors to allocate schedules for breast cancer screening

FREDA E ALEXANDER,<sup>1 5</sup> M MAUREEN ROBERTS,<sup>2 3</sup> ANNE HUGGINS,<sup>3</sup> AND BERENICE B MUIR<sup>3 4</sup>

*From the Medical Statistics Unit<sup>1</sup> and Departments of Clinical Surgery<sup>2</sup> and Medical Radiology,<sup>4</sup> University of Edinburgh, and Edinburgh Breast Screening Clinic<sup>3</sup> and Leukaemia Research Fund Centre for Clinical Epidemiology<sup>5</sup> at the University of Leeds, Department of Pathology*

**SUMMARY** In an effort to reduce the cost of breast cancer screening several studies have explored the possibility of using risk factors to select a high-risk group of women and then restrict screening to that group. The results of these studies have been almost entirely negative and so it is not possible at present to classify any woman as at such low risk that she need not be screened. Nevertheless it is well known that some groups of women can be identified as being at higher risk than the general population. In this study it is assumed that each woman will be offered one screen at which risk factor information will be collected. The usual screening policy is then one of uniform intervention in which the interval to the next screen is the same for all women: the interval that is currently recommended in the UK is three years. An alternative is a risk strategy in which the time to the next screen depends on the woman's risk status; thus the total number of screens available to the population are distributed according to risk status. Using data from the Edinburgh randomised trial of breast cancer screening these policies have been compared. It is estimated that the proportion of cases detected by screening in the three years following the completion of the initial screening round could be raised from 60% to 67% by adopting a risk strategy. Lead time benefits are also quantified as are the comparisons for an established screening programme.

In the Edinburgh randomised trial of breast cancer screening,<sup>1</sup> half of the geographical female population aged 45–64 has been randomly selected for an invitation to screening by 2-view mammography and clinical examination. Thereafter repeat invitations are issued annually to those women who accept the initial invitation; screening is by annual clinical examination and biannual single oblique view mammography given two, four, and six years after the initial screen. Five years of follow-up have been completed. After seven years of follow-up breast cancer mortality in the population who were offered screening will be compared with that in the control population. Randomised trials in other countries<sup>2 3</sup> have already demonstrated a significant reduction in breast cancer mortality in a population invited to screening. The Forrest report<sup>4</sup> has recommended population screening in the UK for women aged 50–64 by mammography and at three yearly intervals.

The current trials in the UK<sup>1 5</sup> are expected to report the results of their mortality comparisons in

1988/89. These will describe the effect of offering annual clinical examination and biannual mammography to British women. Such frequent screening is a major commitment of public resources. Yet it is not certain that an alternative programme with less frequent screening will have a similar impact on mortality. The issues of cost-effectiveness are complex; after considering relevant information the Forrest report recommended the interval of three years between screens. This would be the same for all women of the appropriate age. Since, however, not all women have the same risk of disease it is possible that a policy of allocating women to schedules with different intervals between screens would be preferable. This study examines this possibility.

### Methods

The first step was to derive a discriminant function using the 14 089 women who were cleared of breast cancer at their first screen in the Edinburgh breast

screening programme. In the subsequent five years, 137 of these women were diagnosed as having breast cancer; these were the "cases", and the remaining 13 952 are the "controls". This is the "initial cohort" of the trial. We are confident that the data on subsequent incidence are virtually complete since the women are followed up by two methods: by reference to the pathology registers of Edinburgh hospitals and by flagging at the Scottish General Registry Office. Altogether 100 of the cases were detected at later screening rounds, but all cases arising in the initially "healthy" population were included in the analysis.

The data were prepared for analysis using Fortran 77 programs and the discriminant function was derived by logistic analysis with the statistical package GLIM.

The second step was to use the discriminant function to predict subsequent risk. Its ability to predict would be overfavourably assessed if the existing data were used for this purpose; this is the statistical problem of "shrinkage". So the discriminant function coefficients were preshrunk using the method of Copas (equation 8.5 in ref. 7). Each regression coefficient was multiplied by 0.9 and then the constant term was adjusted to ensure that the mean predicted risk of future disease was correct. This yields a function which is an approximately unbiased predictor of future risk.

As the sample of 14 089 was large we were able to assume that it accurately represented the distribution of the risk factors in any similar future population. The application of the predictor to this population provided estimates of future disease.

The third step was to split the population into three groups: lower, medium, and higher risk groups (the lower and higher risk groups being of equal size). Using a model<sup>8</sup> which has been derived by modifying that of Day and Walter,<sup>9</sup> we were able to estimate the proportions of cases detected and the average amount by which their diagnosis was advanced in two alternative hypothetical screening programmes:

- (a) screening by mammography at three-yearly intervals for all women, and
- (b) screening by mammography at 18-monthly intervals for the high-risk group, at three-yearly intervals for the medium-risk group, and at six-yearly intervals for the low-risk group.

These were evaluated and compared over the first three years after the start of screening, over the first six years, and in a long-established screening programme.

The details of applying the model to the present situation and the assumptions involved are given in the appendix.

## Results

The proportions of cases and controls in each age group and with each risk factor are shown in tables 2 and 3. Crude odds ratios are given in table 3. All are

Table 1 Risk factors and their coding

Factor	Code	Meaning
Age at menarche	-1	15,16
	1	12,13,14
	0	other, including unknown
Age at first full-term pregnancy/parity	-1	FFTP <= 20
	0	FFTP 21-29
	1	FFTP > 30 or nulliparous
Own breast history	0	None
	1	Some
Family history of breast cancer (in mother/sister)	0	None
	1	Some
Menstrual status* (age 45-54)	1	Premenopausal and < 50, or postmenopausal
	2	Menopausal, premenopausal >= 50 or artificial menopause
Age at menopause (age >= 55)	0	Age known to be < 50
	2	Age known to be >= 55
	1	Others (including those reporting a hysterectomy)
Mammographic pattern	0	No
	1	P1, indeterminate
	2	DY, P2
Referral by mammogram reader	0	None
	1	Referred but cleared by radiologist
	2	Some further investigation

\* The aim is to identify women who may be menopausal or menstruating when aged over 50 at the time of later visits from information gathered at this visit. In practice, we took all women who were menopausal or who were premenopausal and aged at least 50 or who had had a previous hysterectomy (since no information was available on the state of the ovaries). For simplicity we did not allow for premenopausal women aged under 50 entering this group as they became older; this in fact lessens the power of our discriminant and weakens our results.

Conventional risk factor information obtained at routine interview and the results of the mammography at the first screen were collected for both cases and controls. The factors that were used are shown in table 1; the epidemiological factors are established risk factors motivated by the previous study of risk factors from the Edinburgh project.<sup>6</sup> The factors were coded in three categories wherever appropriate so as to facilitate the derivation of positive information on low as well as on high risk. The codes were assigned prior to any analysis.

Table 2 Age distribution of cases and controls

Age at first screen	Cases	Controls
45-49	32 (23)	3724 (27)
50-54	36 (26)	3736 (27)
55-59	39 (28)	3672 (26)
60-64	30 (22)	2850 (20)

Percentages in parentheses

Table 3 Risk factors in cases and controls

Factor	Level	% Cases	% Controls	Odds ratio
Age at menarche	15,16	13	20	0.8
	rest	12	15	1.0*
	12,13,14	75	65	1.4
Age at first full-term pregnancy/parity	≤20	2.9	8.4	0.4
	21-29	51.1	56.2	1.0**
	≥30 or nulliparous	46.0	35.4	1.4
Own history of benign breast disease	Some	20	15	1.3
Family history	Some	13	6.4	2.0**
Mammographic pattern	No	15	24	0.6
	P1, etc	61	56	1.0+
	DY,P2	23	20	1.1
Referral of mammogram	None	76	89	1.0
	Some (code 1)	18	10	2.1**
	Some (code 2)	6	3	2.3
Menstrual status	Possibly menopausal	44	32	1.7*
Age at natural menopause	< 50	36	44	0.8
	Neither	51	51	1.0*
	≥55	13	6	2.3

The odds ratios are all calculated without adjustment for any other factors.

\* p < 0.05

\*\* p < 0.01—standard chi-square test

+ p < 0.05 chi-square 1 df test for trend

Table 4 Multivariate logistic analysis

Factor	Regression coefficient (SE)	Odds ratio (adjusted for all other factors)
Age at menarche	0.28 (0.12)	1.23
Age at first full-term pregnancy	0.39 (0.15)	1.48
Family history	0.68 (0.26)	1.97
Own breast history	0.22 (0.21)	1.25
Referral at first screening	0.48 (0.14)	1.62
Mammographic pattern	0.22 (0.13)	1.25
Age at menopause (women > 55)	0.38 (0.20)	1.46
Menstrual status (women < 55)	0.49 (0.24)	1.63

Overall significance of the model:  $\chi^2 = 46.8$  on 9 df; p < 0.0005

statistically significant apart from a woman's own past history of breast disease.

The factors were then combined in the logistic regression; table 4 gives the results showing the regression coefficients, their standard errors, and the odds ratios adjusted for all other factors. As the aim is discrimination, all factors have been retained even though in several cases their independent contribution does not attain conventional levels of statistical significance.

The regression coefficients were preshrunk as described earlier. Then, for each individual, a score

Table 5 Predicted rates of disease in three risk groups

(Rates are averages over a five-year period in a screened population following a negative prevalence screen)

	Risk group		
	High	Medium	Low
Rate/1000 women-yr	3.70	1.73	0.92
Size of group (% of whole)	20%	60%	20%

Note: These correspond to the observed rate of 137/14000 women-years and are higher than would be expected in either an unscreened population or a population with an established screening programme. Corresponding rates for these are 3.52, 1.64, and 0.87 but deriving these requires assumptions that are described in the appendix.

was calculated by adding together the appropriate coefficients; this was the natural logarithm of her predicted future risk of disease. Figure 1 shows the cumulative proportions of cases and of the general population plotted against score. The curves diverge in the central region, showing pictorially the amount of discrimination that is obtained.

Examination of fig 1 suggests that appropriate sizes for high- and low-risk subgroups of the population are 20%. The predicted risks were averaged over each of these and over the medium-risk group, and the results are shown in table 5.

These were taken as underlying incidence rates and applied to the mathematical model developed in ref. 8. Maximum likelihood estimates\* and joint confidence regions for the sojourn time and false negative rates

\* These were: false negative rate = 0.37  
mean sojourn time = 4.98 years

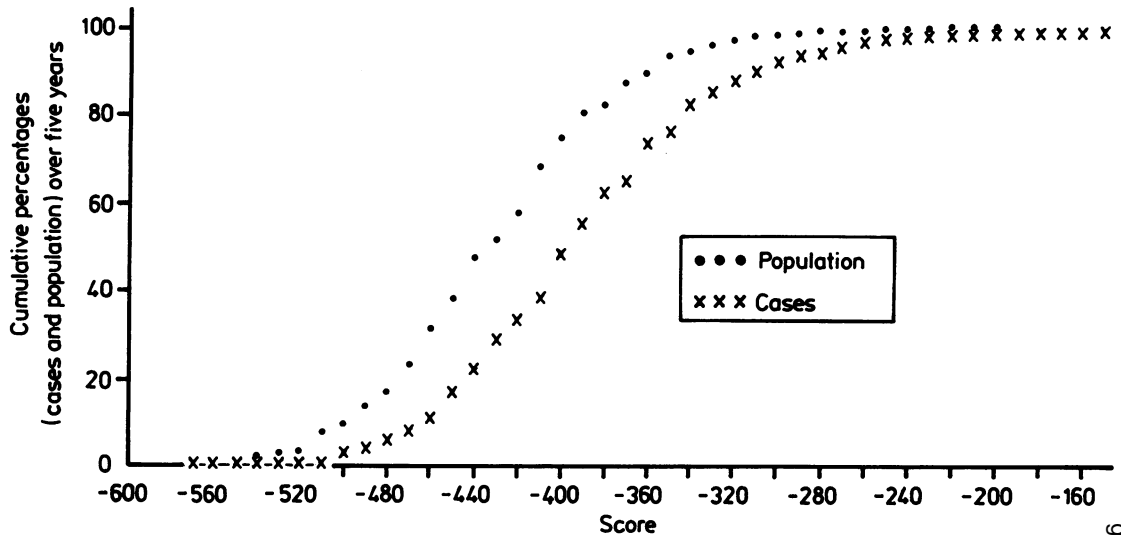


Fig 1 Performance of the discriminant function (predictions for future populations)

Table 6 Proportions (%) of cases detected  
(All cases are those arising after the prevalence screen)

Time	Uniform intervention	Risk strategy	Difference
First 3 yr	59.9 (50–68)	67.3 (59–74)	7.35 (5.7–10.2)
First 6 yr	61.4 (51–71)	68.8 (60–76)	7.2 (5.1–10.1)
Steady state	64.4 (52–75)	70.4 (60–79.5)	6.0 (4.1–7.9)

For the present programme of biannual mammography and annual clinical examination estimates for the first six years and the steady state are 78% and 80% respectively.

Table 7 Mean lead time (years) gained by screening

Time	Uniform intervention	Risk strategy	Difference
First 3 yr	2.5 (1.5–3.7)	2.8 (1.7–4.1)	0.35 (0.23–0.55)
First 6 yr	2.6 (1.5–3.8)	2.9 (1.8–4.2)	0.35 (0.22–0.55)
Steady state	2.7 (1.6–4.0)	3.0 (1.8–4.4)	0.32 (0.21–0.47)

derived in ref. 8 were used to yield estimates of the proportion of cases detected and of the average length of “lead time” (time by which diagnosis was advanced) for the uniform intervention policy and for the risk strategy. The results are shown in tables 6 and 7. This suggests that the risk strategy increases the proportion of cases detected by 7% in the first three-year period without any increase in the number of screens performed. Subsequently both proportions would gradually increase, but the difference between them is only marginally changed. Similar results hold for the

mean lead time. Beyond the six-year period the risk strategy involves a slight increase in the number of screens as the low-risk group is re-screened.

## Discussion

The most obvious way of using risk factor information to reduce the cost of breast cancer screening is to use it to select a high-risk group with the aim of restricting screening to that group. Our previous study<sup>6</sup> showed that this is not possible, at least at present. Similar results have been reported by others.<sup>10–12</sup> Yet the need to minimise the resources committed to successful screening is clear.

Only one recent study has been more optimistic about the possibility of using risk information to select a group of women who need not be screened.<sup>13</sup> This is, however, restricted to volunteer women in whom 50% of the cases have experienced symptoms and so its results must be treated with caution. Moreover it uses sufficiently complex criteria for it to be necessary for a woman to attend for examination to determine whether she is in the high-risk category. It is doubtful whether in such circumstances a woman could ever remain unscreened, though she could perhaps be offered a simpler screen.<sup>6</sup>

Thus it is appropriate for this study to examine a policy in which all women are offered an initial mammographic screen. At that time a variety of information would be collected but we have included only those factors that could be quickly and easily identified at a routine screening visit. The eventual aim

*Use of risk factors to allocate schedules for breast cancer screening*

is to reduce cost, and this would certainly not be achieved if the screening process were made more complex. We would anticipate a computerised record system so that the relatively complex calculations would be performed almost immediately once the results of screening were known. The computer would then automatically issue the next invitation to screening. Use in an actual screening programme would of course be the final test of whether this is feasible or acceptable.

The epidemiological factors we have used are conventional risk factors which are usually associated with aetiological risk. The same is probably true of mammographic pattern. The opinion of the reader of the initial mammogram is entirely different. One of us (MMR) has reviewed the case-notes of every one of the "cases" whose mammogram was referred for radiological opinion. There was no indication that these represented cases that were missed. Thus no "clinical" importance should be attached to this factor. It is only to be seen as one non-specific indicator of subsequent disease which is to be used with others in the scoring system dealing with large numbers, not with individuals.

The discriminant function performs well in selecting a high-risk group—where 45% of the predicted cancers belong to a high-risk group of just 20% of the population. The predicted rate of future disease in the next few years in this group is 3.7/1000 women-years. An interval of three years between screens is likely to be too long for them. It is less good at selecting a truly low-risk group. Here 20% of the population represents 7% of the cases, but the rate of disease in this group is 0.8/1000 women-years, which is far from negligible. These women certainly require some repeat screening.

We have used theoretical methods to quantify the benefits that could be derived from a risk strategy using risk information available at present. The use of the risk strategy would increase the proportion of cases detected by screening in the few years after the start of a screening programme from approximately 60% to approximately 67%. The order of magnitude represents a useful improvement, but to put it in context it should be remembered that the corresponding estimate for the present programme is 78%.

Three conclusions are suggested by this study. Firstly, risk factor information should be collected in some new programme with a three-yearly interval so as to permit prospective evaluation. Secondly, some research should be directed to factors that will positively identify a truly low-risk population of reasonable size.<sup>14</sup> Thirdly, the "high-risk" group should be identified and monitored in future screening programmes to see whether, for them, the three-yearly interval between screens is indeed too long.

Few assumptions have been made in the earlier part of this study except that we disregard any age effect. We did this because at present no definite age relationship is

evident in Edinburgh, but more work is certainly necessary to clarify this. Several assumptions have been made in the modelling process which led to estimates of the proportion of cases detected and of lead time. Three of these merit discussion here.

Firstly, we have assumed that the risk status of each individual woman remained unchanged from the time of her first screen. In practice, it would certainly change: for example, her menstrual status would change, and the information collected at screening would be replaced with new information once she had been screened again. Thus after each screen a woman could be re-allocated to a different risk-category. This would be done automatically by computer.

Secondly, we have taken observed risk of breast cancer diagnosis over a five-year period following a negative screen and used this to estimate underlying incidence rates for subgroups of the population. This assumes that the parameters of the screening programme (false negative rate and mean sojourn time) are independent of the risk factors; this might not be true since, for example, some of the factors could possibly influence tumour growth rate and hence sojourn time. A simulation study<sup>15</sup> has used the same predictor but with otherwise different assumptions and confirmed these results over the first few years of a screening programme.

Thirdly, we are of course assuming a future programme in which the performance of screening is similar to that in the present Edinburgh programme. It may well be that both technology and skills will improve in future.

Whenever a mathematical formula is used to allocate public resources there is concern about "fairness". It has been shown theoretically<sup>16 17</sup> that the allocation which minimises the overall risk of the undesirable event at the same time ensures that each member of the population then has the same chance of experiencing it irrespective of their initial risk status; thus the proposed strategies are "fair" in the sense that they equalise the probabilities of a symptomatic breast cancer diagnosis.

A more fundamental objection lies in our choice of end-point which was detection by screening; ideally we should optimise years of life saved. The aggression displayed by breast tumours covers a wide spectrum, and screening is probably of little benefit to those at either end of the spectrum. Our results extend to benefits for years of life only if this spectrum of aggression is similar in each of the three risk groups. We have, however, searched the literature and can find no definite evidence linking the risk factors with prognosis or with tumour growth rate. We hope that future research will examine this question.

We must finally emphasise that it is unknown whether a policy of allocating different recall times in

mass screening would be acceptable in practice. Anxiety could be generated in those women recalled early and feelings of neglect in those left for the longest time. The staff might find the complexity and the women's queries unacceptable. Research which could elucidate this would be necessary before any proposal is made to adopt a risk strategy as part of public policy.

The Edinburgh Breast Screening Project is funded by the Scottish Home and Health Department and the Cancer Research Campaign. We are very grateful to the members of the Project Committee for their consistent help and encouragement. We particularly wish to thank the other members of the Project Team for their assistance: Professor Sir A P M Forrest, Dr T J Anderson, Mr U Chetty, Dr A E Kirkpatrick, Dr J Lamb and Mr W Lutz.

Address for correspondence: Dr Freda Alexander, Leukaemia Research Fund Centre for Clinical Epidemiology, The University of Leeds, 17 Springfield Mount, Leeds LS2 9NG

## References

- Roberts MM, Alexander FE, *et al.* The Edinburgh randomised trial of screening for breast cancer: Description of method. *Br J Cancer* 1984; **50**: 1-6.
- Shapiro S, Venet W, *et al.* Ten-to-fourteen-year effect of screening on breast cancer mortality. *JNCI* 1982; **69**: 349-55.
- Tabar L, Gad A, *et al.* Reduction in mortality from breast cancer after mass screening with mammography. *The Lancet*; 1985; 829-32.
- Breast Cancer Screening: report to the Health Ministers of England, Wales, Scotland and Northern Ireland by a working group chaired by Sir Patrick Forrest. HMSO, 1987.
- UK Breast Cancer Detection Working Group. Trial of early detection of breast cancer: Description of method. *Br J Cancer* 1981; **44**: 618.
- Alexander FE, Roberts MM, Huggins A. Risk factors for breast cancer with applications to selection for the prevalence screen. *J Epidemiol Community Health* 1987; **41**: 101-6.
- Copas JB: Regression, prediction and shrinkage. *J Roy Statist Soc, Series B*, 1983; **45**: 311-54.
- Alexander FE. Parameter estimation for screening programmes which use two modalities: a modification of the method of Day and Walter. Submitted to *Statistics in Medicine*.
- Day NE, Walter SD. Simplified models of screening for chronic disease: estimation procedures from mass screening programmes. *Biometrics* 1984; **40**: 1-14.
- Shapiro S, Goldber J, *et al.* Risk factors for breast cancer, a prospective study in host environment interactions in the aetiology of cancer in man. Proc. meeting in Princeton (1972), eds. Doll and Vodjoja, Lyons (1973) 169-182.
- De Waard F, Collette HJA, Rombach JJ, *et al.* DOM project for the early detection of breast cancer, Utrecht, The Netherlands *J Chron Dis* 1984; **37**: 1-44.
- Duffy SW, Roberts MM, *et al.* Risk factors for breast cancer: relevance to screening. *J Epidemiol Community Health*, 1983; **37**: 127-31.
- Schechter MT, *et al.* Selection of women at high risk of breast cancer for initial screening, *J Chron Dis* 1986; **39**: 253-60.
- Alexander FE, Roberts MM. The use of risk factors to allocate schedules for breast cancer screening: the need to identify a truly low risk group. (In preparation).
- Alexander FE. The use of risk factors to allocation schedules for breast cancer screening: II-A simulation study. (In preparation).
- Alberman ED, Goldstein H. The 'at risk' register: A statistical evaluation, *Br J Prev Med*, 1970; **24**: 129-35.
- Carpenter RG. Scoring to provide risk-related primary health care: Evaluation and up-dating during use. *J Roy Statist Soc* 1983; Ser A, 146.

Accepted for publication January 1988

## Appendix

### 1. Assigning the underlying incidence rate for the risk groups

The logistic regression model is

$$\text{logit } p = \alpha + \beta^T x$$

where  $x$  is a vector denoting risk factor values and  $p$  is the probability of being diseased.

The preshrunk predictor is then

$$\text{logit } p = \alpha + \beta^T X_0 + k\beta^T (x - x_0)$$

where  $k$  is the shrinkage,  $x_0$  is the vector of mean values, and  $p$  now denotes the probability of future disease. More precisely: since the initial data were for five years' follow-up from the first screen,  $p$  is the probability of disease within the next five years. These divided by 5 yield average annual predicted incidence rates over the next five years and are given for each risk category in table 5.

The screening model in ref. 8 requires knowledge of the underlying incidence rate,  $J$ , in the population. Then in a population in regular screening the incidence rates over five years are  $SJ$  where  $S$  is a function of the screening parameters. If the screening parameters are the same for the three risk groups, it follows that the underlying incidence rates are in the same ratio as the five-year rates. Hence the underlying rates can be obtained from those in table 5 by multiplying by a constant  $b$ , to give underlying rates  $J_H, J_M, J_L$  in the three risk groups satisfying

$$J_H \times 0.2 + J_M \times 0.6 + J_L \times 0.2 = J$$

where  $J = 1.86$  is the underlying rate in the whole population.

It must be emphasised that this involves the assumption that the screening parameters are independent of the risk factors.

2. Estimating the performance of screening

For a population with underlying incidence rates  $J/c \times 1000$  women-years and screened by mammography every  $2c$  years with false negative rate  $\beta$  and mean sojourn time  $m \times c$  years the rates of screen detected and symptomatic cancers over three complete time periods of  $2c$  years are given by the model described in ref. 8.

To estimate the performance of screening when mammography is given every three years we take  $c = 1.5$ ,  $m = 4.98/1.5$  and  $J = 1.86 \times 1.5$ .

To estimate the performance of the risk strategy the three risk groups were considered separately. For the high-risk group  $c = 0.75$ ,  $J = 3.52 \times 0.75$ , for the medium-risk group  $c = 1.5$  and  $J = 1.64 \times 1.5$ , and for the low-risk group  $c = 3$  and  $J = 0.87 \times 3$ . For each risk category this provided rates over three complete cycles. It was assumed that the steady state was represented by cycle 3 for each group. Appropriate rates were then combined in the proportions 0.2:0.6:0.2. This is illustrated in figure 2.

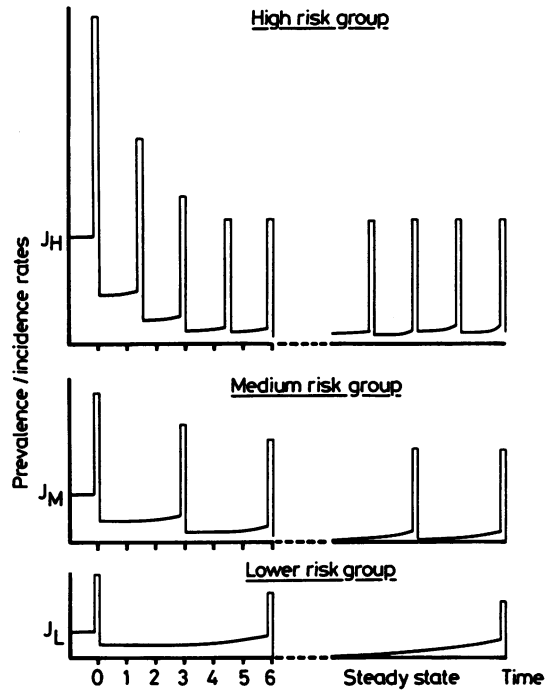


Fig 2 Risk strategy—prevalence and incidence rates.