 International Journal of Epidemiology 1999;**28**:386–390

Interval cancers in a randomized controlled trial of screening for colorectal cancer using a faecal occult blood test

SM Moss,[a] JD Hardcastle,[b] DA Coleman,[a] MHE Robinson[c] and VC Rodrigues[a]

Background	The sensitivity of unhydrated Haemoccult II has been examined in the context of a randomized controlled trial of faecal occult blood screening for colorectal cancer in Nottingham, UK.
Method	Both traditional and proportional incidence methods were used to calculate sensitivity separately for both sexes, for two age groups at entry to the trial, for first screen and repeat screens and for three subsites within the large bowel.
Results	The traditional method of estimation yielded a sensitivity of 59% whereas the corresponding figure obtained using the proportional incidence method was 54%. The difference between the estimates using the two methods was greatest in subjects aged ⩾65 at entry to the trial and in cancers of the distal colon.
Conclusions	The results suggest that there may be a higher proportion of slower growing tumours in subjects aged ⩾65 and that cancers occurring in the distal colon may have a longer mean sojourn time than cancers proximal to the sigmoid colon.

Keywords	Colorectal cancer, screening, sensitivity, proportional incidence
Accepted	20 November 1998

The success of any screening programme depends in part on the ability of the test to pick up progressive cases of disease before they become symptomatic and while they are still curable. Cases of disease presenting after a negative screen will arise either from cases of pre-symptomatic disease present at the time of screening which the screening test failed to detect, or from cases arising *de novo* and progressing to the symptomatic stage within the interval between routine repeat screens. In cancer screening programmes the occurrence of interval cancers among people who had previously been screened negative provides a measure of the sensitivity of the test, and of the interval within which screening needs to be routinely repeated. A traditional method of estimating sensitivity is to calculate the ratio of screen-detected cancers to the total number of cancers diagnosed among acceptors of screening during periodic screening.[1] A second method is to compare the incidence rate of interval cancers in successive periods after a negative screen with the expected incidence in the absence of screening. The difference between these two rates gives the number and proportion of cancers whose diagnosis has been advanced by screening; this is known as the proportional incidence method.[2] The expected cancer incidence in the absence of screening can most reliably be obtained from the control group incidence in a randomized controlled trial of screening.

Sensitivity may be calculated for a single screening test (test or screen sensitivity) or for a programme of screening (programme sensitivity). The latter has been defined as 'the probability that a case in the detectable preclinical stage at any time during an ongoing screening program (and ending with the last screen) will test positive in at least one of the screens'.[3] The traditional method of calculation can estimate programme sensitivity by the inclusion of all screen-detected and interval cancers, or test sensitivity by including only cases detected at or occurring in the interval after one screening round. However, since some screen-detected lesions might not otherwise have proceeded to clinical disease within the given interval, test sensitivity may be overestimated by this method.

The proportional incidence method provides an estimate of test sensitivity, but, like the traditional method, assumes that all interval cancers were missed by screening. The longer the period during which interval cancers are counted, the greater will be the proportion which are 'newly arising' as opposed to those missed by screening. Fitting various mathematical models to the data[3–6] can take this into account.

A randomized controlled trial of faecal occult blood screening for colorectal cancer, which has been conducted in Nottingham,[7] provides an opportunity to examine both test and programme sensitivity which will help inform decisions on screening interval.

[a] Cancer Screening Evaluation Unit, Block D, Institute of Cancer Research, Cotswold Road, Sutton, Surrey SM2 5NG, UK.

[b] Section of Surgery, Floor E, West Block, Queen's Medical Centre, Nottingham NG7 2UH, UK.

[c] Department of Surgery, City Hospital, Nottingham NG5 1PB, UK.

Method

Between February 1981 and January 1991, 152 850 men and women aged 45–74 were recruited into the trial and randomly allocated to a study group who were invited to be screened or to a control group who were not invited. The screening test was a Haemoccult II faecal occult blood test in which subjects collected stool samples on a test card and mailed them to a screening centre where they were processed without rehydration. Subjects with positive test results were fully investigated, in the great majority of cases by colonoscopy.[7] After exclusion of study group subjects in whom cancer or adenoma was found, the remainder were routinely re-invited to carry out a repeat test every 2 years. Of the study group, 53.4% accepted the first test and 6.1% of individuals who had not previously responded accepted a re-invitation for screening. Thus, of those invited for screening, almost 60% completed at least one test. Any cancers diagnosed within 2 years of a negative screen or after a positive screen where further investigation was negative or was refused have been classified as interval cancers.

Using the proportional incidence method, test sensitivity can be estimated using the formula:

$$s = \frac{1 - (I_T/I)}{1 - \int_0^T F(t)dt}$$

where I_T = observed interval cancers in time T, I is the expected incidence in the absence of screening and F(t) is the cumulative distribution function of f(t), the distribution of the preclinical detectable phase (PCDP).[2]

Since F(t) is unknown, the approximate formula: $s = 1 - (I_T/I)$ is used here. Clearly the validity of this approximation will depend on the value of $\int_0^T F(t)dt$, which is determined by the length of the interval compared with the mean sojourn time of the PCDP. This is addressed further in the Discussion.

For proportional incidence estimates, the expected incidence of colorectal cancer in the absence of screening has been calculated from the overall control group incidence. However, as test group subjects who did not take up the offer of screening had a higher incidence of colorectal cancer than the control group, the expected incidence in screened subjects was derived after adjusting for this bias. The adjustment was carried out separately for each sex, but not for individual age groups or subsites because of variability due to small numbers. This was done using the formula:

$$r_a = [r_c - (1 - p)r_n]/p$$

where r_a is the incidence rate in acceptors of screening, r_c is the control group incidence, r_n is the incidence in non-acceptors and p is the proportion of the intervention group who accepted screening.

Calculations of test sensitivity using the proportional incidence method have been made separately for males and females, for two age groups at entry to the trial: 45–64 years and ⩾65 years, for first screens and routine repeat screens, and for three subsites within the large bowel: rectum and rectosigmoid, sigmoid and descending colon, and transverse and ascending colon and caecum, including both flexures.

Table 1 Incidence rates (per 1000 person-years) of colorectal cancer in control group, in follow-up period[a] and adjusted expected rates for acceptors of screening

	45–64 years		⩾65 years		All ages	
	Male	Female	Male	Female	Male	Female
Control group incidence	1.31	0.92	2.94	1.96	1.74	1.26
Adjusted rate for acceptors	1.07	1.10	2.69	1.65	1.50	1.30

[a] Follow-up period is from trial entry to 31 December 1996.

The 'traditional' method estimates sensitivity according to the formula:

Sensitivity (%) = screen detected cancers/
(screen detected + interval cancers)

where interval cancers are those clinically detected within a short period following a negative screen, 2 years in this instance.

The traditional method has been used to calculate programme sensitivity according to sex, age at entry and subsite, and also to give an estimate of test sensitivity for different screening rounds and average test sensitivity.

For every accepted screen, person-years are calculated for the period of observation until the next accepted screen, censoring at 2 years if this is earlier.

The number of interval cancers used in calculating sensitivity by the traditional method includes only those following screening carried out before 1 January 1995 (allowing two completed years of follow-up since last screen), whereas for the proportional incidence method all interval cancers diagnosed up to 31 December 1996 are included, regardless of date of last screen.

Results

Proportional incidence method

The age- and sex-specific incidence rates of colorectal cancer in the control group, and the adjusted expected underlying rates in acceptors of screening are shown in Table 1. The incidence rates of colorectal cancer in the control group (all ages) were 1.74 and 1.26 per 1000 person-years for males and females, respectively. The corresponding figures for the adjusted underlying rates in screening acceptors were 1.50 and 1.30 per 1000 person-years. Apart from women in the younger age group, rates in acceptors are lower than in the control group, indicating that except for younger women, people who are at higher risk of colorectal cancer are less likely to accept screening. The difference in younger women reflects the higher acceptance rate of screening in this group as compared to other age and sex groups in the Nottingham trial. In the following Tables the adjusted incidence rates for acceptors of screening are applied to the person-years of observation following screening to calculate the number of cancers expected within the 2-year period between repeated screens.

Table 2 shows the numbers of cancers detected at each screening round, with a total of 139 cancers detected in males and 97 in females.

Up to 31 December 1996, there were 255 438 person-years of observation of screened subjects within 2 years following

Table 2 Detection of colorectal cancers by screening examination and sex of screened individuals

Screening examination	No. screened		Cancers detected	
	Male	Female	Male	Female
1	20 997	23 840	63	41
2	16 264	19 145	42	23
3	12 704	15 475	14	17
4	7808	9678	13	13
5	3113	3953	7	3

Table 3 Test sensitivity of screening by sex and person-years of observation within the 2-year period following each screening examination

	Person-years	Observed interval cases	Expected cases	% detected by screening
Males				
Year 1	59 675	39	89.5	56.4
Year 2	56 326	58	84.5	31.3
Total		97	174.0	44.3
Females				
Year 1	71 340	34	92.7	63.3
Year 2	68 097	33	88.5	62.7
Total		67	181.2	63.0

their first screen (Table 3); 164 interval cancers were diagnosed, 73 within 12 months of screening and 91 between 12 and 24 months after screening. The proportion of expected cancers whose diagnosis had been advanced by screening was 54% over the 2-year period, 60% of those expected in the first 12 months and 47% of those expected in the second 12-month period.

When the sexes were analysed separately (Table 3), the proportion of cancers whose diagnosis was advanced by screening was significantly higher in women than in men (χ^2 = 6.4, P = 0.012). This difference mainly emerged in the second year after screening, with only 31.3% of the expected cancers among men detected by screening in this period.

Separating the first screen from subsequent routine repeat screens it was found that the proportion of expected cases whose diagnosis had been advanced was higher following the first screen (62.7%) than following re-screens (50.0%) (Table 4), however, this difference did not attain statistical significance (χ^2 = 2.58, P = 0.108).

Table 5 shows that the proportion of cancers whose diagnosis was advanced by screening was slightly higher among people aged under 65 at entry to the trial, than in those over 65 among whom it fell off markedly in the second year after screening. However, the difference between age groups was not statistically significant (χ^2 = 0.05, P = 0.828).

Table 6 shows that screening apparently advanced the diagnosis of cancer arising in the transverse and ascending colon (48.7%) less often compared with the rectum and rectosigmoid (59.3%) and the sigmoid and descending colon (55.8%). Although these results were consistent for both sexes (data not shown), sensitivity was considerably higher among females for all three subsites.

Table 4 Test sensitivity of first and subsequent screens by person-years of observation within the 2-year period following each screening examination

	Person-years	Observed interval cases	Expected cases	% detected by screening
First screen				
Year 1	44 079	21	61.9	66.2
Year 2	41 948	24	58.9	59.3
Total		45	120.8	62.7
Rescreen				
Year 1	86 936	52	122.1	57.4
Year 2	82 475	67	115.9	42.2
Total		119	238.0	50.0

Table 5 Test sensitivity of screening by age at entry to the study and person-years of observation within the 2-year period following each screening examination

Age at entry	Person-years	Observed interval cases	Expected cases	% detected by screening
45–64 years				
Year 1	93 531	43	104.1	58.7
Year 2	89 288	49	99.4	50.7
Total		92	203.5	54.8
⩾65 years				
Year 1	37 484	30	78.5	61.8
Year 2	35 135	42	73.5	42.8
Total		72	152.0	52.6

Table 6 Test sensitivity of screening by subsite of cancer and person-years of observation within the 2-year period following each screening examination

Subsite[a]	Observed interval cases	Expected cases	% detected by screening
Rectum and rectosigmoid			
Year 1	27	70.6	61.8
Year 2	29	66.9	56.7
Total	56	137.5	59.3
Sigmoid and descending colon			
Year 1	18	53.4	66.3
Year 2	28	50.7	44.8
Total	46	104.1	55.8
Transverse and ascending colon			
Year 1	26	66.3	60.8
Year 2	31	44.8	30.8
Total	57	55.8	48.7

[a] Subsite not coded for five interval cancers.

Traditional method

Among people screened up to 31 December 1994, 236 cancers had been found at screening and 164 interval cancers had been diagnosed in those with a negative screen within 2 years. This gives a programme sensitivity of 236/400 or 59.0% (95%

Table 7 Programme sensitivity of screening by sex, age at entry to study, type of screen and subsite of cancer using the traditional method of estimation

	Screen-detected cancers	Interval cancers	Sensitivity %
Sex			
Males	139	97	58.9
Females	97	67	59.1
Age at entry			
45–64 years	107	92	53.8
⩾65 years	129	72	64.2
Subsite[a]			
Rectum and rectosigmoid	78	56	58.2
Sigmoid and descending	103	46	69.1
Transverse and ascending	55	57	49.1
Total	236	164	59.0

[a] Subsite not coded for five interval cancers.

Table 8 Estimate of test sensitivity by traditional method

Screening round	Screen-detected cancers	Interval cancers	Sensitivity %
1	104	45	69.7
2	65	48	57.5
3	31	34	47.7
4	26	28	48.1
5	10	11	47.6

CI : 54.2–63.8). If only the 73 interval cancers occurring within 12 months of a negative screen are included, the estimated programme sensitivity is 76.4% (95% CI : 71.7–81.1).

Table 7 shows programme sensitivity estimates by sex, age at entry, and subsite. The difference between males and females observed in test sensitivity using the proportional incidence method is no longer apparent, but the estimated programme sensitivity is higher (64.2%) in those aged ⩾65 at entry than in those aged under 65 (53.8%) (χ^2 =4.48, P = 0.034). In the analysis by subsite it is the programme sensitivity for the sigmoid and descending colon (69.1%) which is notably higher than the estimate above.

Table 8 gives the test sensitivity for individual screening rounds estimated by the traditional method. The crude average test sensitivity is 54.1%, whilst the average weighted by the number of screen-detected cases, is 60.1% at each round.

Discussion

While opinions on what is an 'acceptable' level of sensitivity for a screening programme may vary, the findings in this trial suggest that biennial screening using unhydrated Haemoccult will only detect around half of the progressive cancers in an unselected population. In a randomized controlled trial in Denmark sensitivity of unhydrated Haemoccult II was estimated to be 43.8% using the traditional method[8] whereas calculations using a mathematical model yielded a test sensitivity of 62.1%.[6] Experience with large numbers of interval cancers in American and Swedish faecal occult blood studies led the investigators to alter their original protocol by rehydrating the stool samples before testing. Programme sensitivity, including cases occurring within 12 months of a negative screen as interval cases, increased from 80.8% to 92.2% with rehydration in Minnesota.[9] A recent paper has estimated sensitivity in the annual screening arm of the Minnesota trial using several methods.[4] The estimate of the average test sensitivity obtained by the traditional method is 89.6%. An adjustment which aims to include only those cancers detectable at screening, and also to adjust for those cases not clinically diagnosed within 12 months, gives an estimated test sensitivity of 74.8% (95% CI : 65.2–84.4%), whilst the estimate based on a model which incorporates the chance effect of colonoscopy is about 95%. In Gothenburg, sensitivity following the first screen, counting interval cancers occurring in the mean period of 20 months before the second screen, estimated by the traditional method was 22% in the unhydrated group and 86% in the rehydrated group, although the numbers in each group are small.[10] However, this level of sensitivity could only be achieved with a substantial fall in specificity with many false positive results. As many as 38% of screened subjects in the Minnesota trial underwent at least one colonoscopy during the trial compared with 4% in Nottingham.

In another recent paper based on the first round of screening in Calvados, France, sensitivity was estimated by the traditional method, and also jointly with mean sojourn time using maximum likelihood and log-linear modelling techniques.[5] The traditional method, with cancers occurring within 12 months defined as interval cancers, gives an estimated sensitivity for all sites of 75%, whilst the modelling approaches give estimates of 47–48%. Data are also presented on the proportional incidence of interval cancers, from which an estimate of sensitivity, over a 2-year period of 34% can be derived. This is considerably lower than the results obtained here; one possible explanation is that the participation rate in Calvados was 43% overall, and no adjustment has been made for selection bias, possibly leading to a difference in underlying incidence between acceptors and controls. However, the estimated sensitivity of 76.4% obtained in the present study using the traditional method and a 12-month interval is similar to that observed in Calvados.

The validity of the estimate of sensitivity using the proportional incidence method will depend on the magnitude of $\int_0^T F(t)dt$, which is assumed to be 0 here. For example, if f(t) was exponentially distributed with mean 5 years, then $\int_0^T F(t)dt$ is approximately 0.35, meaning that the sensitivity of the test in detecting preclinical disease would be underestimated by 54% ($1/(1 - 0.35) \times 100$) using a 2-year interval. Using a one-year interval it is underestimated by around 10%. Nevertheless, comparisons between sensitivity calculated by the different methods will still be informative about the natural history of the disease.

The relatively low sensitivity and 2-year screening interval will be partly responsible for the lower reduction in mortality from colorectal cancer observed in Nottingham (15%) and Fünen (18%), compared with the Minnesota trial,[11] which found a significant 33% reduction in colorectal cancer mortality in subjects randomized to annual screening but not in those screened biennially, when compared to an unscreened control group. Lack of compliance with screening, which was not a problem in the Minnesota trial's volunteer population, was a further major constraint in both the Nottingham and the Danish trial.

The differences between test sensitivity measured by the traditional and proportional incidence methods are of relevance to understanding the natural history of colorectal cancer. The proportional incidence method measures the test's ability to advance the date of diagnosis of cancers which would otherwise arise in a defined interval—2 years in this case. It thus gives an indication of the effect of screening on cancer mortality in the relatively short-term, but ignores the effect of the detection and early treatment of cancers which would have later presented beyond this 2-year interval. The traditional method includes these as screen-detected cases, but it can be argued that they are likely to be slower-growing, less malignant variants, more likely to be curable even if left until symptomatic, and perhaps including some which might otherwise never have been diagnosed in the person's lifetime.

The higher estimated sensitivity for distal lesions obtained using the traditional method compared with the proportional incidence method therefore suggests a slower progression of such lesions. This is in agreement with the results of Launoy,[5] who found an estimated mean sojourn time of 6.4 years for distal lesions compared with 3.5 years for proximal cancers. This suggestion will be investigated further using more sophisticated, multivariate methods on the Nottingham data.

Another possible explanation for the discrepancy between the two methods according to subsite could arise from variability in definition of the subsites particularly between sigmoid and rectosigmoid. The control group cancers in this study (on which estimated expected numbers are based) were managed by many different surgical teams without standardization of criteria for classifying subsite, whereas the screen-detected cancers and a majority of interval cancers were managed by a single team, which defined subsite by measured distance from the anal margin.

For people over 65 when first invited to screening, the relatively high programme sensitivity (64%) contrasts with the low proportional incidence rate (53%), and with the lower observed mortality reduction (10% compared with 19% in those under 65 at entry). This suggests the possibility that proportionally more 'unimportant' malignant lesions are being detected and treated in this age group. The significantly higher sensitivity in females by the proportional incidence method is not observed in the estimate using the traditional method, or reflected in the mortality reduction, where the slightly higher reduction in females is probably explained by a higher uptake of screening.

In conclusion, the sensitivity of unhydrated Haemoccult as used in the Nottingham trial is estimated as 59% using the traditional method and 54% using the proportional incidence method. The difference between estimates using the two methods is greatest for those aged ⩾65 at entry to the trial, and in cancers of the distal colon, suggesting that there may be a higher proportion of slower growing tumours in these groups.

Acknowledgements

This work was a joint collaboration between the Section of Surgery, Queen's Medical Centre, Nottingham and the Cancer Screening Evaluation Unit, Institute of Cancer Research, Sutton, Surrey. The latter receives support from the Department of Health; the views expressed in this publication are those of the authors and not necessarily those of the Department of Health. We would like to thank Professor JO Chamberlain, Mrs C Mangham and Ms TJ Mapp for their contribution to this work.

References

[1] Chamberlain J, Clifford RD, Nathan BE, Price JL, Burn I. Error rates in screening for breast cancer. *Clin Oncol* 1979;**5:**135–46.

[2] Day NE. Estimating the sensitivity of a screening test. *J Epidemiol Community Health* 1985;**39:**364–66.

[3] Church TR, Ederer F, Mandel JS. Fecal occult blood screening in the Minnesota study: sensitivity of the screening test. *J Natl Cancer Inst* 1997;**89:**1440–48.

[4] Day NE, Walter SD. Simplified models of screening for chronic disease: estimation procedures from mass screening programmes. *Biometrics* 1984;**40:**1–14.

[5] Launoy G, Smith TC, Duffy SW, Bouvier V. Colorectal cancer mass-screening: estimation of faecal occult blood test sensitivity, taking into account cancer mean sojourn time. *Int J Cancer* 1997;**73:**220–24.

[6] Gyrd-Hansen D, Sogaard J, Kronborg O. Analysis of screening data: colorectal cancer. *Int J Epidemiol* 1997;**26:**1172–81.

[7] Robinson MHE, Moss SM, Hardcastle JD, Whynes DK, Chamberlain JO, Mangham CM. Effect of retesting with dietary restriction in Haemoccult screening for colorectal cancer. *J Med Screening* 1995; **2:**41–44.

[8] Kronborg O, Fenger C, Olsen J, Bech K, Sondergaard O. Repeated screening for colorectal cancer with faecal occult blood test. *Scand J Gastroenterol* 1989;**24:**599–606.

[9] Mandel JS. The University of Minnesota's colon cancer control study: design and progress to date. In: Chamberlain J, Miller AB (eds). *Screening for Gastrointestinal Cancer*. Stuttgart: Hans Huber Publishers, 1988, pp.17–24.

[10] Kewenter J, Bjork S, Haglind E, Smith L, Svanvik J, Ahren C. Screening and rescreening for colorectal cancer. A controlled trial of fecal occult blood testing in 27 700 subjects. *Cancer* 1988;**62:**645–51.

[11] Mandel JS, Bond JH, Church TR *et al.* For the Minnesota Colon Cancer Control Study. Reducing mortality from colorectal cancer by screening for fecal occult blood. *N Engl J Med* 1993;**328:**1365–71.