Correspondence

Re: Participation in Colorectal Cancer Screening: a Review

Vernon (1) has reviewed the published literature on adherence to screening for colorectal cancer.

The present correspondence relates to our experience with colorectal cancer screening, which was not mentioned in Vernon’s review. In a rural area of the Province of Florence, Italy, a population-based screening by fecal occult blood testing (FOBT) has been carried out since 1982. All residents of that area who are aged 40–70 years (about 70,000 inhabitants) have been invited every other year to undergo a screening protocol (2).

The non-rehydrated guaiac test (Hemoccult II, SmithKline Diagnostics, Palo Alto, CA) was used until 1992, after which the rehydrated test (Hemoccult II) was used until 1995 and, more recently, immunologic tests (Hemeselect, SmithKline Diagnostics; Immudia Hem SP, Fujirebio Inc. Tokyo, Japan) have been adopted.

In the framework of a case–control study on the effectiveness of FOBT (3), a questionnaire was mailed to 1030 control subjects who were randomly selected from municipality residence archives. Control subjects were matched to the case subjects by sex, age, and place of residence. In the questionnaire, detailed information was requested about marital status, educational level, current and usual occupation, cigarette smoking habits, and number of first-degree relatives dead from any cause and from colorectal cancer. Furthermore, detailed information was requested about body mass index, dietary habits, and previous enterologic examinations (FOBT, x ray, and endoscopy) performed outside the screening program. Among the control subjects, 378 (36.7%) had had at least one screening during their lifetime and 301 (29.2%) had screening within the past 3 years. Answers to the questionnaire indicated no differences among participants in the screening and nonparticipants with regard to the following factors: age, place of birth, educational level, body mass index, food and wine consumption per day, cigarette smoking habits, and number of relatives dead from any cause. Participants were more likely to be females (in the first round, FOBT screening was combined with mammographic screening); consequently, the percentage of housewives was greater among participants. Furthermore, compared with nonparticipants, participants were slightly more likely to report some diagnostic intestinal examinations performed outside the program (12.4% for participants versus 9.2% for nonparticipants; P = .18). The major difference among participants and nonparticipants was that a larger proportion of participants had at least one relative dead from colorectal cancer (10.6% versus 3.7%; P < .01). Thus, as Vernon (1) remarked, adherence was highest in relatives of colorectal cancer case subjects, and no other statistically significant determinants of compliance emerged from our survey.

In conclusion, in one of the few experiences of colorectal cancer screening in southern Europe, the goal of Healthy People 2000 (4), which is that at least 50% of the population will have had FOBT within the past 2 years, is far from reaching its mark. The use of a 1-day immunologic test in our program may increase the compliance (to approximately 40% (5) because there are no dietary restrictions imposed by the protocol, and a shorter period is required for collection of feces.

Marco Zappa
Guido Castiglione
Daniela Giorgi
Grazia Grazzini
Eugenio Paci
Stefano Ciatto

References


Notes

Affiliations of authors: M. Zappa, G. Castiglione, G. Grazzini, E. Paci, S. Ciatto, Centro per lo Studio e la Prevenzione Oncologica, Azienda Ospedaliera di Careggi, Florence, Italy; D. Giorgi, Istituto Scientifico Tumori Genova Satellite Unit, e/o Centro per lo Studio e la Prevenzione Oncologica, Florence, Italy.

Correspondence to: Marco Zappa, M.D., Unità Operativa Epidemiologia, Centro per lo Studio e la Prevenzione Oncologica, via di San Salvi 12, 50135 Florence, Italy.

Re: Fecal Occult Blood Screening in the Minnesota Study

In a recent issue of the Journal, Church et al. (1) reanalyzed data from the Minnesota study to estimate the sensitivity of the screening test (fecal occult blood test composed of six slides) and the sensitivity of the screening program. Developed approaches agreed with the crude estimate of test and program sensitivities of about 90%. We agree with Church et al. with regard to the great importance of distinguishing screening test and screening program sensitivity. However, we believe that the study by Church et al. has some limitations with regard to the methodologies proposed and the public health implications.

Methodological limitations. As the investigators (1) suggested, the proportion of cancers detected by a particular test among all cancers diagnosed in complying people within 1 year of their screen in the Minnesota study provides a proper estimate for program sensitivity but not for screening test sensitivity, since it would suppose that all cancers existing at a screen emerge within 1 year after the screen and that all cancers emerging 1 year after the screen existed
at the screen. To distinguish cancers entering the detectable preclinical phase after the screen from those existing at screen, Church et al. (1) suggested an adjustment for the incidence of cancers occurring after a false-positive test, assuming that all prevalent cases among positive tests have been detected. This method would be completely relevant if all positive case patients had undergone a complete colonoscopy. In fact, only 827 (83.9%) of 986 positive case patients underwent a complete colonoscopy so that some prevalent cases remained undiscovered among the positive case patients (2). Therefore, such an adjustment underestimates the number of false-negatives and consequently overestimates the test sensitivity.

Church et al. (1) provide a relationship between the sensitivity of the program ($S_p$) and the sensitivity of a screen ($S$) as follows:

$$S_p = 1 - \prod_{i=1}^{k} [1 - S(t_i)]$$

for a cancer with a preclinical duration of $k$ screens, with $S(t_i)$ being the screen sensitivity for the time $t_i$ in the disease progression. As Church et al. indicate, this formula is relevant only if the cancers having a preclinical duration longer than the interval between two screens represent the vast majority of cases. Thus, such a formula can be applied in a very limited number of situations like those of the Minnesota study where the interval between two screens is 1 year and the preclinical duration of the cancer is assumed to be of about 5 years (1).

**Limitations in public health implications.** We believe that the major limitation of this study is not its validity but its interest for public health. Even if the re-application of the Lang–Ransohoff method with the real data from the Minnesota study, published in the same issue of the Journal (3), showed that the role of chance detection in the reduction of mortality from colorectal cancer in the Minnesota study is smaller than that previously estimated by Lang and Ransohoff, the cost of the numerous colonoscopies generated by a program with a positive rate of about 10% and annual screening is prohibitive for the majority of countries. From a public health point of view, the interest in the assessment of test and program sensitivity derived from the Minnesota study is very limited for all societies that cannot support the cost generated by such a program.

We have recently estimated the sensitivity of the fecal occult blood test and the duration of the preclinical detectable phase of the cancer reflected by the mean sojourn time (MST), using a Bayesian technique of Gibbs sampling (4) with data from a French ongoing mass screening program where the positivity rate was 2.8%, very close to conditions of randomized European trials (5–7). Since MST was estimated to be much shorter than 5 years for certain subsites and the interval between two screens was longer than 1 year in European trials, the formula proposed by Church et al. (1) could not be used in estimating the program sensitivity. We suggest below another method for estimating the sensitivity of the program:

Under the assumption of an exponential distribution of sojourn time with parameter $\lambda$ (with $S_i$ being the sensitivity of the test and $r$ being the interval between two screens), the probability $P$ that a cancer is detected by a given test can be expressed by

$$P = S_i (1 - \exp(-\lambda r))/\lambda r.$$

With further mathematical developments, we can express the sensitivity of the program as follows:

$$S_p = S_i [1 - \exp(-\lambda r)])/\lambda r] [1/(1 - (\exp(-\lambda r)(1 - S_j))).]$$

Applying this formula, Table 1 shows the values of program sensitivity according to subsite for different assumptions of test sensitivity and interval between two screens.

We believe that these results provide useful additional information to the estimates derived from the Minnesota study for countries involved in mass screening for colorectal cancer strategies that differ from those of the Minnesota study.

**GUY LAUNOY
STEPHEN W. DUFFY
TERESA C. PREVOST
VÉRONIQUE BOUVIER**

**References**


**Table 1.** Sensitivity of colorectal cancer mass screening program according to cancer subsite under different hypothesis for test sensitivity ($S_i$) and interval between screening ($r$) values

<table>
<thead>
<tr>
<th>Subsite</th>
<th>$r = 1$</th>
<th>$r = 2$</th>
<th>$r = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S_i = 40%$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal colon</td>
<td>0.62</td>
<td>0.46</td>
<td>0.36</td>
</tr>
<tr>
<td>Distal colon</td>
<td>0.76</td>
<td>0.60</td>
<td>0.50</td>
</tr>
<tr>
<td>Rectum</td>
<td>0.57</td>
<td>0.39</td>
<td>0.30</td>
</tr>
<tr>
<td>All subsites</td>
<td>0.70</td>
<td>0.54</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>$S_i = 50%$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal colon</td>
<td>0.69</td>
<td>0.53</td>
<td>0.43</td>
</tr>
<tr>
<td>Distal colon</td>
<td>0.82</td>
<td>0.66</td>
<td>0.61</td>
</tr>
<tr>
<td>Rectum</td>
<td>0.64</td>
<td>0.46</td>
<td>0.36</td>
</tr>
<tr>
<td>All subsites</td>
<td>0.76</td>
<td>0.60</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>$S_i = 60%$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal colon</td>
<td>0.74</td>
<td>0.59</td>
<td>0.48</td>
</tr>
<tr>
<td>Distal colon</td>
<td>0.86</td>
<td>0.80</td>
<td>0.75</td>
</tr>
<tr>
<td>Rectum</td>
<td>0.69</td>
<td>0.52</td>
<td>0.41</td>
</tr>
<tr>
<td>All subsites</td>
<td>0.80</td>
<td>0.66</td>
<td>0.56</td>
</tr>
</tbody>
</table>

*MST = mean sojourn time value as estimated in Launoy et al. (5).
Selection of positive subjects for a complete colon work-up may create bias, which could lead either to underestimation or overestimation of the rate of new cancers among test-positive subjects. Commonly, cooperation with study procedures is a marker for healthy outcomes, so those incompletely diagnosed subjects may be more prone to develop cancer. However, any selection effect would have to be large to create sizable bias, since the fraction of positive subjects with an incomplete work-up is small (17%) and the positive predictivity even smaller (2%).

Launoy et al. also imply that expression 1 (1) was used in our estimates of sensitivity and that it applied to all screening cases. However, the expression was intended only to illustrate that cumulative (i.e., program) sensitivity can be greater than the largest screen sensitivity; it did not imply that each tumor is exposed to multiple screens. As we state (1), “expression 1 can be used to calculate program sensitivity for a typical or plausible sensitivity function.” Indeed, some tumors may not be exposed to any screens if the detectable, preclinical phase duration for the tumor is shorter than the screening interval and the tumor initiation time is such that the phase does not intersect a screen. Furthermore, if screen sensitivity is high, few prevalent tumors will be available for detection at a later screen.

The question of cost-effectiveness of fecal occult blood screening for colorectal cancer is not settled. We are unaware of any thorough evaluation of the costs, taking into account both the outlays for screening itself versus usual care and the costs of care for screened versus unscreened cancers. A preliminary assessment (3) found the cost-effectiveness of fecal occult blood testing for colorectal cancer to compare favorably with estimates of cost-effectiveness for programs of breast cancer screening for women over the age of 50 years. Moreover, ongoing follow-up in the Minnesota study may eventually indicate that the trend seen in our earlier analysis (4) toward lower incidence in the screened groups is real. This could greatly affect the cost-effectiveness of screening. For example, a 20% decrease in incidence would lead to a decrease in the cost of treating 20% of the cancers, which could offset some (if not all) of the costs of screening. Until these values are better known, it is premature to decide that screening is not affordable.

Finally, the model that Launoy et al. offer is interesting and certainly plausible in situations in which the sensitivity is constant over the progression of the tumor. However, it is unlikely that the potential for bleeding remains constant over the natural history of a tumor; therefore, such a model may be less useful than models that allow screen sensitivity to vary and integrate over individual cases. Nonetheless, even with this caveat, we are gratified that their model reinforces the finding that program sensitivity can be considerably higher than screen sensitivity, especially when average preclinical duration is long compared with the screening interval.

Response

We thank Dr. Launoy and colleagues for their interest in and comments on our article (1). We would like to clarify aspects of the methodology that pertain to their criticisms. Our analysis (1) distinguishes average values from distributions of individual values yielding those averages. Not only do the preclinical durations vary among tumors, but also the sensitivity of the test changes as the tumor progresses. We have attempted estimation in face of these highly variable interactions in a way that is free of distributional assumptions. The validity of any models with additional constraints is questionable.

Specifically, Launoy et al. assume that, in the estimate of screen sensitivity adjusted for incidence among false-positives (as well as among all negatives), false-positives comprised all positive tests without diagnosis. This procedure would underestimate the incidence rate, deflate the denominator, and bias the sensitivity upward. However, we designated as false-positives only those positive subjects submitting either to a colonoscopy or to a flexible sigmoidoscopy and a barium enema x ray. Cancers in test-positive subjects not receiving a diagnosis were omitted from both the denominator and the numerator; hence, they could not bias the calculation in the way Launoy et al. describe.

However, as we note (1), the self-selection of positive subjects for a complete colon work-up may create bias, which could lead either to underestimation or overestimation of the rate of new cancers among test-positive subjects. Commonly, cooperation with study procedures is a marker for healthy outcomes, so those incompletely diagnosed subjects may be more prone to develop cancer. However, any selection effect would have to be large to create sizable bias, since the fraction of positive subjects with an incomplete work-up is small (17%) and the positive predictivity even smaller (2%).

Launoy et al. also imply that expression 1 (1) was used in our estimates of sensitivity and that it applied to all screening cases. However, the expression was intended only to illustrate that cumulative (i.e., program) sensitivity can be greater than the largest screen sensitivity; it did not imply that each tumor is exposed to multiple screens. As we state (1), “expression 1 can be used to calculate program sensitivity for a typical or plausible sensitivity function.” Indeed, some tumors may not be exposed to any screens if the detectable, preclinical phase duration for the tumor is shorter than the screening interval and the tumor initiation time is such that the phase does not intersect a screen. Furthermore, if screen sensitivity is high, few prevalent tumors will be available for detection at a later screen.

The question of cost-effectiveness of fecal occult blood screening for colorectal cancer is not settled. We are unaware of any thorough evaluation of the costs, taking into account both the outlays for screening itself versus usual care and the costs of care for screened versus unscreened cancers. A preliminary assessment (3) found the cost-effectiveness of fecal occult blood testing for colorectal cancer to compare favorably with estimates of cost-effectiveness for programs of breast cancer screening for women over the age of 50 years. Moreover, ongoing follow-up in the Minnesota study may eventually indicate that the trend seen in our earlier analysis (4) toward lower incidence in the screened groups is real. This could greatly affect the cost-effectiveness of screening. For example, a 20% decrease in incidence would lead to a decrease in the cost of treating 20% of the cancers, which could offset some (if not all) of the costs of screening. Until these values are better known, it is premature to decide that screening is not affordable.

Finally, the model that Launoy et al. offer is interesting and certainly plausible in situations in which the sensitivity is constant over the progression of the tumor. However, it is unlikely that the potential for bleeding remains constant over the natural history of a tumor; therefore, such a model may be less useful than models that allow screen sensitivity to vary and integrate over individual cases. Nonetheless, even with this caveat, we are gratified that their model reinforces the finding that program sensitivity can be considerably higher than screen sensitivity, especially when average preclinical duration is long compared with the screening interval.

Notes

Affiliations of authors: G. Launoy, V. Bouvier, Registe des Cancers Digestifs du Calvados (CIRI Institut National de la Sante et de la Recherche Medicales [INSERM] 9603), Caen, France; S. W. Duffy, T. C. Prevost, Medical Research Council Biostatistics Unit, Cambridge, U.K.

Correspondence to: Guy Launoy, M.D., Ph.D., CJF INSERM 9603, Faculte Medecine, CHU Caen 14933, France.

Response

We thank Dr. Launoy and colleagues for their interest in and comments on our article (1). We would like to clarify aspects of the methodology that pertain to their criticisms. Our analysis (1) distinguishes average values from distributions of individual values yielding those averages. Not only do the preclinical durations vary among tumors, but also the sensitivity of the test changes as the tumor progresses. We have attempted estimation in face of these highly variable interactions in a way that is free of distributional assumptions. The validity of any models with additional constraints is questionable.

Specifically, Launoy et al. assume that, in the estimate of screen sensitivity adjusted for incidence among false-positives (as well as among all negatives), false-positives comprised all positive tests without diagnosis. This procedure would underestimate the incidence rate, deflate the denominator, and bias the sensitivity upward. However, we designated as false-positives only those positive subjects submitting either to a colonoscopy or to a flexible sigmoidoscopy and a barium enema x ray. Cancers in test-positive subjects not receiving a diagnosis were omitted from both the denominator and the numerator; hence, they could not bias the calculation in the way Launoy et al. describe.

However, as we note (1), the self-selection of positive subjects for a complete colon work-up may create bias, which could lead either to underestimation or overestimation of the rate of new cancers among test-positive subjects. Commonly, cooperation with study procedures is a marker for healthy outcomes, so those incompletely diagnosed subjects may be more prone to develop cancer. However, any selection effect would have to be large to create sizable bias, since the fraction of positive subjects with an incomplete work-up is small (17%) and the positive predictivity even smaller (2%).

Launoy et al. also imply that expression 1 (1) was used in our estimates of sensitivity and that it applied to all screening cases. However, the expression was intended only to illustrate that cumulative (i.e., program) sensitivity can be greater than the largest screen sensitivity; it did not imply that each tumor is exposed to multiple screens. As we state (1), “expression 1 can be used to calculate program sensitivity for a typical or plausible sensitivity function.” Indeed, some tumors may not be exposed to any screens if the detectable, preclinical phase duration for the tumor is shorter than the screening interval and the tumor initiation time is such that the phase does not intersect a screen. Furthermore, if screen sensitivity is high, few prevalent tumors will be available for detection at a later screen.

The question of cost-effectiveness of fecal occult blood screening for colorectal cancer is not settled. We are unaware of any thorough evaluation of the costs, taking into account both the outlays for screening itself versus usual care and the costs of care for screened versus unscreened cancers. A preliminary assessment (3) found the cost-effectiveness of fecal occult blood testing for colorectal cancer to compare favorably with estimates of cost-effectiveness for programs of breast cancer screening for women over the age of 50 years. Moreover, ongoing follow-up in the Minnesota study may eventually indicate that the trend seen in our earlier analysis (4) toward lower incidence in the screened groups is real. This could greatly affect the cost-effectiveness of screening. For example, a 20% decrease in incidence would lead to a decrease in the cost of treating 20% of the cancers, which could offset some (if not all) of the costs of screening. Until these values are better known, it is premature to decide that screening is not affordable.

Finally, the model that Launoy et al. offer is interesting and certainly plausible in situations in which the sensitivity is constant over the progression of the tumor. However, it is unlikely that the potential for bleeding remains constant over the natural history of a tumor; therefore, such a model may be less useful than models that allow screen sensitivity to vary and integrate over individual cases. Nonetheless, even with this caveat, we are gratified that their model reinforces the finding that program sensitivity can be considerably higher than screen sensitivity, especially when average preclinical duration is long compared with the screening interval.

References

(2) Ederer F, Church TR, Mandel JS. Sample sizes for prevention trials have been too small. Am J Epidemiol 1993;137:787–96.

Notes

Affiliations of authors: T. R. Church, J. Mandel, Division of Environmental and Occupational Health, School of Public Health, University of Minnesota, Minneapolis; F. Ederer, Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, and The EMMES Corporation, Potomac, MD.

Correspondence to: Timothy R. Church, Ph.D., Division of Environmental and Occupational Health, Box 807 Mayo, 420 Delaware St. S.E., Minneapolis, MN 55455. E-mail: trc@cccs.umn.edu