ISSUES IN MEASURING EFFECTIVENESS

To be judged effective, a cancer screening protocol must either increase life spans, improve the quality of life, or both. Changes in the length and quality of life associated with a screening protocol can be both positive and negative; the net effectiveness of a strategy would depend on how such changes balance out. For example, if a positive screening test result leads to risky or uncomfortable confirmatory tests, the increased life expectancy and decreased morbidity resulting from early detection would have to be weighed against the increased mortality, morbidity, or discomfort for those who undergo the followup testing.

Accurate assessment of the full effects (both positive and negative) of a screening strategy requires controlled experiments in which observed differences in mortality and morbidity between those who undergo screening and those who do not can be validly ascribed to the screening program and not to uncontrolled differences between the screened and unscreened groups. When such studies are not available, judgments about the importance of departures from full validity must be made, and studies of more intermediate measures of effectiveness are often used.

One intermediate measure of effectiveness commonly used in evaluating colorectal cancer (CRC) screening is the ***positive predictive value*** (PPV), the percent of all positive screening tests that lead to a diagnosis of cancer or polyps. If the screening test has a high false positive rate[1], the PPV will be low. Even with a low false positive rate, however, if the disease is rare, the PPV will be low, because the vast majority of people who are screened will be disease free and the number (though not the rate) of false positive findings high. The PPV can also be expected to decline with increasing frequency of periodic screening, because the prevalence of previously undetected cases would be lower in more frequent screening programs. A low PPV implies that for every cancer (or polyp) found, a large number of people will be subjected to followup testing with its inherent medical risks and costs.

Although it is a useful indirect indicator of effectiveness, by itself the PPV is insufficient. A screening procedure with a low PPV can still be effective if the reductions in mortality or morbidity resulting from early detection are great compared to the morbidity and mortality associated with the screening and followup procedures. Consequently, the use of PPV to guide screening decisions involves implicit judgments about the relative importance of a screening strategy's benefits and risks, which must be based on other information.

Another measure of effectiveness often used in evaluations of CRC screening is the ***stage-distribution of cancers (or neoplasms) found.*** If a screening program detects a high proportion of cancers relative to the rate expected in the general population, particularly a high proportion in early stages with effective treatment available, it is sometimes reasonable to assume that this shift in the distribution of lesions found toward earlier stages (or even toward precancerous stages) will ultimately be translated into changes in mortality and morbidity, as fewer cancers progress to more serious stages. Without additional information, it is impossible to know to what degree the increase in early-stage cancers (or precancerous lesions) detected will actually translate into reductions in later cancers, because such studies are potentially biased in three ways:

- o Lead time bias - the shift in cancer stage at detection may reflect earlier diagnosis unaccompanied by equal improvements in benefit. Earlier detection of a completely incurable cancer, for example, will improve survival time but will not help the patient. Indeed, the patient may suffer unnecessary anxiety from knowing early about a cancer with effective therapy.
- o Length bias[2] - Because the length of time spent in pre-clinical stages is longer for slow-growing

[1] **The false positive rate is the percent of all people free of disease whose screening test is positive.**

[2] **This kind of bias has also been referred to as "overdetection" bias (116).**

lesions than for fast-growing lesions, slow-growing lesions have a greater chance of being detected in a periodic screening program.

These slow-growing lesions are not as invasive or lethal as fast-growing cancers. Thus, the stage shift will overestimate the number of late cancers averted or the years of life gained.[3]

- o Volunteer bias - people who agree to participate in screening (or even in a cancer screening trial) may have a different clinical course from those who do not, possibly leading to a different distribution of cancers found by screening.

In addition, reliance on the shift in the stage distribution of detected lesions as the principal indicator of effectiveness ignores the medical risks and inconvenience of the screening and followup testing required to find the early cancers.

The problem with stage-specific case finding rates as measures of effectiveness is even greater if the focus is on colorectal polyps, the suspected precursors to cancer. Since only a small minority (perhaps 5 to 10 percent) of colorectal adenomas progress to cancer (106), the potential impact of length bias is even greater, and screenees will be subjected not only to the medical risks of followup testing, but also to the risks of removal of many polyps that would not have progressed to cancer.

Because of these problems, most experts would agree that fully valid evidence on the effectiveness of any CRC screening program requires randomized clinical trials comparing mortality and morbidity rates in those offered screening with such rates in those not offered screening (39,157,158). Such studies are difficult and costly to mount, however. CRC is relatively rare and takes as long as 10 to 15 years to progress from polyps to clinically detectable stages (106), so that measurement of the effects of screening requires many participants and many years to follow the medical histories of study subjects. Despite these problems, several well-designed studies of selected CRC screening protocols are currently underway and may in time provide highly valid information on the effectiveness of certain colorectal screening protocols.

The inadequate evidence on the net effectiveness of CRC screening underlies the present disparity among experts in conclusions about the appropriate place of CRC screening in average-risk older adults. Those who require high standards of validity in studies of screening techniques generally conclude that no CRC screening protocol has been shown to be effective (48,49,50,116,133,137). Others who examine the shift in the distribution of lesions found to early or precancerous stages have concluded that the potential biases are unlikely to account for all of the benefit afforded by early detection (35,171). Eddy has observed that "There is a conceptual issue here -- how certain do you have to be before you say it is beneficial?" (81).

EVIDENCE ON EFFECTIVENESS OF FECAL OCCULT BLOOD TEST (FOBT)

Although a large literature exists on the use of the FOBT as a strategy for CRC screening, only six controlled studies of FOBT screening in asymptomatic individuals have been reported, and four of these are still underway.[4] The researchers in each of these studies have reported the PPV of the FOBT and most have studied differences between the intervention and control groups in the stage of cancers detected. Interim mortality data are available from only one study, a large ongoing study of FOBT screening in a community in Denmark begun in 1985. Table 3 summarizes the study designs and results of the trials to date.

All but one of the studies are large randomized clinical trials conducted in older average-risk individuals, beginning at ages 45- to 60-year-olds. The exception is a study of volunteers over 40 years old attending a cancer prevention clinic in New York City who were assigned to experimental and control groups according to the month in which they presented at the clinics

[3] An even stronger argument can be made with respect to screening and removal of colonic polyps. If the cancers that tend to progress rapidly are not those arising from polyps, then removing a large number of polyps may not have much effect on the incidence of cancer (101).

[4] For reviews of uncontrolled studies of FOBT screening, see Simon (137); Frank (48,49,50); and Fletcher and Dauphinee (46).

[5] In this trial, both groups received sigmoidoscopy as part of a cancer checkup. Only the experimental group received FOBT.

All six studies have reported on the PPV for cancer, adenomatous polyps, or neoplasms (cancer + adenomas). As discussed above, the PPV for a specific condition is directly related to the prevalence of the sought-after condition in the screened population and inversely related to the false positive rate. The prevalence of previously undetected disease would decrease as screening frequency increases; therefore, programs with more frequent screening intervals should have lower PPVs. Also, the first screen in a new FOBT screening program should have a higher PPV than subsequent screens. Thus, studies undertaken for longer periods of time, with periodic rescreening, should report lower PPVs overall and declining PPVs as the trial progresses. These trends are apparent in the studies. The University of Minnesota study has had the longest period of screening (at 1- and 2-year intervals) and has reported the lowest PPV for FOBT, only 2.5 percent for cancer (i.e., a positive FOBT resulted in a cancer diagnosis less than 3 percent of the time). The other controlled studies reported PPV for cancer on first and second screens in the neighborhood of 10 to 20 percent.

Because PPV depends on prevalence, it should be higher in populations with higher prevalence of CRC. Where data on PPV are available by age, the results are consistent with expectations. In the Strang Clinic study, the PPV for cancer in screenees 70 years of age or older (23 percent) was almost twice as high as the PPV in the screened population as a whole (170).

The method of preparing FOBT specimens for analysis also affects the PPV because it alters the false positive rate. Dehydration (adding a few drops of water to the test slide prior to analysis) is frequently practiced to increase the test's sensitivity to blood in the stool. But dehydration also increases the false positive rate. Consequently, the PPV of FOBT under dehydration is lower than the PPV without such a procedure. The Swedish study showed that dehydration reduced the PPV for all neoplasms by 10 percentage points. Dehydration of test slides was gradually introduced in the University of Minnesota study in order to increase test sensitivity; in all, approximately two-thirds of all slides were dehydrated. Thus, the low PPV in that study may be partly due to dehydration.

Because the prevalence of adenomatous polyps is much higher than the prevalence of cancer, particularly in elderly people, the PPV is substantially higher for neoplasms than it is for cancer alone. In the Danish study of biannual FOBT screening in 45- to 70-year-olds, 52 percent of all positive FOBTs were diagnosed either with a cancer or adenoma, compared to 7 percent for cancer alone. Among 60- to 64-year-olds in Sweden, the PPV for neoplasm (i.e., cancer plus polyps) with dehydrated slides on the second screen was 24 percent. Ransohoff and Lang have observed that the calculated PPV of the FOBT may actually reflect a random selection of elderly people for followup and detection of their polyps. To the extent that false positive FOBTs occur serendipitously in patients who happen to have polyps, the PPV will give the FOBT credit for "finding" the polyp even though it occurred by chance (123).

The success of an FOBT screening program in detecting early cancer or altering mortality rates depends in large measure on the rate of compliance with screening regimens in the population. If few people avail themselves of the opportunity to be screened, then the potential for detecting and treating cancer early is compromised. Compliance appears to vary widely across the studies, depending on the age of the screenee (older people are less compliant); the age of the program (compliance with rescreening is lower than with the first screen); the population on which randomization was based (volunteers are more compliant, at least in the beginning); and the kinds of recruitment efforts made by the program.

Despite imperfect compliance, rates of detection of CRC are consistently higher in the intervention groups than in the control groups, and a higher proportion of those found are early cancers. For example, an ongoing British trial begun in 1984 found 58 percent more cancers in the group offered screening than in the control group after 2 screening periods. These extra cancers detected were heavily concentrated in Stage A, the most curable stage of CRC. The Swedish study also found a much higher rate of cancers in the intervention group after 27 months of study, but differences in the distribution of cancer stage, which favored earlier cancers, were not

Table 3-Fecal Occult Blood Teat (FOBT) Controlled Clinical Trials

Study/site	Years	Study population	Intervention group(s)	Control group	Screening compliance rates	Dehydration status	FOBT positive predictive value	Cancer stage et detection	Mortality
Strang Clinic Colon Project, New York, NY a,b,c	1975-1079	21,008 symptomatic volunteers › 40 years old (33% › 60) attending cancer prevention clinic and followed up in 1984	Annual medical exam, rigid sigmoidoscopy and FOBT (group selected by calendar period of entry to clinic (Hemoccult and Hemobcult II)	Same as intervention except no FOBT	60-80% compliance in year 1, declining to 20-40% in year 5	Unspecified, but most No	PPV on all screens [initial + followup]: Tot PPV: 12% for cancer 36% for cancer + adenomas PPV (60-69 years old) 13% for cancer 42% for cancer + adenomas PPV › 70 years old) 23% for cancer 42% for cancer + adenomas	Percent of all cancer detected in Dukes' A and B: 65% in intervention group 33% in control group	NA
Funen, Denmark d,e	1985-ongoing	60,00045 to 70-year-olds asymptomatic for colorectal cancers, adenomas, or metastasis from all cancers	Randomized: 30,000 offered screening every 2 years with reminders	No screening offered	First screen: 65-69 70-74 men 63% 54% women 64% 50%	No	First screen: 18% cancer 58% cancer + adenomas Second screen: 7% cancer 52% cancer + adenomas	Rate of CRC detection after 34 months Test grp Control grp All CRC 0.428% 0.306% (p<0.01) Stage A CRC 0.122%0 002% (p‹0.0001) Other Stages 0.204% 0.286% (p<0.002)	First 38 months of Study: 27% reduction in CRC mortality, not statistically significant (p=0.16)
Hardcastle, Nottingham, England f	1983	20,525 patients 45 to 74-year-olds in 9 general practices in Nottingham, England without known bowel disease or cancer	Randomized: 10,253 sent instructions on Hemoccult II	No screening	35.1% men 34.8% women 27% for › 70 year	No	52% neoplasm	Rate of CRC detection Test grp Control grp All CRC 0.225% 0.097% Stage A CRC 0.092% 0.000%	NA

ABBREVIATIONS: CRC = colorectal cancer FOBT = fecal occult blood test; NA = not available; NS = not significant; PPV = positive predictive value.

a B J Flehinger, E. Herbert, S.J. Winawer et al., "Screening for Colorectal Cancer With Fecal Occult Blood Test and Sigmoidoscopy: Preliminary Report of the Colon project of Memorial Sloan-Ketting Cancer Center and PMI-Strang Clinic," J Chamberlain and A.B. Miller (eds.) *Screening for Gastrointestinal Cancer* (Lewiston, NY: Hans Huber Publishers, 198S).

b S.J Winawer, M. Baldwin, E. Herbert et al., "screening& Experience With Fecal Occult Blood Testing as a Function of Age," in *Perspectives on Prevention and Treatment of Cancer In the Elderly*, R Yancik, P.P. Cat'bone, W.B. Patterson et al. (eds.) (New York, NY: Raven Press, 1983).

c S.J. Winawer, J St. John, J. Bond et al., "Position Paper: Risk and Screening of Average Risk individuals for Colorectal Cancer," forthcoming in WHO *Bulletin*.

d o Kronborg, C. Fenger, O. Sondergaard et al., "Initial Mass Screening for Colorectal Cancer With Fecal Occult Blood Test," *Scand. J. Gastroenterol.* 22:677-686, 10S7.

e OKronborg, "Mass Screening for Colorectal Cancer With Hemoccult-II at Funen in Denmark," Interim Report, June 19SS, unpublished.

f J.D Hardcastle, P.A. Farrands, T.W. Balfour et al., "Controlled Testing in the Detection of Colorectal cancer," *Lancet* 2:1-4, 1983.

g J.D. Hardcastle, "Randomized Controlled Trial of Fecal Occult Blood: Screening for Colorectal Cancer," unpublished paper. no date.

h J.S Mandel, J.H. Bond, M Bradley et al., "Sensitivity, Specificity and Positive Predictivity of the Hemoccult Test in Screening for Colorectal cancers: The Univereityof Minnesotans Colon Cancer Control Study," unpublished paper, undated.

i J S Mandel, J.H Bond, D C Snover et al., "screening for Colorectal Cancers: The University of Minnesota's Study," unpublished paper, undated.

j J.S Mandel, University of Minnesota, personal communication, July, 1990.

k K. Klaaborg, M.S. Madsen, O. Sondergaard et al., "Participating in Mass screening Colorectal Cancer With Fecal Occult Blood Test" *Scand. J. Gastroent.* 21:1 1S0-1 1S4, 19S6.

l J, Kewenter, S.Bjork, E.Haglind et al., "Screening and Rescreening for Colorectal Cancer: A Controlled Trial of Fe@ Occult Blood Testing in 27,700 Subjects," *Cancer* 62(3):645651, 1988.

Table 3-Fecal Occult Blood Teat (FOBT) Controlled Clinical Trials (continued)

Study/site	Years	Study population	Intervention group(s)	Control group	Screening compliance rates	Dehydration status	FOBT positive predictive value	Cancer stage at detection	Mortality
Hardcastle, England [a]	1984-ongoing	107,000 50 to 74-year-olds	Randomized: 53,464 offered A) 3day Hemoccult II B) 6-day Hemoccult II with reminder; positive tests repeated; rescreen at 2-year intervals offered to those accepting first screen	No screening	Initial screen: 52.9% First rescreen: 77.0% Second rescreen: 80.0%	No	Initial screen: 10% for cancer First rescreen: 8% for cancer Second rescreen: 12% for cancer All screens: 10% for cancer	Rate of CRC detection after 2 rescreens Teat grp / Control grp All CRC: 0.33% / 0.22% (p<.0.01) Stage A CRC: o 10% / o 02% (p<.0.01) Other Stages CRC: 0.23% / 020%	NA
University of Minnesota [a,h,i]	1070-1982 (phase 1) 1 9 8 6 - (phase II - ongoing)	46,622 volunteers aged 50-80 recruited from community	Randomized: 1) annual FOBT 2) biannual FOBT	No screening	Phase I (1976-1982) 1) 75.7% 2) 76.7% declined preciplously for those ›80 years old (@ 55%)	No/Yes (gradually introduced during phase i) overall,	1st phase: 2.5% cancer 16.2% adenomatous polyps <60 = 1.690 2/3 = yes	Percent of all CRCs detected in Stage A in screen group: 35%	NA
Gothenberg, Sweden [c,j]	1982-1983; 1984-1985; and ongoing	27,503 residents of Gothenberg aged 60-64 in 1982	Randomized: First screen – mailed Hemocult II with mail return &two reminders 1) dehydrated 2) not-dehydrated Second screen – dehydrated only	No screening	First screen: 66% Second screen: 58%	First screen: random allocation of specimens to dehydration and no dehydration; on second screen dehydration only	First screen: Dehydrated – 22'% for neoplasms Not dehydrated – 32% for neoplasms Second screen: 24% for neoplasm	Rate of CRC detection after 27 months Test grp / Control grp All CRC: 0.44% / 0.15% (p<0.001) Stage A: 0.09% / 0.02% (NS) Stage B: 0.1 1% / O 04% (NS) (C& D): 0.24% / 0.09 (NS)	NA

ABBREVIATIONS: CRC = colorectal cancer; FOBT = fecal occult blood test; NA = not available; NS = not significant; PPV = positive predictive value.

[a] B.J. Flehinger, E. Herbert, S.J. Winawer et al., "Screening for Colorectal Cancer With Fecal Occult Blood Test and Sigmoidoscopy: Preliminary Report of the Colon Project of Memorial Sloan-Kettering Cancer Center and PMI-Strang Clinic," J. Chamberlain and A.B Miller (eds.) *Screening for Gastrointestinal Cancer* (Lewiston, NY: Hans Huber Publishers, 198S).

[b] S.J. Winawer, M, Baldwin, E. Herbert et al., "Screening Experience With F@ Occult Blood Testing aaa Function of Age," in *Perspectives on Prevention and Treatment of Cancer in the Elderly*, R. Yancik, P. P. Carbine, W.B. Pattereon et al. (eds.) (New York, NY: Raven Press, 19S3).

[c] S.J. Winawer, J. St. John, J Send et al., "Position Paper: Risk and Screening of Average Risk Individuals for Colorectal Cancer," forthcoming in *WHO Bulletin*.

[d] O. Kronborg, C. Fenger, O. Sondergaard et al., "Initial Mass Screening for Colorectal Cancer With Fecal Occult Blood Test," *Scand. J. Gastroentero* 1. 22:677-686, 1987.

[e] Kronborg, "Mass Screening for Colorectal Cancer wnh Hemoccult-II at Funen in Denmark," Interim Report, June 19SS, unpublished.

[f] J.D. Hardcastle, P.A. Farrands, T.W. Balfour et al., "Controlled Testing in the Detection of Colorectal Cancer," *Lancet* 2:1-4, 1983.

[g] J.D. Hardcastle, "Randomized Controlled Trial of Fecal Occult Blood: Screening for Colorectal Cancer," unpublished paper. no date.

[h] J.S. Mandel, J.H. Bond, M.Bradley et al., 'Sensitivity, Specificity and Positive Predictivity of the Hemoccult Teef in Screening for Colorectal Cancers: The University of Minnesota's Colon Cancer Control Study," unpublished paper, undated.

[i] J.S. Mandel, J.H. Bond, D.C. Snover et al., "Screening for Colorectal Cancers: The University of Minnesota's Study," unpublished paper, undated.

[j] J.S. Mandel, University of Minnesota, personal communication, July, 1990.

[k] K. Klasborg, M.S. Mad-,0, Sondergaard et al., "Participating in Mass Screening Colorectal Cancer With Fecal Occult Blood Test" *Scand. J. Gastroent.* 21:11s0-11s4, 1SS6.

[l] J. Kewenter, S. Bjork, E. Haglind et al., "Screening and Rescreening for Colorectal Cancer: A Controlled Trial of Fecal Occult Blood Testing in 27,700 Subjects," *Cancer* 62(3):645651, 1sss

statistically significant. The Swedish study is much smaller than the British trial, however, which may account for the lack so far of statistically significant differences in cancer stage distribution.

Only one of the trials has reported on mortality differences between intervention and control groups. A large trial of biannual FOBT screening of 45- to 70-year-olds in Denmark found a 27 percent lower CRC mortality rate in the group offered screening after about 3 years of study, but the number of deaths in the study so far is very small and the difference is not statistically significant by conventional standards ($p = 0.16$).

To summarize, the six controlled studies of FOBT screening suggest that in an ongoing screening program, a large number of screenees will undergo followup diagnostic tests for every CRC found, but this number decreases with the age of the screenee. It is clear that FOBT screening improves the stage distribution of cancers detected, which should translate into decreases in cancer mortality. However, even in very large trials, no such mortality effect has been identified to date. This leads to the possibility that length bias may have a strong influence on the screening programs.

EVIDENCE ON THE EFFECTIVENESS OF SIGMOIDOSCOPY

The argument for effectiveness of sigmoidoscopy as a screening tool is most forcefully made through direct comparison with the FOBT. Unlike the FOBT, which has many false negatives for cancer, sigmoidoscopy has high sensitivity and specificity for rectal or colonic lesions within its reach into the colon. Indeed, endoscopic examination is the diagnostic standard against which most other CRC detection methods are assessed (133). In studies comparing sigmoidoscopy with barium enema, sigmoidoscopy generally had a very high sensitivity -- on the order of 90 to 95 percent in detecting lesions found by either method in asymptomatic or symptomatic persons (148,167).

Because polyps and cancers are directly visualized in sigmoidoscopy, a positive finding is always a true positive. The virtual non-existence of false positives would imply a high PPV for sigmoidoscopy. If, however, one considers some of those positive findings to be clinically insignificant, then sigmoidoscopy may have a substantial false positive rate and, hence, a lower PPV. Hyperplastic polyps, for example, do not progress to cancer, but it is impossible to accurately differentiate hyperplastic polyps from neoplastic polyps without a biopsy. Consequently, when such polyps are found they are typically removed and biopsied. The clinical significance of very small polyps (i.e., those smaller than 5 mm) is also questionable (106); some researchers believe that these are highly unlikely to progress to cancer, yet they, too, are typically removed when found on an endoscopic examination.[6] Thus, for every "positive" sigmoidoscopic screening examination, a relatively small number may actually be at risk for developing into CRC.

The rapid change in endoscopic technology that occurred in the mid-1970s increased the tension between detection capability and clinical significance of lesions detected. The development of flexible fiberoptic sigmoidoscopes with lengths of up to 60 cm, compared to the 25 cm length of the rigid sigmoidoscope, increased the potential proportion of polyps and cancers that are detectable with high sensitivity and specificity with sigmoidoscopy at the same time that it increased the number of clinically insignificant lesions found and removed. A review of studies comparing flexible with rigid sigmoidoscopes found about 2.6 times as many cancers and 2.5 times as many polyps with a 60 cm flexible sigmoidoscope as with a rigid sigmoidoscope (76).

Studies of the impact of sigmoidoscopic screening on cancer incidence or mortality are even fewer than for FOBT. Only three studies of outcomes of screening programs using sigmoidoscopy have been reported and all used the rigid sigmoidoscope. Two of these were long-term observational studies of screened subjects without comparison groups. The third was a randomized clinical trial of rigid sigmoidoscopy as part of a program of periodic preventive health services offered to non-elderly

[6] This position is controversial. Tedesco found that almost 50 percent of very small polyps were adenomatous, but the study was in a symptomatic group of people (147).

enrollees in an HMO. These studies have been carefully reviewed and critiqued by several experts (101,116,133).

The two observational studies of CRC sigmoidoscopy screening programs showed dramatic shifts of detected cancers to early stages. A study of annual sigmoidoscopy examinations offered from 1946 to 1954 at the Strang Clinic in New York to 26,126 patients 45 years of age and older, most of whom (89 percent) were asymptomatic at the time of sigmoidoscopy, detected 81 percent of CRCs in Stage A or B (67). In a study of annual sigmoidoscopy offered to people 45 years of age and older in Minnesota between 1948 and 1974, all cancers detected on the second or subsequent screens were in stage A or B (54).

An analysis of cancer incidence and mortality in the Minnesota program's followup period suggested that, after eliminating cancers found on the first screen, the rate of CRC detected in subsequent years was much lower than would be expected in a like unscreened Minnesota population (54). This would imply that removal of polyps found on sigmoidoscopy prevented CRC. But the cancers found at the first screen were not prevented; rather they were found early (101,105,114,133). In a reanalysis that included pre-existing cancers, Miller concluded that the CRCs detected over the period were strikingly similar to the age-adjusted rate in an unscreened population (101). Selby and Friedman have also pointed out that the reported incidence rate in the screened group was based on the number of person-years of observations, and people may have dropped out of followup if CRC was discovered. Thus, the actual CRC incidence rate in the population offered screening was probably higher than reported in the study (133). Finally, as a program that enrolled volunteers, the cancers in the Minnesota study may have had an unrepresentative incidence and stage distribution (114), although the direction of such "volunteer bias" cannot be predicted.

The one randomized clinical trial involving annual rigid sigmoidoscopic screening for 40- to 54-year-olds as part of a multiphasic health examination for enrollees in an HMO found significantly lower death rates from CRC over an n-year period in the group offered screening than in the control group (27). On its face, this finding from a randomized trial would be strong evidence of an effect on mortality from sigmoidoscopy. But several reviewers, including one of the investigators on the original study, have called these results into question. First, the difference between the study group and control group in the use of sigmoidoscopy over the 10 year period (31 percent vs. 26 percent) was not great enough to account for the two-fold observed difference in CRC mortality (133). Second, most tumors found in both groups were detected from symptoms, not through screening (133). Third, given the design of the study, the authors probably used too lenient a test for statistical significance (101, 114,133). Thus, several reviewers have concluded that whatever real differences existed in CRC incidence and death between the study group and control group were due to factors other than the availability of sigmoidoscopy (101,114,133).

Taken as a whole, the evidence on sigmoidoscopy suggests a major shift in the stage at which CRCs are detected, but inadequate evidence that this stage shift actually reduces death rates from cancer over time. How can these two seemingly contradictory findings be reconciled? First, there has never been a good trial of the effect of screening flexible fiberoptic signodoscope (FSIG) on cancer mortality, so the lack of evidence on outcomes should not be equated with the existence of negative evidence. But, second, if future randomized studies do confirm that stage shift is unaccompanied by changes in mortality from CRC, then one must look for possible biases in screening programs. Only if therapy is no more effective in early cancers than in late cancers or if length and volunteer biases are strong can the stage shift coincide with no impact on mortality. Experts agree that therapy is much more effective in early CRC than it is once cancer has spread beyond the wall of the colon or rectum (23,145). Thus, increased lead time would not explain the lack of mortality differences between screened and unscreened groups. Many experts believe that length and volunteer biases can be powerful influences on outcome and consequently discount the evidence on stage shift as inadequate. Others see the dramatic shifts in stage at detection as unlikely to be caused simply by length

and volunteer bias; they accept this evidence as sufficient to justify periodic CRC screening with sigmoidoscopy or FOBT.

Virtually all CRCs are removed promptly after detection, so there is no direct evidence on how the length of time in each stage of cancer varies among patients. If the distribution of time spent in early stages is very wide, with some cancers progressing rapidly but many progressing slowly or hardly at all, then the **case** for length bias is strengthened. Numerous experts have commented on the high variability in the speed with which cancers grow in size or progress (29,145). Periodic screening would be likely to pick up relatively few of the rapidly progressing cancers (a high proportion of which would develop and grow in the time interval between screenings) but a large number of indolent cancers. On the other hand, if all CRCs progress at a similar speed, then the argument for length bias would be weaker, and the evidence on the proportion of early cancers detected would be compelling. Because it is unethical to leave CRCs detected but untreated, direct observation of the distribution of cancer progression rates is infeasible; only controlled cancer screening trials of sufficient size and duration will provide definitive information on the extent of length bias. The National Cancer Institute (NCI) has recently announced plans for a 5-year randomized clinical trial of sigmoidoscopy screening in men 60- to 74-year-olds (152). This trial, which will enroll enough men to detect a 20 percent decrease in mortality, will test whether a sigmoidoscopy every 3 years will affect outcomes in the elderly. The results of the trial will probably not be available for at least 10 to 15 years.

EFFECTIVENESS ISSUES SPECIFIC TO THE ELDERLY

Although all CRC screening programs have targeted older people (generally over 45 or 50 years of age), screening in the elderly (people 65 years of age and older) raises issues that may not be so important in middle-aged people.

High Incidence of CRC in the Elderly--The incidence of CRC rises dramatically with age (see Chart I); the incidence of CRC **at ages** 70-74, for example, is 5.7 times higher than at ages 50-54 (152). At the same time, it does not appear that CRC progresses at a different rate in the elderly; the distribution of stage at detection is virtually the same in the elderly as in the non-elderly (58,73,75,104). Thus, the potential burden of illness in those over 65 years of age is much higher than for other groups, and the potential effectiveness of screening in reducing morbidity and mortality is higher as well.

The distribution of cancers throughout the large intestine appears to be different in the elderly. The elderly tend to have more CRC located proximal to the splenic flexure than do the non-elderly (45,63,139). In a Swedish study of 264 patients with polyps found through colonoscopy, polyps in patients over 65 years of age were much more uniformly distributed throughout the large intestine than were polyps found in people 55 years of age and under. There, almost 70 percent of all polyps were found in the sigmoid colon and rectum, compared to only about 35 percent of polyps in the elderly (59,161). Consequently, within their limited reach, sigmoidoscopes would probably detect a smaller fraction of all CRCs in the elderly than in non-elderly screenees.

High Prevalence of Colorectal Polyps in the Elderly--The prevalence of asymptomatic benign adenomatous polyps increases dramatically in the elderly. Autopsy studies conducted in the United States and other countries over the years have consistently found an increase in the prevalence of polyps with age up to approximately 60- to 70-year-olds. Beyond that point, the prevalence of polyps shows no systematic relationship to age (table 4).

Although the frequency of adenomas increases dramatically with age, the average size of these lesions does not vary with age, which suggests that "while new adenomas develop with aging most tend to remain static in size after reaching a diameter of less than 10 mm" (128).

If adenomas are precursors of the vast majority of CRCs, as many researchers have suggested (25,35,106,108), then removal of adenomas would appear to be prudent even if the great majority of them will not develop into cancer, but this would imply that a large number of elderly people entering

Table 4-Age-Specific Polyp Prevalence: Autopsy Studies

Study	Year	population	Polyp type	Age-specific prevalence rate (number) 60-69 Male	60-69 Female	70-79 Male	70-79 Female	80 and over Male	80 and over Female
United States:									
Rickert et al., 1979	NG	518 autopsies of males and females between ages 20 and 102 not previously diagnosed with carcinoma of the bowel	Adenomatous	59.8% (58)	48.8% (19)	68.996 (51)	40.0% (20)	61.1% (22)	63.0% (17)
Stemmermann and Yatani, 1973	1960-1972	Autopsies from 202 Hawaiian Japanese at the Kuakini hospital in Honolulu, HI	Adenomatous	60.0% (22)	59.0% (10)	70.0% (16)	60.0% (18)	66.0% (23)	83.0% (12)
			Hyperplastic	81.0% (30)	77.0% (13)	87.0% (20)	67.0% (21)	86.0% (30)	16.0% (16)
Arminski and McLean, 1964	NG	1,000 autopsies of men and women aged 20 and over at the Grace Hospital, Detroit, MI	Adenomatous	39.9% (79)	25.3% (28)	46.1% (53)	47.3% (35)	46.3% (19)	42.5% (17)
Chapman, 1963	NG	Autopsies of 443 adults in New York hospital	Adenomatous	43.0%[a] (NG)	37.0% (NG)	65.0% (NG)	37.0% (NG)	63.0% (NG)	50.0% (NG)
Blatt, 1961	1960	556 autopsies performed during a 9-month period in NY (446 colons used in the study)	Adenomatous	35.0% (23)	37.5% (15)	45.0% (42)	46.0% (29)	46.7% (21)	40.0% (20)
Correa et al., 1977	1970-75	301 autopsies performed in New Orleans	Hyperplastic	14.3%[b] (177)	10.4%[b] (124)				
Other countries:									
Williams et al., 1982	NG	365 autopsy specimens in a l-year period in Liverpool, England. 134/365 came from cases dying in hospital.	Adenomatous	44.0%[d] (25)	35.0% (18)	52.0% (22)	33.0% (21)		
			Hyperplastic	47.0% (27)	13.0% (7)	33.0% (14)	34.0% (22		
Restrepo et al., 1981	1971-1973	529 specimens from consecutive autopsies of persons age 10 and over in a hospital in Medellin, Colombia	Adenomatous	7.0%[b] (32)	16.7% (24)				
			Hyperplastic	31,2% (lo)	29.2% (7)				
Hughes, 1968	1964	200 colons from autopsies done in Queensland, Australia examined for polyps	Not Specified	22%[c] (10)		17% (11)		32.4% (12)	

NG = not given.
a In this study groups are divided at age 65 to 74, 75 to 84, and 85 and over.
b In this study the prevalence rate for elderly people is reportad as one group, age 60 and over.
c Male and female prevalence rates are combined.
d In this study groups are divided at age 65 to 74, and 75 and over.

SOURCE: T.C. Arminski end D.W. McLean, "Incidence and Distribution of Adenomatous Polyps of the Colon and Rectum Based on 1,000 Autopsy Examinations," *Dis. Colon Rectum* 7:249–261, 1964; LJ. Blatt, "Polyps of the Colon and Rectum: Incidence and Distribution," *Dis. Col.* Rec. 4:277-282, 1961; 1. Chapman, "Adenomatous Polypi of Large Intestine: Incidence and Distribution, "Ann. *Surg.* 157(2):223-226, 19S3; P. Correa, J.P. Strong, A. Reif, et al., "The Epidemiology of Colorectal polyps: Prevalence in New Orleans and International Comparisons," *Cancer* 39:2258-2264, 1977; LE. Hughes, "The Incidence of Benign and Malignant Neoplasms of the Colon and Rectum: A Post Mortem Study, " *N.Z. J. Surg.* 38(1):30-35, 19SS; C. Restrepo, P Correa, E. Duque, et al., "polyps in a Low-Risk Colonic Population in Columbia, South America. " *Dis. Colon Rectum*, 24:29-36, 1981; R.R. Rickert, O. Auerbackk, L Garfinkel, et al., "Adenomatous Leaions of the Large Bowel: An Autopsy Survey," *Cancer* 43:1847-1857, 1979; G.N. Stemmerman and R. Yatani, "Diverticulosis and Polyps of the Large Intestine: A Neocropsy Study of Hawaii Japanese, " *Cancer* 31(5):1260-1270, 1973; A.R. Williams, B.A. Balascoriya, D.W. Day, et al., "Polyps and Cancer of the Large Bowel: A Necropsy Study in Liverpool," Gut 23:835–842, 19S2.

a screening program for the first time would have one or more polyps removed, with consequent risks and costs. In addition, with FOBT and sigmoidoscopy as the screening tools, elderly people would not be likely to have as high a proportion of polyps found and removed as would the non-elderly, mainly because more are located beyond the reach of the sigmoidoscope and are not likely to be picked up by the FOBT. If the detection and removal of polyps are important for the success of a CRC screening program, then perhaps a screening tool with greater sensitivity and reach into the colon, such as DCBE or colonoscopy, would be more effective (but also more costly) in the elderly.

Decreased Acceptability of Screening Procedures--There is some question as to whether the elderly find the current screening methods acceptable. The discomfort of sigmoidoscopy and the preparation required for all screening, including dietary restrictions, purging, etc. may be more difficult for elderly people to undertake than younger people.

Increased Fragility in the Aged--The ability of elderly people to withstand the discomfort and risks of sigmoidoscopy and colonoscopy is highly variable, of course, but the issue of increased complications with age needs to be considered. The elderly tend to have more underlying diseases that may make it more difficult to pass these instruments (135), and rates of post-colonoscopy hemorrhage may be higher in the elderly than in the non-elderly (30). In addition, the dietary preparation needed both for adequate endoscopy and DCBE may be more difficult for frail or very old people to withstand (135). Especially in the very elderly (those 80 years of age and above), the question must be asked how much morbidity is associated with the screening tests themselves and any followup procedures, such as the DCBE, colonoscopy, or surgery (120). Studies of resection for CRC in the elderly showed increasing postoperative complication and mortality rates with advanced age (42,89).

Decreased Life Spans in the Elderly--The progression from polyp to cancer and from early cancer to late cancer is not well understood, but the process is generally thought to be gradual, and does not change with age. Studies of the growth rates of small CRCs have indicated that colorectal tumors may grow from 5 mm to 1 cm in a median time of 30 to 40 months, with a lower bound of 12 to 16 months (18). Some experts claim that it takes from 5 to 10 years for an adenomatous polyp of 1 cm in size to develop into cancer (35,135). The slow progression of polyps to cancer raises the question whether screening for polyps in patients over 70 years of age can increase longevity (135). Yet, avoiding CRC has benefits that are independent of effects on length of life. The real question is whether detecting and removing the polyps and early cancers are worth the inconvenience, medical risk, and cost that are implied for the very old.