MINI-REVIEW

The unsolved problem of surveillance for colorectal cancer in ulcerative colitis

Fabiola Delcò MD MPH, Amnon Sonnenberg MD Msc

THE RATIONALE FOR SURVEILLANCE

Patients with long-standing extensive ulcerative colitis and the increased risk for developing colorectal cancer suggests that surveillance by frequent and regular colonoscopies may provide a means for reducing cancer-related morbidity and mortality. A crude calculation suggests that such a surveillance program would also be cost effective. None of several clinical trials, however, has been able to provide unequivocal evidence in favour of surveillance. The major reason for this failure relates to the prohibitively large number of patients with ulcerative colitis who need to be followed over a prolonged time period before statistically sound results would be obtained. Because models of decision analyses themselves have to rely on medical evidence, they cannot provide a substitute for deficient clinical data. The issue of surveillance colonoscopy cannot be resolved by the available knowledge or analytical tools. Hopefully, new techniques of surveillance or even a cure for ulcerative colitis will render the question mute of whether or not to screen patients with ulcerative colitis by frequent colonoscopy.

Key Words: Colon cancer; Colonoscopy; Cost-effectiveness; Decision analysis; Sample size; Screening; Surveillance; Ulcerative colitis

This mini-review was prepared from a presentation made at the World Congress of Gastroenterology, September 6 to 11, 1998, Vienna, Austria

Department of Veterans Affairs Medical Center, and The University of New Mexico, Albuquerque, New Mexico

Correspondence and reprints: Dr A Sonnenberg, Department of Veterans Affairs Medical Center 111F, 1501 San Pedro Drive SE, Albuquerque, New Mexico 87108, USA. Telephone 505-265-1711, fax 505-256-5751, e-mail sonnenb@unm.edu

Received for publication January 15, 1999. Accepted January 20, 1999

Can J Gastroenterol Vol 13 No 8 October 1999
Delcò and Sonnenberg

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Endoscopies invested per cancer prevented or life-year saved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Value</td>
</tr>
<tr>
<td>Efficacy of surveillance in preventing death (eff)</td>
<td>50%</td>
</tr>
<tr>
<td>Incidence of colorectal cancer (inc)</td>
<td>1% per year</td>
</tr>
<tr>
<td>Frequency of colonoscopy per year (freq)</td>
<td>1 per 2 years</td>
</tr>
<tr>
<td>Life expectancy at cancer prevention (LE)</td>
<td>30 years</td>
</tr>
<tr>
<td>Colonoscopies per cancer prevented (colon)</td>
<td>100</td>
</tr>
<tr>
<td>Colonoscopies per life-year saved</td>
<td>3.3</td>
</tr>
</tbody>
</table>

with lesion or mass. The expected shift of cancer diagnosis toward an earlier stage when treatment is more likely to be effective (ie, secondary prevention) may be an additional benefit of surveillance colonoscopy.

Few authors differentiate between the terms ‘screening’ and ‘surveillance’. The term ‘screening’ refers to the first colonoscopy done to assess the initial severity of illness and the prevalence of mucosal dysplasia. The term ‘surveillance’ refers to the continuous sequence of colonoscopies done at regular time intervals, independently of patient symptoms, to evaluate the incidence of mucosal dysplasia. In general, the yield of a single screening procedure with respect to pathological findings is much higher than the yield of each surveillance colonoscopy. Although in a few instances the two terms may be used interchangeably, the contents of the present article relate only to surveillance as a sequence of frequent colonoscopies performed in asymptomatic patients to prevent or reduce cancer-related mortality.

The investment in colonoscopies per prevented death from colorectal cancer and saved life year were calculated (Table 1). The joint influence of histological examination and timely proctocolectomy is assumed to prevent 50% of all deaths from colorectal cancer. In a surveillance program with biannual colonoscopy, 100 procedures are invested to prevent one death from colorectal cancer. If each prevented death is associated with a remaining life expectancy of 30 years, about three colonoscopies are invested per saved life-year. According to these crude calculations, surveillance for colorectal cancer appears to be a rather cost effective strategy. How well supported are these theoretical considerations by the existing clinical evidence?

**THE PROBLEMS OF CLINICAL EVIDENCE**

An historical cohort of patients with pancolitis of at least nine years’ duration was assembled from the University of Chicago Inflammatory Bowel Disease Registry (9). Rates of colectomy, cancer incidence and death were compared among 91 cases undergoing surveillance and 95 controls outside the surveillance program. No statistically significant difference was observed in the cancer rates of the two groups. Although the distribution of the cancerous lesions in the surveillance group was shifted slightly toward earlier stages, the difference was not significant. Surveillance was found to be associated with delayed colectomy and improved survival. The reduction in mortality achieved in the surveillance group, however, could not be ascribed to a reduction in mortality from colorectal cancer, which was actually lower in the unscreened population. The results were unaffected by adjustment for the potentially confounding influences of sex and age at symptom onset. These findings, therefore, failed to confirm a beneficial effect of surveillance. The scientific merits of the study are somewhat limited by the fact that the data were generated retrospectively. The authors had no control over the potential bias of patients with distinctive characteristics being selected preferentially into the case or control population.

Choi and co-workers (10) compared cancer stage with survival among 19 cases with and 22 controls without colonoscopic surveillance. The 19 cases underwent a median of two surveillance colonoscopies before developing cancer, while the 22 controls underwent colonoscopy only after developing symptoms from their cancer. Colon cancer was detected at a significantly earlier Dukes’ stage in the surveillance group. After adjustment for the time difference in the length of follow-up, the overall five-year survival rate was 77% in the surveillance group compared with 36% in the control group without surveillance. The authors concluded that colonoscopic surveillance improved the survival of patients with ulcerative colitis in whom colorectal cancer developed. The observed survival benefit appeared to be mainly derived from the detection of cancer at an earlier stage. The study outcome, however, did not provide clear-cut proof in favour of surveillance. Enrolment in a surveillance program and frequent physician contacts rather than colonoscopy itself may have been responsible for the seeming benefit of early detection. There was no prospective random enrolment of patients into either study group. No protective influence of surveillance against cancer was detected because the analysis was confined to patients with ulcerative colitis who had already developed cancer. Because the majority of patients from the two study populations with and without surveillance were excluded from the analysis, the results leave open the possibility that incidence rates of colorectal cancer were similar in the two populations.

Between 1978 an 1990, a cohort of 160 patients with long-standing and extensive ulcerative colitis was enrolled in an endoscopic surveillance program at the Leeds Royal Infirmary (11). The authors kept records of all colitis-related cancers seen among patients outside the surveillance program at their institution during the same time period. Only one cancer was diagnosed as a result of surveillance, whereas nine cancers were diagnosed in patients outside the program. These nine cancers occurred in patients who abandoned the program, who were not eligible for enrolment and who were lost to follow-up before the program was initiated. The authors concluded that their findings cast doubts on the effectiveness of endoscopic surveillance. The study design
did not include comparison with a prospective control group, and the size of the unscreened population was unknown.

The experience with surveillance in patients with extensive ulcerative colitis at the St Mark’s Hospital was reported in two articles published in 1994 (12,13). In the first analysis, the study population was comprised of 120 patients who were operated for colorectal cancer between 1947 and 1992 (12). The surveillance program was offered to patients with long standing extensive colitis only since 1966. The five-year survival rate in 16 patients in whom surveillance was practised was significantly higher than that of 104 patients who did not undergo surveillance (87% versus 55%). The main source of bias in this study stems from comparing survival in a small group of recent cancers diagnosed between 1966 and 1992, with a much larger group of cancers diagnosed between 1947 and 1992.

To overcome this obvious shortcoming, in their second study, Connell and co-workers (13) restricted their analysis to the cohort of patients who underwent surveillance after 1971, when colonoscopy became available as the primary screening tool. In 248 patients in the surveillance program, 11 cancers associated with no deaths were diagnosed as a direct benefit of surveillance, while six cancers associated with four death were diagnosed because of symptoms. In 48 patients who left the surveillance program, three cancers associated with one death were diagnosed. Eleven cancers developed in an unknown number of patients outside the surveillance program. Although these data may suggest a benefit of surveillance, the lack of any prospective control group, again, makes it rather difficult to assess its true impact in an unambiguous way.

A nested case-control study from Sweden tested the influence of surveillance on mortality from colorectal cancer in patients with ulcerative colitis (14). The case population comprised all 40 patients who died from colorectal cancer since 1975. The control population of 102 patients was randomly chosen from all patients with ulcerative colitis for at least five years who did not die from colorectal cancer. Case and control subjects were matched by sex, age at diagnosis, duration of disease and extent of disease at diagnosis. Surveillance exerted a protective influence against death from colorectal cancer with a relative risk (RR) of 0.29 (CI 0.06 to 1.31) for one procedure and an RR of 0.22 (CI 0.03 to 1.74) for two or more procedures during surveillance. As indicated by the wide confidence intervals, including unity, the protective influence did not reach statistical significance.

Overall, the clinical evidence is inconclusive at best. The larger studies using control groups for comparison failed to show a statistically significant benefit of surveillance colonoscopy. Those studies that alluded to a benefit of surveillance can all be blamed for subgroup analysis and potential selection bias. Because, at least on theoretical grounds, the benefit of surveillance appears to be obvious, why have clinical trials failed to provide clear-cut evidence in its favour? Part of the difficulty may relate to an underlying fundamental statistical problem.

Can J Gastroenterol Vol 13 No 8 October 1999

### The Statistical Problem

Until now, no randomized clinical trial has been undertaken to assess the effectiveness of endoscopic surveillance in ulcerative colitis. During the past two decades, surveillance by colonoscopy became the standard of care for many gastroenterologists. As a result, the incentives for setting up such a trial have decreased. Moreover, the trial would require a very large number of patients followed over prolonged periods of time to reach statistical power. We calculated the number of patients with ulcerative colitis and the number of person-years needed to prove prospectively that surveillance by colonoscopy for precancerous lesions is truly beneficial (Table 2). It is assumed that surveillance halves the mortality rate compared with no surveillance. The goal of a prospective trial would be to demonstrate, for instance, that the rate difference between 45% without and 23% with surveillance is statistically significant. Fleiss’ book (15) contains a readily applicable formula for calculating the number of cases needed in each study arm (16). To avoid a false positive error of alpha = 0.05 and false negative error of beta = 0.2 in the trial, a sample size of 77 would be required in each of the two study arms. The average annual incidence rate of colorectal cancers in ulcerative colitis serves to translate the overall number of 154 cancer cases into a total number of 15,400 person-years.

One hundred and fifty-four colon cancers arising from ulcerative colitis could still be recruited in a multicentre case-control study. The main obstacle to a prospective cohort study evaluating the impact of surveillance arises from the relatively low incidence of colorectal cancer. Large populations need to be followed for long time periods before the necessary numbers of cancer cases accumulate in each study arm. In the present example, 1540 subjects with ulcerative colitis would need to be followed over 10 years. The expenses associated with the recruitment and prospective long term follow-up of such a large patient population are prohibitive. The prevalence of extensive ulcerative colitis is relatively small, and even large medical centres committed to the study of inflammatory bowel disease rarely follow more than 200 patients at any given point in time. It is not foreseeable that any randomized clinical trial will become available to provide a conclusive answer. Considering the

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case fatality rate without surveillance</td>
<td>45%</td>
<td></td>
</tr>
<tr>
<td>Detected reduction of case fatality rate</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>Case fatality rate with surveillance</td>
<td>23%</td>
<td>45%/30%</td>
</tr>
<tr>
<td>Cancers in each study arm</td>
<td>77</td>
<td>From reference 13</td>
</tr>
<tr>
<td>Total number of cancers</td>
<td>154</td>
<td>2x77</td>
</tr>
<tr>
<td>Incidence rate of cancer in ulcerative colitis</td>
<td>1% per year</td>
<td></td>
</tr>
<tr>
<td>Person-years of follow-up</td>
<td>15,400</td>
<td>154/1%</td>
</tr>
</tbody>
</table>
obstacles to a clinical solution, is it possible that medical decision analysis could ascertain the feasibility of surveillance?

THE PROBLEM WITH MEDICAL DECISION ANALYSIS

The reader interested only in the main findings of the decision analysis could skip this and the following two paragraphs. The present paragraph explains the structure of the decision tree shown in Figure 1, proceeding from left to right and from top to bottom. The filled square on the very left side symbolizes the initial decision in favour of or against surveillance. The chances for or against the development of colorectal cancer are denoted as $p$ and $1-p$, respectively. In the case of cancer, surveillance colonoscopy plus histology can yield true positive (TP) or false negative test results. Cancers prevented or detected as a consequence of surveillance are associated with proctocolectomy. The cancer-related mortality in patients with surveillance ($\text{mort}_1$) reflects the

Figure 1) Decision tree of surveillance for colorectal cancer (CRC) in patients with long-standing extensive ulcerative colitis. FN False negative; FP False positive results of surveillance colonoscopy; HRQL Health-related quality of life; $\text{mort}_1$ Cancer-related mortality in patients with surveillance; $\text{mort}_2$ Cancer-related mortality in patients without surveillance; $p$ Cumulative probability of developing CRC; TN True negative; TP True positive

Can J Gastroenterol Vol 13 No 8 October 1999

Delcò and Sonnenberg
impact of cancers that could not be prevented through surveillance and proctocolectomy. The life gained becomes reduced by the impaired health-related quality of life (HRQL) after proctocolectomy. The final outcomes of TP and false negative surveillance tests are quite similar, except for the higher mortality rate (cancer-related mortality in patients without surveillance [mort2]) associated with cancers missed during surveillance. Colonoscopy in patients without dysplasia can yield true negative (TN) or false positive tests. Life and its quality remain unaffected by TN surveillance procedures. False positive tests lead to an unnecessary proctocolectomy and a reduction in HRQL. The main lower branch of the tree represents the decision against surveillance. It has only two possible outcomes governed by the probability of developing cancer. In the case of cancer, the outcome is identical to that of a missed cancer as a consequence of false negative surveillance tests. Without cancer, the outcome is identical to the outcome associated with TN surveillance tests, that is, life unaffected by cancer or proctocolectomy.

The decision tree was applied to the example of a 45-year-old patient with ulcerative colitis first diagnosed at age 35 years. Because the patient’s life expectancy is 34 years (17), a biannual surveillance program would require 17 colonoscopies at a total cost of surveillance of 17 × US$911 = US$15,486. The costs for a single colonoscopy (US$911) were estimated from the average reimbursement schedule of the United States Health Care Finance Administration during fiscal year 1998. They include all cost items associated with each individual procedure, such as facility costs, physician fees, costs of surgical pathology and the expected average costs of complications such as bleeding or perforation. By using the human capital approach, the value of life was equated with the average annual earnings (18) multiplied by the life expectancy, that is,

\[\text{Life} = 34 \times US\$24,294 = US\$826,000\]

The HRQL after proctocolectomy was assumed to be 95% compared with 100% for an unoperated healthy individual. A recent study reported a mort1 of 15% in a population with surveillance as opposed to a mort2 of 45% in a population without surveillance (12). The sensitivity of colonoscopy in detecting premalignant lesions and preventing cancer-related death was estimated to be TP=80%, and the specificity was estimated as a TN=60% based on data taken from Connell et al (13). As described in the first paragraph of this article, the cumulative probability P for developing cancer depends on the length of the history of the disease.

For surveillance to be the preferred management strategy, the upper main branch of the decision tree should result in a higher yield than the lower main branch, that is,

\[
P = \frac{\text{surveillance} + (\text{TN} \cdot c + \text{FP} \cdot d - f) - (\text{FP} \cdot a + \text{FN} \cdot b - e)}{(\text{TN} \cdot c + \text{FP} \cdot d - f) - (\text{FP} \cdot a + \text{FN} \cdot b - e)}
\]

where the letters ‘a’ through ‘f’ are used as short forms to indicate the various outcomes. The cost of surveillance enters the equation with a minus sign, as opposed to the plus sign associated with the benefit of life-years saved. The equation from above can be solved for the probability value of P:

\[
P = \frac{\text{surveillance} + (\text{TN} \cdot c + \text{FP} \cdot d - f)}{(\text{TN} \cdot c + \text{FP} \cdot d - f) - (\text{FP} \cdot a + \text{FN} \cdot b - e)}
\]

Although the formula may look daunting, P is readily calculated on a spreadsheet. We obtained a probability of P≥16% by using the values introduced in the preceding paragraph. This P value suggests that, if the probability of developing cancer exceeds 16%, surveillance would represent a decision preferred over no surveillance. In the present example of a 45-year-old patient with a 10-year history of ulcerative colitis, the cumulative probability of developing cancer over the patient’s remaining lifetime of 34 years is as follows:

\[
P = 10 \text{ years} \times 0.5% + 10 \text{ years} \times 10% + 14 \text{ years} \times 15% = 36%\]

The threshold of 16% would be reached in 21 years from now at the age 66, because:

\[
P = 10 \text{ years} \times 0.5% + 11 \text{ years} \times 10% = 16%\]

If the risk of developing colitis-related cancer is considered, biannual surveillance appears to be the better medical decision. The decision model outlined in the preceding three paragraphs uses the concept of a threshold analysis (19). If the probability of developing cancer exceeds a threshold value P, surveillance becomes the preferred management strategy. Accordingly, the strength of the argument in favour of surveillance is directly proportional to the threshold value. A very low threshold value would be evidence strongly in favour of surveillance. A very high threshold value exceeding the lifetime probability of developing cancer would be evidence against the use of surveillance colonoscopy.

Because the value of life appears in the final outcomes of all six branches in the decision tree of Figure 1, the actual costs calculated by the human capital approach exert little influence on the outcome of the analysis. The cost of surveillance pales in comparison with the benefit of life-years saved. Therefore, variations in the cost of surveillance also exert relatively little influence on the threshold probability. However, the outcome of the analysis very much depends on the other assumptions built into the model. It has been suggested, for instance, that HRQL remains largely unaffected by proctocolectomy (20). Increasing HRQL from baseline 95% to 100% halves the threshold value from baseline 16% to 8%. In the baseline analysis shown in Figure 1, the following set of values were chosen: mort1=15%, HRQL=95% and TP=80%. Slight variations lead to a second set of values, such as mort1=25%, HRQL=85% and TP=70%. The resulting threshold value P=44% lies outside the cumulative lifetime risk of the patient ever developing colorectal cancer. The second set of assumptions is by no means extreme and seems to fall well within a reasonable range that might be expected by a widely distributed surveillance program.

As these examples show, one can conceive similarly reasonable sets of assumptions that result either in excessively...
high or low thresholds. Depending on one’s preferences, the decision analysis can be used to either defend or refute the usefulness of surveillance colonoscopy. To narrow down the possible range of each assumption built into the decision analysis, a better set of data are needed that can only be obtained through clinical studies. The authors have come full circle. Because decision analysis is based on clinical data, it cannot substitute for their deficiency. More refined decision models can be envisaged that account for the time-dependent development of dysplastic lesions and their multistep transition into cancer (21-23). Such models can also consider the patient compliance, the impact of the surveillance procedure itself on the HRQL, the expenditures arising from medical and surgical therapy, as well as the indirect costs arising from missed work days secondary to surveillance. However, complicated such models are, they cannot dodge the issue of incomplete clinical information.

SUMMARY OF THE PROBLEMS

The association between long standing extensive ulcerative colitis and the increased risk of developing colorectal cancer suggests that surveillance by frequent and regular colonoscopies may provide a means for reducing cancer-related morbidity and mortality. A crude calculation suggests that such a surveillance program would also be quite cost effective. None of several clinical trials, however, has been able to provide unequivocal evidence in favour of surveillance. The major reason for this failure relates to the prohibitively large number of patients with ulcerative colitis who need to be followed over a prolonged time period, before statistically sound results are obtained. Medical decision analysis cannot help in resolving this clinical dilemma. To provide precise and reliable estimates regarding the usefulness of surveillance, decision analysis has to rely on the very clinical evidence that is missing.

During the past two decades, many pages in gastroenterology and medicine journals have been devoted to editorials and reviews written in favour of and against endoscopic surveillance in ulcerative colitis. The issue of surveillance colonoscopy cannot be resolved by the available knowledge or analytical tools. Gastroenterologists may feel uncomfortable with such an important problem of their subspecialty waiting for its solution. However, historical experience indicates that most true problems have never been solved but only rendered irrelevant by the passage of time (24). Chances are good that eventually completely new surveillance techniques or even a cure for ulcerative colitis will turn mute the question of whether to screen patients with ulcerative colitis by frequent colonoscopy.

ACKNOWLEDGEMENTS: Fabiola Delcò was supported by the Swiss Foundation for Grants in Medicine and Biology.