Users’ Guides to the Medical Literature

CLINICAL SCENARIO
You are a family physician seeing a 47-year-old woman and her husband of the same age. They are concerned because a friend recently found out that she had bowel cancer and has urged them both to undergo screening with fecal occult blood tests (FOBTs) because, she says, prevention is much better than the cure she is now undergoing. Both your patients have no family history of bowel cancer and no change in bowel habit. They ask whether you agree that they should be screened.

You know that trials of FOBT screening have demonstrated that screening can reduce mortality from colorectal cancer (CRC), but you also recall that FOBTs can have a high false-positive rate that then requires investigation by colonoscopy. You are unsure whether screening is worthwhile.5-8 Sometimes screening is clearly effective, with large benefits and negligible harms, as is the case with phenylketonuria screening and screening for systolic hypertension (>160 mm Hg) among the elderly.9 In other situations, clinicians must often weigh the benefits and harms when considering whether to screen.10 This guide extends earlier approaches by providing a framework for assessing the methodological strength of guidelines on screening and by demonstrating the importance of weighing the benefits and harms of screening when they are closely balanced. The final decision about whether to screen is greatly influenced by the values different individuals place on each of the possible benefits and harms.

Our criteria for reviewing a guideline (or a meta-analysis) about screening for CRC that might help you.

THE SEARCH
Since you know there is more than 1 randomized controlled trial (RCT), you look first for a systematic review. Your MEDLINE search (using the terms fecal occult blood test and colorectal or colonic neoplasms and mass screening and systematic review) produces a systematic review by Towler et al.1 However, there may be ancillary evidence that would influence your decision about whether to recommend screening to your patient (such as the false-positive rate of the test, the adverse effects of subsequent investigation and treatment, and costs) so you also check for a practice guideline. You find the American Gastroenterological Association (AGA) guideline on CRC screening,2 which is based on the same trials as the systematic review but also provides the additional information you were hoping to find. The full text is provided so you print off a copy to take home and read.

INTRODUCTION
When assessing a guideline or recommendation about screening you should apply the criteria suggested earlier in this series about assessment of health care interventions.3,4 You may also consider other criteria for evaluating whether screening is worthwhile.5-8
Are the recommendations valid? Are the data identified, selected, and combined in an unbiased fashion? Were the data indentified, selected, and combined in an unbiased fashion? Was the evidence of benefit overwhelmed? RCT assessment is required. The study may assess the entire screening process (early detection and early intervention, Figure 1, left), in which case people are randomized to be screened and treated if early abnormality is detected or not screened and treated (and treated only if symptomatic disease occurs). Trials of mammographic screening have used this design.12-14 Alternatively, everyone may participate in screening and those with positive test results are randomized to be treated or not treated (Figure 1, right). If those who receive treatment do better, then one can conclude that early treatment has provided some benefit. Investigators usually use this design when screening detects not the disease itself, but factors that increase the risk of disease. Tests of screening programs for hypertension and high cholesterol levels have used this design.15,36 The principles outlined in this article apply to both screening for occult disease and screening for risk factors for later disease.
The AGA guideline\(^1\) on colorectal screening used explicit inclusion and exclusion criteria and a comprehensive search to identify all the RCTs of FOBT screening. The authors include a critical appraisal of the trials and conclude that the trials provide strong evidence of effectiveness, though they are limited in that they do not consider the effect of screening on health-related quality of life.

**What Are the Benefits?**

What outcomes need to be measured to estimate the benefits of a screening program? Benefits will usually be experienced by some of those with positive test results, as either a reduction in mortality or an increase in quality of life. The benefit can be estimated as an absolute risk reduction (ARR) or a relative risk reduction (RRR) in adverse outcomes. (Readers desiring a full discussion of these concepts can refer back to an earlier Users’ Guide.\(^2\)) Briefly, the ARR depends on the baseline risk of disease and thus presents a more realistic estimate of the size of the mortality benefit. The RRR, in contrast, is independent of baseline risk and can lead to a misleading impression of benefit (TABLE 3). The number of people needed to screen to prevent an adverse outcome provides another way of presenting benefit.

In addition to prevention of adverse outcomes, people may also regard knowledge of the presence of an abnormality as a benefit as in antenatal screening for Down syndrome. Another potential benefit of screening comes from reassurance afforded by a negative test result, if a person is experiencing anxiety because a family member or friend has developed the target condition or from discussion in the media. However, if the anxiety is a result of the publicity surrounding the screening program itself, we would not view anxiety reduction as a benefit.

The AGA guideline reports that the RRRs from 3 trials of FOBT screening are 33% (annual screening) and 15% and 18% (biennial screening). An estimate of the uncertainty associated with these estimates (as one would get from the 95% confidence interval [CI] around a pooled RRR) would help the reader appreciate the range within which the true RRR plausibly lies. Based on a computer simulation, the AGA guideline estimates an ARR of 1330 deaths prevented per 100,000 (13.3 per 1000) people screened annually using FOBT from 50 to 85 years of age, assuming 100% participation (TABLE 4).

**What Are the Harms?**

Among those with positive test results, harms may include the following:

- complications arising from investigation
- adverse effects of treatment
- unnecessary treatment of persons with true-positive test results who have inconsequential disease
- adverse effects of labeling or early diagnosis
- anxiety generated by the investigations and treatment
- costs and inconvenience incurred during investigations and treatment.
The AGA guideline reports that of the patients who do not have CRC, 8% to 10% will have false-positive test results (specificity, 90%-92% using rehydrated slides). In the trials, only 2% to 6% of those with positive test results actually had colon cancer (positive predictive value, 2%-6%). Thus, of every 100 screening participants with a positive test result, only 2 to 6 will have cancer, but all 100 will be exposed to colonoscopy and its attendant risks (Table 4). While the colonoscopies will reveal few cancers, they will show many polyps (25% of people aged 50 years or older have polyps, some of which will be judged to need removal depending on the size of the polyp). Part of the benefit of screening will come from removal of the small proportion of polyps that would have progressed to invasive cancer. Part of the harm of screening will come from regular colonoscopies that are recommended for people who have had a benign or inconsequential polyp removed.

Among those with negative test results, harms may include the following:
- anxiety generated by the screening test (waiting for result)
- false reassurance (and delayed presentation of symptomatic disease later)
- costs and inconvenience incurred during the screening test.

Of those who have cancer, FOBT screening using rehydrated slides will correctly identify 90% and miss the other 10% (sensitivity of 90%), according to the AGA guideline. Those who present with symptoms after a false-negative screen may experience a sense of anger and betrayal that they would not suffer in the absence of a screening program.

Using the computer simulation, the AGA guideline presents data on the frequency of some of these harms. These data are summarized in Table 4 for 1000 people participating in annual screening by FOBT from 50 to 85 years of age. The model assumes those who test positive have a colonoscopy.

We now know the magnitude of both benefits and harms (as presented in Table 4). This balance sheet tells us that screening 1000 people annually with FOBT from 50 years of age will prevent 13.3 deaths from CRC, but will cause 0.5 deaths from the complications of investigation and surgery. There will also be 10.4 major complications (perforations and major bleeding episodes) and 7.7 minor complications. The authors provide no data on anxiety, but we could assume that some people will feel anxious prior to colonoscopy. Figure 2 presents these data as a flow diagram.

These data assume that the screening programs will deliver the same magnitude of benefit and harms as found in RCTs; this will be true only if the program is delivered to the same standard of quality as in the trials. Otherwise, benefits will be smaller and the harms greater.

### How Do Benefits and Harms Compare in Different People and With Different Screening Strategies?

The AGA guideline recommends that people at average risk and older than 50 years of age be offered screening for CRC. The guideline discusses several screening strategies (FOBT, flexible sigmoidoscopy, barium enema, and colonoscopy) and, in relation to FOBT, recommends offering annual screening.

The magnitude of benefits and harms will vary in different patients and under different screening strategies, as the following discussion reveals.

**Risk of Disease.** Assuming that the RRR is constant over a broad range of risk of disease, benefits will be greater for people at higher risk of disease. For example, mortality from CRC rises with age, and the mortality benefit achieved by screening rises accordingly (Figure 3, top). But the life years lost in the population to CRC are related both to the age at which mortality is highest and the length of life still available. Thus, the number of life years that can be saved by CRC screening increases with age to about 75 years and then decreases again as life expectancy declines (Figure 3, bottom). The number of deaths averted by screening over 10 years for those aged 40, 50, and 60 years at first screening (0.2, 1.0, and 2.4, respectively, per 1000 people) reflects these differences. Because of a greater benefit, it may be rational for a 60-year-old person to decide screening is worthwhile, while a 40-year-old person (or 80 years old) with smaller potential benefit might decide it is not worthwhile.

Risk of disease, and therefore benefits from screening, may be increased by other factors, such as a family history. The AGA guideline reports that people with 1 or more first-degree relatives (parent, sibling, child) with CRC, but without one of the specific genetic syndromes, have approximately twice the risk of developing CRC as average-risk individuals without a family history. This means that for people aged 40 years who have a first-degree relative with CRC, the incidence of CRC is comparable to that for people aged 50 years without a family history. The guideline also notes that within each age group, the risk is greatest in those whose relatives developed cancer at a younger age.

**Screening Interval.** As the screening interval is shortened, the effectiveness of a screening program will tend to improve, although there is a limit to the amount of improvement that is possible.

### Table 4. Clinical Consequences for 1000 People Entering a Program of Annual Fecal Occult Blood Test Screening for Colorectal Cancer at Age 50 Years and Remaining in the Program Until 85 Years of Age or Death*

<table>
<thead>
<tr>
<th>Clinical Consequences</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harms</td>
<td></td>
</tr>
<tr>
<td>Screening tests</td>
<td>27,030</td>
</tr>
<tr>
<td>Diagnostic evaluations (by colonoscopy)</td>
<td>2,263</td>
</tr>
<tr>
<td>False-positive screening tests</td>
<td>2,158</td>
</tr>
<tr>
<td>Deaths due to colonoscopy complications</td>
<td>0.5</td>
</tr>
<tr>
<td>Biopsy perforations from colonoscopy</td>
<td>3.0</td>
</tr>
<tr>
<td>Major bleeding episodes from colonoscopy</td>
<td>7.4</td>
</tr>
<tr>
<td>Minor complications from colonoscopy</td>
<td>7.7</td>
</tr>
<tr>
<td>Benefits</td>
<td></td>
</tr>
<tr>
<td>Deaths averted</td>
<td>13.3</td>
</tr>
<tr>
<td>Years of life saved</td>
<td>123.3</td>
</tr>
<tr>
<td>Years of life gained per person whose cancer death was prevented</td>
<td>9.3</td>
</tr>
</tbody>
</table>

*Adapted from Winawer et al.*

©1999 American Medical Association. All rights reserved.
sible. For example, screening twice as often could theoretically double the relative mortality reduction obtainable by screening, but in practice, the effect is usually much less. Cervical cancer screening may, for instance, reduce the incidence of invasive cervical cancer by 64%, 84%, and 94% if screening is conducted at 10-year, 5-year, and annual intervals, respectively.18

The frequency of harms will also increase with more frequent screening, potentially directly in proportion to the frequency of screening. Thus, we will see diminishing marginal return as the screening interval is shortened. Ultimately, the marginal harms will outweigh the marginal benefit of further reductions in the screening interval.

**Test Characteristics.** If the sensitivity of a new test is greater than the test used in the trials and is detecting significant disease earlier, the benefit of screening will increase. But it may be that the new, apparently more sensitive, test is detecting more cases of inconsequential disease (for example, by detecting more low-grade prostate cancer or more low-grade cervical epithelial abnormalities19), which will increase the harms. On the other hand, if specificity is improved and testing produces fewer false-positive results, net benefit will increase and the test may now be useful in groups in which the old test was not.

Ideally, clinicians would look to RCTs of the new test compared with the old test. However, new tests often appear in profusion, and randomized trials are expensive and often only interpretable after long follow-up. Being pragmatic, we will usually need to accept that the trials have shown that earlier detection works and a comparison of a new vs the old test only needs to examine test characteristics. Returning to CRC screening, since we have RCT data of mortality reduction, we may assume that earlier detection using other methods such as flexible sigmoidoscopy will also reduce mortality from CRC even though there are no published reports of RCTs of screening with flexible sigmoidoscopy.

**Figure 2. Flow Diagram of the Clinical Consequences for 1000 People Entering a Program of Annual Fecal Occult Blood Test (FOBT) Screening for Colorectal Cancer (CRC) at Age 50 Years and Remaining in the Program Until 85 Years of Age or Death**

<table>
<thead>
<tr>
<th>27,030 Annual FOBT Screens in 1000 People Aged 50 Years Until Age 85 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>2263 Colonoscopies</td>
</tr>
<tr>
<td>2158 No Cancer</td>
</tr>
<tr>
<td>18.6 Complications</td>
</tr>
<tr>
<td>0.5 Deaths</td>
</tr>
<tr>
<td>3.0 Perforations</td>
</tr>
<tr>
<td>7.4 Major Hemorrhages</td>
</tr>
<tr>
<td>7.7 Minor Hemorrhages</td>
</tr>
<tr>
<td>106 Cancers</td>
</tr>
<tr>
<td>28.7 Deaths</td>
</tr>
<tr>
<td>65.0 Usual Survivors</td>
</tr>
<tr>
<td>13.3 Extra Survivors</td>
</tr>
</tbody>
</table>

Usual survivors are those who would have survived with or without screening. Extra survivors are those in whom the earlier detection of cancer averts death. Adapted from Winawer et al.7

**What Is the Impact of People’s Values and Preferences?**

People will value benefits and harms of screening differently. For example, pregnant women who are considering screening for Down syndrome may make different choices depending on the value they place on having a Down syndrome baby vs the risk of iatrogenic abortion from amniocentesis.20

Individuals who choose to participate in screening programs are benefiting (in their view) from screening, and other individuals are benefiting (in their view) from not participating. Individuals can only make the right choice for themselves if they have access to high-quality information about the benefits and harms of screening and are able to weigh that information. This probably will require much better educational materials and decision support materials; some examples are already available.21,22

**What Is the Impact of Uncertainty Associated With the Evidence?**

There is always uncertainty about the benefits and harms of screening. The 95% CIs around the magnitude of each benefit and harm provides an indication of the amount of uncertainty in each estimate. Where sample size is limited, the CIs will be wide and clinicians should alert potential screening participants that the magnitude of the benefit or harm could be considerably smaller or greater than the point estimate.

**What Is the Cost-effectiveness?**

While clinicians will be most interested in the balance of benefits and harms for their individual patients, policymakers must consider issues of cost-effectiveness and local resources in their decisions. Clinicians can look to previous Users’ Guides to help them
evaluate studies addressing these economic issues.\textsuperscript{23,24}

The AGA guideline reports that the estimated cost-effectiveness of FOBT screening is approximately $10,000 per life year gained among people older than 50 years (although, like the absolute size of the benefit, it will vary with risk of disease). The AGA guideline also notes that all CRC screening strategies examined (FOBT, flexible sigmoidoscopy, barium enema, colonoscopy) cost less than $20,000 per life year saved.\textsuperscript{20}

These cost-effectiveness ratios are within the range of what is currently paid in some countries for the benefits of other screening programs such as mammographic screening for women aged 50 to 69 years (estimated at $39,495\textsuperscript{20}) and ultrasound screening for abdominal aortic aneurysm in men aged 60 to 80 years (estimated $41,550 per life year gained\textsuperscript{22}).

\textbf{RESOLUTION OF THE SCENARIO}

The guideline should quantify the benefit of screening according to age so you can inform your patients as accurately as possible about the benefits of screening for them. The AGA guideline does not provide age-specific mortality reductions attributable to screening; therefore, you cannot easily quantify the benefit for your patients. From the guideline, all you could say is that screening a group of 100 people with FOBT beginning at 50 years of age and continuing annually to 85 years of age will avert about 13 deaths from CRC. However, we know from the systematic review by Towler et al\textsuperscript{1} that the mortality benefit for people between 40 and 50 years of age is about 0.2 to 1.0 deaths averted over 10 years per 1000 people screened. Next you could outline the potential harms of screening. As noted earlier, the harms are mostly related to the colonoscopy. According to the AGA guideline, the risks of colonoscopy are about 0.1 to 0.3 per 1000 for death, and 1 to 3 per 1000 for perforation and hemorrhage. In addition, there would also be issues of cost, inconvenience, and anxiety.

It is up to your patients to weigh whether the benefit of reduced risk of death from CRC is worth the risks. If they feel unable to do this, then you could consider helping them to clarify their values about the possible outcomes. For example, if they are not bothered by the prospect of a colonoscopy, they would probably choose to be screened. But if either of them places a high value on avoiding colonoscopy now, he or she may prefer to reconsider screening in a few years’ time when the benefits will be greater.

\textbf{REFERENCES}