Screening for Colorectal Cancer
Updated Evidence Report and Systematic Review for the US Preventive Services Task Force

Jennifer S. Lin, MD; Margaret A. Piper, PhD; Leslie A. Perdue, MPH; Carolyn M. Rutter, PhD; Elizabeth M. Webber, MS; Elizabeth O'Connor, PhD; Ning Smith, PhD; Evelyn P. Whitlock, MD

**IMPORTANCE**
Colorectal cancer (CRC) remains a significant cause of morbidity and mortality in the United States.

**OBJECTIVE**
To systematically review the effectiveness, diagnostic accuracy, and harms of screening for CRC.

**DATA SOURCES**

**STUDY SELECTION**
English-language studies conducted in asymptomatic populations at general risk of CRC.

**DATA EXTRACTION AND SYNTHESIS**
Two reviewers independently appraised the articles and extracted relevant study data from fair- or good-quality studies. Random-effects meta-analyses were conducted.

**MAIN OUTCOMES AND MEASURES**
Colorectal cancer incidence and mortality, test accuracy in detecting CRC or adenomas, and serious adverse events.

**RESULTS**
Four pragmatic randomized clinical trials (RCTs) evaluating 1-time or 2-time flexible sigmoidoscopy (n = 458 002) were associated with decreased CRC-specific mortality compared with no screening (incidence rate ratio, 0.73; 95% CI, 0.66-0.82). Five RCTs with multiple rounds of biennial screening with guaiac-based fecal occult blood testing (n = 419 966) showed reduced CRC-specific mortality (relative risk [RR], 0.91; 95% CI, 0.84-0.98, at 19.5 years to RR, 0.78; 95% CI, 0.65-0.93, at 30 years). Seven studies of computed tomographic colonography (CTC) with bowel preparation demonstrated per-person sensitivity and specificity to detect adenomas 6 mm and larger comparable with colonoscopy (sensitivity from 73% [95% CI, 58%-84%] to 98% [95% CI, 91%-100%]; specificity from 89% [95% CI, 84%-93%] to 91% [95% CI, 88%-93%]); variability and imprecision may be due to differences in study designs or CTC protocols. Sensitivity of colonoscopy to detect adenomas 6 mm or larger ranged from 75% (95% CI, 63%-84%) to 93% (95% CI, 88%-96%). On the basis of a single stool specimen, the most commonly evaluated families of fecal immunochemical tests (FITs) demonstrated good sensitivity (range, 73%-88%) and specificity (range, 90%-96%). One study (n = 9989) found that FIT plus stool DNA test had better sensitivity in detecting CRC than FIT alone (92%) but lower specificity (84%). Serious adverse events from colonoscopy in asymptomatic persons included perforations (4/10 000 procedures, 95% CI, 2.5 in 10 000) and major bleeds (8/10 000 procedures, 95% CI, 5.14 in 10 000). Computed tomographic colonography may have harms resulting from low-dose ionizing radiation exposure or identification of extracolonic findings.

**CONCLUSIONS AND RELEVANCE**
Colonoscopy, flexible sigmoidoscopy, CTC, and stool tests have differing levels of evidence to support their use, ability to detect cancer and precursor lesions, and risk of serious adverse events in average-risk adults. Although CRC screening has a large body of supporting evidence, additional research is still needed.

Published online June 15, 2016. Last corrected on October 4, 2016.
Although colorectal cancer (CRC) incidence has been declining over the past 20 years in the United States, it still causes significant morbidity and mortality. Despite increases in screening rates over the past 30 years, in 2012 an estimated 28% of eligible US adults had never been screened for CRC. A variety of tests are available for screening, including stool-based tests (eg, guaiac-based fecal occult blood testing [gFOBT], immunochemical-based fecal occult blood testing [FIT], stool DNA [sDNA] testing), endoscopy (eg, flexible sigmoidoscopy [SIG], colonoscopy), and imaging (eg, double-contrast barium enema, computed tomographic colonography [CTC]).

Currently, most US guideline organizations, including the US Preventive Services Task Force (USPSTF), recommend that options for CRC screening include colonoscopy every 10 years, an annual high-sensitivity gFOBT or FIT, and SIG every 5 years with high-sensitivity gFOBT or FIT. In 2008, the USPSTF recommended CRC screening using fecal occult blood testing, sigmoidoscopy, or colonoscopy beginning at age 50 years and continuing until age 75 years (A recommendation); selectively offering screening in adults aged 76 to 85 years (C recommendation); and against screening for colorectal cancer in adults older than 85 years (D recommendation). At that time, the USPSTF had insufficient evidence to assess the benefits and harms of CTC and sDNA testing as screening modalities. A systematic review was conducted to update relevant evidence since 2008 and to help inform a separate modeling exercise, which together were used by the USPSTF in its process of updating the 2008 CRC screening recommendations.

**Methods**

**Scope of Review**

This review addressed 3 key questions (KQs) as shown in Figure 1. Additional methodological details regarding search strategies, detailed study inclusion criteria, quality assessment, excluded studies, and description of data analyses are publicly available in the full evidence report at http://www.uspreventiveservicestaskforce.org/Page/Document/final-evidence-review/colorectal-cancer-screening.
Data Sources and Searches
MEDLINE, PubMed, and the Cochrane Central Register of Controlled Trials were searched to locate primary studies informing the key questions (Methods in the Supplement) that were published from the end of the previous review6 (January 1, 2008) through December 31, 2014. The database searches were supplemented with expert suggestions and by reviewing reference lists from all other relevant systematic reviews, including the 2008 USPSTF evidence report. The search also included selected gray literature sources, including ClinicalTrials.gov and the World Health Organization International Clinical Trials Registry Platform, for ongoing trials. Since December 2014, we continued to conduct ongoing surveillance through article alerts and targeted searches of high-impact journals to identify major studies published in the interim that may affect the conclusions or understanding of the evidence and therefore the related USPSTF recommendation. The last surveillance was conducted on February 23, 2016. Although several potentially relevant new studies were identified, none of these studies would substantively change the review’s interpretation of findings or conclusions.

Study Selection
Two investigators independently reviewed 8492 titles and abstracts and 696 articles against the specified inclusion criteria (Figure 2). Discrepancies were resolved through consensus and consultation with a third investigator. Inclusion criteria were fair- and good-quality English-language studies of asymptomatic screening population of individuals who were 40 years or older, either at average risk for CRC or not selected for inclusion based on CRC risk factors. Studies were included that evaluated the following screening tests: colonoscopy, SIG, fOBT, FIT, FIT plus sDNA, or a blood test for methylated SEPT9 DNA (mSEPT9).

For KQ1, randomized clinical trials (RCTs) or otherwise controlled trials of CRC screening vs no screening, as well as trials comparing screening tests, that included outcomes of cancer incidence, CRC-specific mortality, or all-cause mortality were reviewed for inclusion. For tests without trial-level evidence (ie, colonoscopy, FIT), well-conducted prospective cohort or population-based nested case-control studies were examined.

For KQ2, diagnostic accuracy studies that used colonoscopy as a reference standard were included. Studies whose design was sub-
Evidence Report: Screening for Colorectal Cancer

US Preventive Services Task Force Clinical Review & Education

For KQ3, all trials and observational studies that reported serious adverse events requiring unexpected or unwanted medical attention or resulting in death were included. These events included, but were not limited to, perforation, major bleeding, severe abdominal symptoms, and cardiovascular events. Studies designed to assess for extracolonic findings (ie, incidental findings on CTC) and the resultant diagnostic yield and harms of workup were also included. Studies reporting extracolonic findings generally used the CT Colono-graphy Reporting and Data System (C-RADS). Under C-RADS, extracolonic findings are categorized as E0 (limited examination), E1 (normal examination or normal variant), E2 (clinically unimportant finding in which no workup is required), E3 (likely unimportant or incompletely characterized in which workup may be required), or E4 (potentially important finding requiring follow-up).15

Data Extraction and Quality Assessment
Two reviewers each critically appraised all articles that met inclusion criteria using the USPSTF design-specific quality criteria16 supplemented by the National Institute for Health and Clinical Excellence methodology checklists.17 A Measurement Tool to Assess Systematic Reviews (AMSTAR) for systematic reviews,18 New-castle Ottawa Scales for cohort and case-control studies.19 and Quality Assessment of Diagnostic Accuracy (QUADAS) and QUADAS-2 for studies of diagnostic accuracy (eTable 1 in the Supplement).20,21 Poor-quality studies and those with a single fatal flaw or multiple important limitations that could invalidate results were excluded from this review. Disagreements about critical appraisal were resolved by consensus and, if needed, consultation with a third independent reviewer. One reviewer extracted key data from included studies; a second reviewer checked the data for accuracy.

Data Synthesis and Analysis
For each KQ, the number and design of included studies, overall results, consistency or precision of results, reporting bias, study quality, limitations of the body of evidence, and applicability of findings were summarized. The results were synthesized by KQ, type of screening test, and study design. Studies from the 2008 review that met the updated inclusion criteria were incorporated. The analyses for test performance focused primarily on per-person (ie, by individual patient rather than by lesion) test sensitivity and specificity to detect adenomas (by size, where reported, <6 mm, ≥6 mm, ≥10 mm), advanced adenomas (as defined by the study), and CRC. The studies used several kinds of FITs, which were grouped as qualitative (fixed cutoff) or quantitative (adjustable cutoff), as well as into families (tests produced by the same manufacturer, using the same components and method, or compatible with different automated analyzers). Tests were compared using similar cutoff values expressed in μg hemoglobin (Hb)/g feces.

Because of the limited number of studies and the clinical heterogeneity of studies, the analyses were largely descriptive. Random-effects meta-analyses were conducted using the profile likelihood method22 to estimate the effect of SIG based on the pooled incidence rate ratio (events/person-year) for CRC incidence and mortality across the 4 major SIG trials. Random-effects models were also conducted using the restricted maximum likelihood estimation method to estimate rates of serious adverse events for colonoscopy and SIG. The presence and magnitude of statistical heterogeneity were assessed among pooled studies using the I² statistic. All tests were 2-sided with a P value less than .05 indicating statistical significance. Meta-analyses were performed using R version 3.0.2 (R Project for Statistical Computing).23,24

Results

Effectiveness of Screening
Key Question 1. What is the effectiveness of screening programs based on the prespecified screening tests (alone or in combination) in reducing incidence of and mortality from colorectal cancer?

Twenty-five unique fair- or good-quality studies25-49 (published in 47 articles25-71) were found that assessed the effectiveness or comparative effectiveness of screening tests on CRC incidence and mortality. These studies included 1 cohort study of screening colonoscopy,36,4 4 RCTs of SIG (in 7 articles),35,39,41,50,60,66,71 and 6 trials (in 11 articles) of Hemoccult II gFOBT (References 29, 33-35, 40, 44, 59, 62-64, 67). In addition, 15 comparative effectiveness studies (in 22 articles) were found that were primarily designed to assess the relative uptake and CRC yield between different screening modalities (References 26-28, 30-32, 37, 38, 42, 43, 45-49, 54-58, 65, 69). Due to limitations in study designs, the observational colonoscopy study and comparative effectiveness studies are not discussed further in this article. Summarized below are the results for CRC-specific mortality, as results for CRC incidence were consistent with CRC mortality findings.

Flexible Sigmoideoscopy
Four large, fair-quality, pragmatic RCTs (n = 458 002) evaluated the effectiveness of 1 or 2 rounds of SIG in average-risk adults aged 50 to 74 years (Table I).25,39,41,50,60,66,71 Adherence to SIG in these trials ranged from 58% to 84%, and rates of diagnostic colonoscopy ranged from 5% to 33% due to differences in referral criteria. Based on pooled intention-to-treat analyses, SIG was associated with lower CRC-specific mortality compared with no screening at 11 to 12 years of follow-up (incidence rate ratio, 0.73; 95% CI, 0.66-0.82; I² = 0%) (Figure 3); however, the association with mortality benefit was limited to distal CRC (incidence rate ratio, 0.63; 95% CI, 0.49-0.84; I² = 44%) (eFigure 1 in the Supplement). In 1 trial, conducted in Norway, half of the participants randomized to SIG also received a single FIT test; the SIG-plus-FIT group had lower CRC mortality than the SIG-only group did (hazard ratio, 0.62; 95% CI, 0.42-0.90).60

gFOBT
Five older, large, pragmatic RCTs (n = 419 966) with 11 to 30 years of follow-up evaluated the effectiveness of annual or biennial screening programs with Hemoccult II (Table I) (References 29, 33, 34, 40, 44, 59, 63, 64, 67). Based on intention-to-treat analyses, compared with no screening, biennial screening with Hemoccult II resulted in a reduction in CRC-specific mortality after 2 to 9 rounds of
screening (relative risk [RR], 0.91; 95% CI, 0.84-0.98, at 19.5 years to RR, 0.78; 95% CI, 0.65-0.93, at 30 years). Based on 1 trial, conducted in the United States, annual screening with Hemoccult II after 11 rounds of screening resulted in greater reductions (RR, 0.68; 95% CI, 0.56-0.82) than biennial screening at 30 years did (RR, 0.78; 95% CI, 0.65-0.93).44

Diagnostic Accuracy of Screening

Key Question 2. What are the test performance characteristics of the prespecified screening tests (alone or in combination) for detecting colorectal cancer, advanced adenomas, or adenomatous polyps based on size?

Thirty-three unique diagnostic accuracy studies72-104 (published in 44 articles72-115) were found that evaluated the 1-time test performance of a screening test compared with an adequate reference standard, including 9 studies of screening CTC (in 10 articles), (References 81, 82, 85-87, 89, 93, 99, 101, 114) 3 studies of gFOBT Hemoccult Sensa,72,73,90 20 studies of various FITs (References 72- 78, 80, 82-84, 88, 90, 91, 94-98, 100, 102-104) (1 of which evaluated a FIT plus sDNA test83), and 1 study of a blood test to detect

Table 1. Effectiveness of Screening to Reduce Colorectal Cancer Mortality: Flexible Sigmoidoscopy and Hemoccult II RCTs (Key Question 1)*

<table>
<thead>
<tr>
<th>Screening Tool and Reference</th>
<th>Qualityb</th>
<th>Country</th>
<th>Patient Age Range, y</th>
<th>No. of Participants</th>
<th>No. of Screening Rounds</th>
<th>Screening Interval, y</th>
<th>Follow-up Period, y</th>
<th>Positive Screening Results, %c</th>
<th>No. of CRC Deaths/100 000 Person-Years</th>
<th>CRC Mortality, RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flexible Sigmoidoscopy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NORCCAP,60 2014</td>
<td>Fair</td>
<td>Norway</td>
<td>50-64</td>
<td>Intervention: 20 572 Control: 78 220</td>
<td>1</td>
<td>NA</td>
<td>11.0</td>
<td>20.4</td>
<td>1.4</td>
<td>Intervention: 31</td>
</tr>
<tr>
<td>PLCO,59,71 2012</td>
<td>Fair</td>
<td>United States</td>
<td>55-74</td>
<td>Intervention: 77 445 Control: 77 455</td>
<td>2</td>
<td>3-5</td>
<td>12.1</td>
<td>32.9</td>
<td>1.5</td>
<td>Intervention: 29</td>
</tr>
<tr>
<td>SCORE41,66 2011</td>
<td>Fair</td>
<td>Italy</td>
<td>55-64</td>
<td>Intervention: 17 136 Control: 17 136</td>
<td>1</td>
<td>NA</td>
<td>11.4</td>
<td>8.6</td>
<td>1.6</td>
<td>Intervention: 35</td>
</tr>
<tr>
<td>UKFSS,75,50 2010</td>
<td>Fair</td>
<td>United Kingdom</td>
<td>55-64</td>
<td>Intervention: 57 099 Control: 112 939</td>
<td>1</td>
<td>NA</td>
<td>11.2</td>
<td>5.2</td>
<td>1.5</td>
<td>Intervention: 30</td>
</tr>
<tr>
<td>Hemoccult II*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minnesota Colon Cancer Control Study,2013</td>
<td>Good</td>
<td>United States</td>
<td>50-80</td>
<td>Intervention (biennial): 15 587 Control: 15 394</td>
<td>6</td>
<td>2</td>
<td>30</td>
<td>NR†</td>
<td>2.9†</td>
<td>Intervention: 50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>1</td>
<td>30</td>
<td>NR†</td>
<td>2.9†</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CRC, colorectal cancer; NA, not applicable; NR, not reported; RCT, randomized clinical trial; RR, relative risk.

* The comparator for each of these RCTs was a control group that was not offered any CRC screening.

† Assessed using criteria from the US Preventive Services Task Force.26

‡ Median follow-up time for flexible sigmoidoscopy, longest follow-up time for Hemoccult II.

§ For flexible sigmoidoscopy, this refers to the percentage of patients who were referred to colonoscopy out of those who received their flexible sigmoidoscopy. For Hemoccult II, it refers to the percentage of patients who tested positive out of those who took the test in round 1 only.

* The percentage of CRC cases that occurred during the follow-up period among those included in the study at baseline.

† Calculated RR (not study reported).

§ NORCCAP reported a statistically significant decrease in CRC mortality for the screened group vs the control (hazard ratio, 0.73; 95% CI, 0.56-0.94; P = .02). To present consistent results across studies, we show unadjusted results here.

One trial in Finland has not reported CRC mortality.95-97

† Study included rehydrated tests: in Göteborg, 91.7% of all tests were rehydrated; in the Minnesota Colon Cancer Control Study, 82.5% of all tests were rehydrated.

† Refers to all 3 groups of the trial (annual, biennial, and control).
circulating mSEPT9. The study of mSEPT9 (not approved by the US Food and Drug Administration [FDA] for screening) and studies evaluating Hemoccult Sensa and FITs that only applied the colonoscopy reference standard to positive stool tests are not discussed further in this article.

**Direct Visualization Tests**

Nine fair- or good-quality studies (n = 6497) evaluated the diagnostic accuracy of multidetector CTC in average-risk screening populations (Table 2) (References 81, 82, 85-87, 89, 93, 99, 101, 114). The 2 largest and best-quality studies were multicenter trials.

---

**Table 2. Prospective Diagnostic Accuracy Studies of Screening Computed Tomographic Colonography (Key Question 2)**

| Study | Quality* | Site | Cohort Size | Mean Patient Age, y | Fecal Tag* | No. of Readers, Training* | Reading Strategy* | Reference Standard | Adenoma ≥6 mm, % (95% CI) | Adenoma ≥10 mm, % (95% CI) |
|-------|----------|-----|-------------|---------------------|-----------|-------------------------|------------------|-------------------|----------------|----------------|----------------|
| **With Bowel Preparation** | | | | | | | | | | |
| Lefere et al, 2013 | Fair | Portugal | 496 | 60 | Yes | 1, >5000 exams | 3D (with 2D) | Repeat colonoscopy if indicated | 98 (91-100) | 91 (88-93) | NR | NR |
| Grasser et al, 2009 | Good | Germany | 307 | 60 | No | 3, >300 exams | 3D (with 2D) | Repeat colonoscopy, segmental unblinding* | 91 (80-97) | 93 (90-96) | 92 (76-98) | 98 (96-99) |
| Johnson et al, 2008 | Good | United States | 2531 | 58 | Yes | 15, >500 exams | 3D (with 2D) | Repeat colonoscopy if indicated | 78 (72-83) | 90 (88-91) | 90 (83-95) | 86 (85-87) |
| Kim et al, 2008 | Fair | South Korea | 241 | 58 | No | 2, >100 exams | 2D (with 3D) | Single colonoscopy | 68 (55-80) | 89 (84-93) | 87 (64-97) | 97 (95-99) |
| Johnson et al, 2007 | Fair | United States | 452 | 65 | No | 3, >1000 exams | 3D (with 2D) | Single colonoscopy | NR | NR | 67 (45-84) | 98 (96-99) |
| Macari et al, 2004 | Fair | United States | 68 | 55 | No | 1, 5 y | NR | Single colonoscopy | NR | NR | 100 (46-100) | 98 (93-100) |
| Pickhardt et al, 2003 | Good | United States | 1233 | 58 | Yes | 6, >25 exams | 3D (with 2D) | Colorectal cancer | 89 (83-93) | 80 (77-82) | 94 (84-98) | 96 (95-97) |
| **Without Bowel Preparation** | | | | | | | | | | |
| Fletcher et al, 2013 | Good | United States | 564 | NR | Yes | 2, >150 exams | 2D and 3D | Single colonoscopy | 75 (59-87) | 92 (90-94) | 67 (42-86) | 97 (96-98) |
| Zalis et al, 2012 | Good | United States | 605 | 60 | Yes | 3, >200 exams | 2D and 3D | Colorectal cancer | 58 (46-69) | 88 (85-91) | 90 (70-98) | 85 (82-88) |

Abbreviations: CT, computed tomographic; exams, examinations; NR, not reported.

* Quality assessed using criteria from Quality Assessment of Diagnostic Accuracy Studies (QUADAS)(29) and QUADAS 2(2) instrument.

* Oral ingestion of high-density oral contrast agent so that residual colonic contents can be differentiated from soft tissue density polypos.

* Number of examinations or years of training required by each reader or radiologist.

* Reader or radiologist procedure for using 2- and 3-dimensional images.

* CT colonography enhanced colonoscopy, in which endoscopist was shown respective segment of colon on CT colonography after examination of segment by colonoscopy.

f National CT Colonography Trial.

* Any histology ≥6 mm; sensitivity for adenomas =6 mm, 72.7% (95% CI, 58.4%-84.1%); specificity not reported.

* Any histology ≥10 mm; sensitivity for adenomas ≥10 mm, 90.0% (95% CI, 61.9%-99.0%); specificity not reported.

i Study evaluated different reading strategies; data shown reflect primary 3D strategy.

j Any histology ≥10 mm.
conducted in the United States evaluating CTC with bowel preparation and fecal tagging.85,99 Overall, the studies were not powered to estimate test performance to detect CRC. Based on 7 studies of CTC with bowel preparation (n = 5328), the per-person sensitivity to detect adenomas 10 mm and larger ranged from 67% (95% CI, 45%-84%) to 94% (95% CI, 84%-98%), and specificity ranged from 98% (95% CI, 96%-99%) to 96% (95% CI, 95%-97%). The per-person sensitivity to detect adenomas 6 mm and larger ranged from 73% (95% CI, 58%-84%) to 98% (95% CI, 91%-100%), and specificity ranged from 89% (95% CI, 84%-93%) to 91% (95% CI, 88%-93%). Two studies (N = 169) evaluated CTC without bowel preparation.83,101 Although the data were limited, the sensitivity of CTC without bowel preparation to detect adenomas 6 mm and larger appeared to be lower than the sensitivity of CTC protocols including bowel preparation.

Four (n = 4821) of the 9 CTC studies allowed for the estimation of sensitivity of colonoscopy generalizable to community practice.85,86,99,101 Compared with CTC or colonoscopy plus CTC (eg, segmental unblinding), the sensitivity for colonoscopy to detect adenomas 10 mm and larger ranged from 89% (95% CI, 78%-96%) to 98% (95% CI, 74%-100%) and for adenomas 6 mm and larger ranged from 75% (95% CI, 63%-84%) to 93% (95% CI, 88%-96%) (see full report). Therefore, CTC with bowel preparation had sensitivity to detect adenomas 6 mm and larger comparable with colonoscopy, albeit with wider variability in estimated performance. It is unclear whether the observed variation in CTC performance was due to differences in study design, populations, bowel preparation, CTC technologies, or differences in reader experience or reading protocols.

Stool Tests

Fourteen fair- or good-quality studies (n = 59425) that used colonoscopy reference standard in all participants reported sensitivity and specificity for 19 different types of qualitative or quantitative FITs, including 1 FIT plus sDNA test (Table 3) (References 74, 77, 78, 80, 82, 83, 91, 94, 95, 97, 98, 100, 103, 104, 107, 108, 115). Overall, the sensitivity for CRC and advanced adenomas varied widely, including a discontinued test with very low sensitivity.100 Given the heterogeneity among FITs and their test performance, focus was placed on the performance characteristics of currently available tests evaluated in more than 1 study. Two families of FDA-cleared tests, OC-Light (qualitative, No. of studies = 3, n = 25924) and OC FIT-CHEK (eg, OC Sensor Diana, OC-Micro, OC-Auto) (quantitative, No. of studies = 5, n = 12794), had relatively high sensitivity and specificity. With a single stool specimen, the lowest sensitivity demonstrated for CRC was 73% (95% CI, 48%-90%) and specificity was 96% (95% CI, 95%-96%). Similarly, the highest sensitivity with paired specificity for CRC was 88% (95% CI, 55%-99%) and 91% (95% CI, 89%-92%), respectively. In the largest studies, sensitivity ranged from 74% (95% CI, 62%-83%) for quantitative test categories (n = 9989) to 79% (95% CI, 61%-90%) for qualitative test categories (n = 18296). In a small study (n = 770) that tested 3 stool specimens, sensitivity was 92% (95% CI, 69%-99%), but specificity was 87% (95% CI, 85%-89%). OC-Light or OC FIT-CHEK test sensitivity and specificity for advanced adenomas ranged from 22% (95% CI, 17%-28%) to 40% (95% CI, 30%-51%), and specificity ranged from 97% (95% CI, 97%-98%) to 91% (95% CI, 91%-92%). Although higher sensitivities to detect advanced adenomas were obtained for certain other FITs or by using 3 stool specimens, the corresponding specificities were lower.

Cologuard (Exact Sciences) is an FDA-approved stool test that combines stool DNA with a proprietary FIT component. One fair-quality diagnostic accuracy study (n = 9989) evaluated Cologuard compared with OC FIT-CHEK.83 In that study, Cologuard had a statistically significant higher sensitivity to detect CRC and advanced adenoma compared with OC FIT-CHEK. The higher sensitivity for CRC (92% [95% CI, 84%-97%]) and for advanced adenoma (42%; 95% CI, 39%-46%) was accompanied by lower specificity (84%; 95% CI, 84%-85% for CRC and 87%; 95% CI, 86%-87% for advanced adenoma). In our active surveillance of the literature, we identified 1 additional diagnostic accuracy study of FIT plus sDNA (n = 661) in asymptomatic Alaska Native adults.9 This study was not powered to find a difference in detection of CRC; nonetheless, findings were generally consistent with the included study on FIT plus sDNA.83

Harms of Screening

Key Question 3a. What are the adverse effects of the different screening tests (either as single application or in a screening program)?

Key Question 3b. Do adverse effects vary by important subpopulations (eg, age)?

Ninety-eight fair- or good-quality studies (References 27, 29, 37-39, 45, 48, 64, 66, 77, 78, 81, 82, 85-87, 89, 92, 93, 97, 99, 101, 116-119) in 113 articles (References 27, 29, 34, 37-39, 44, 45, 48, 50, 53, 64, 66, 70, 77, 78, 81, 82, 85-87, 89, 92, 93, 97, 99, 101, 114, 116-200) were included that evaluated the harms of CRC screening. These studies included 14 studies of screening programs using stool testing or SIG, 55 studies of colonoscopy in asymptomatic adults, (References 37, 45, 77, 78, 85, 97, 101, 116, 117, 119-124, 126, 128-131, 133, 136, 140, 142, 144, 147, 148, 150, 151, 153-156, 158, 159, 161-163, 170-178, 180-183, 187-190) 18 studies of screening SIG, (References 27, 38, 39, 43, 48, 50, 66, 143, 146, 151, 157, 162, 176, 183, 185, 186, 191-194, 200) and 15 studies of screening CTC in asymptomatic adults (References 45, 81, 82, 85, 87, 89, 101, 118, 135, 145, 150, 162, 169, 179). Twelve CTC studies provided estimates of radiation exposure per examination, (References 81, 82, 85-87, 89, 93, 99, 101, 118, 135, 162) and another 21 CTC studies reported information on extracolonic findings (References 45, 85, 89, 101, 114, 125, 127, 134, 137-139, 141, 150, 152, 160, 164, 166-168, 184, 195, 198).

Endoscopy Harms

Approximately half of colonoscopy harms studies (29/55 studies) were in explicitly screening or asymptomatic populations (eTable 2 in the Supplement). By pooling 26 studies (n = 3 414108) in screening populations or generally asymptomatic persons, (References 37, 45, 77, 78, 85, 97, 101, 120, 121, 124, 126, 130, 131, 136, 150, 156, 163, 170, 174, 176, 180-182, 188-190) it was estimated that the risk of perforations from colonoscopy was 4 in 10 000 procedures (95% CI, 2.5-6 in 10 000; I2 = 86%) (Figure 4). On the basis of 22 of those studies (n = 3 347101), (References 37, 45, 77, 85, 97, 101, 120, 121, 124, 126, 130, 131, 156, 163, 170, 174, 180-182, 188-190) it was estimated that the risk of major bleeding from colonoscopy was 8 in 10 000 procedures (95% CI, 5.4-14 in 10 000; I2 = 97%) (Figure 5). Only eight studies (n = 204614) explicitly reported if perforation or major bleeding was related to polypectomy or...
Table 3. Prospective Diagnostic Accuracy Studies of FIT Tests (With or Without Stool DNA Test) Using Colonoscopy Reference Standard (Key Question 2)

<table>
<thead>
<tr>
<th>Study</th>
<th>Quality*</th>
<th>Mean Patient Age, y</th>
<th>Cohort Size</th>
<th>Test Family Name</th>
<th>Cutoff, μg Hb/g Feces</th>
<th>No. of Stool Samples per Person</th>
<th>CRC, %</th>
<th>Advanced Adenomas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sensitivity, % (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sensitivity, % (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qualitative FIT Tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Levy et al,91 2014b</td>
<td>Fair</td>
<td>56.9</td>
<td>308</td>
<td>Clearview (cassette)</td>
<td>6</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Chiu et al,78 2013</td>
<td>Good</td>
<td>18.296</td>
<td>10</td>
<td>OC-Light</td>
<td>0.15</td>
<td>54 (32-74)</td>
<td>89 (88-90)</td>
<td>37 (30-44)</td>
</tr>
<tr>
<td>Ng et al,97 2013</td>
<td>Fair</td>
<td>57.7</td>
<td>4539</td>
<td>Hemosure</td>
<td>0.48</td>
<td>54 (32-74)</td>
<td>89 (88-90)</td>
<td>37 (30-44)</td>
</tr>
<tr>
<td>Brenner et al,107 2010</td>
<td>Good</td>
<td>63c</td>
<td>1319</td>
<td>Bionexia Hb-Hp</td>
<td>0.8</td>
<td>52 (44-61)</td>
<td>80 (77-82)</td>
<td>52 (44-61)</td>
</tr>
<tr>
<td>Cheng et al,77 2002</td>
<td>Fair</td>
<td>46.8</td>
<td>7411</td>
<td>OC-Light</td>
<td>0.22</td>
<td>88 (66-97)</td>
<td>91 (90-92)</td>
<td>40 (30-51)</td>
</tr>
<tr>
<td>Nakama et al,95 1999</td>
<td>Fair</td>
<td>NR</td>
<td>4611</td>
<td>Monohaem</td>
<td>NR</td>
<td>86 (57-98)</td>
<td>94 (93-95)</td>
<td>NR</td>
</tr>
<tr>
<td>Quantitative FIT Tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hernandez et al,103 2014</td>
<td>Good</td>
<td>57.6</td>
<td>779</td>
<td>OC FIT-CHEK</td>
<td>0.6</td>
<td>100 (62-100)</td>
<td>92 (90-94)</td>
<td>NR</td>
</tr>
<tr>
<td>Imperiale et al,83 2014</td>
<td>Fair</td>
<td>64.2</td>
<td>9989</td>
<td>OC FIT-CHEK</td>
<td>0.65</td>
<td>74 (62-83)</td>
<td>93 (93-94)</td>
<td>24 (21-27)</td>
</tr>
<tr>
<td>Lee et al,104 2014</td>
<td>Good</td>
<td>58c</td>
<td>NR</td>
<td>Hemo Tech NS-Plus C system</td>
<td>6.3</td>
<td>86 (57-98)</td>
<td>94 (93-95)</td>
<td>NR</td>
</tr>
<tr>
<td>Brenner and Tao,74 2013</td>
<td>Good</td>
<td>62.7</td>
<td>2220</td>
<td>OC FIT-CHEK</td>
<td>0.67</td>
<td>73 (48-90)</td>
<td>96 (95-96)</td>
<td>22 (17-28)</td>
</tr>
<tr>
<td>de Wijkerslooth et al80 2012</td>
<td>Good</td>
<td>60c</td>
<td>1256</td>
<td>OC FIT-CHEK</td>
<td>1.7</td>
<td>92 (69-99)</td>
<td>87 (85-89)</td>
<td>44 (32-57)</td>
</tr>
<tr>
<td>Park et al,98 2010</td>
<td>Fair</td>
<td>59.3</td>
<td>770</td>
<td>OC FIT-CHEK</td>
<td>1.7</td>
<td>92 (69-99)</td>
<td>87 (85-89)</td>
<td>44 (32-57)</td>
</tr>
<tr>
<td>Graser et al,82 2009</td>
<td>Good</td>
<td>60.5</td>
<td>285</td>
<td>FOB Gold</td>
<td>0.33</td>
<td>100 (87.8-92.1)</td>
<td>89 (88-90)</td>
<td>32 (22.8-46.5)</td>
</tr>
<tr>
<td>Morikawa et al,99 2005</td>
<td>Fair</td>
<td>48</td>
<td>21805</td>
<td>Magstream/HemeSelect</td>
<td>0.4</td>
<td>65.8 (54.9-75.6)</td>
<td>94.6 (94.3-94.9)</td>
<td>NR</td>
</tr>
<tr>
<td>Sohn et al,100 2005</td>
<td>Fair</td>
<td>48.9</td>
<td>3794</td>
<td>OC Hemodia</td>
<td>0.3</td>
<td>25.0</td>
<td>NR</td>
<td>6.0</td>
</tr>
</tbody>
</table>

Abbreviations: CRC, colorectal cancer; FIT, fecal immunochemical test; Hb, hemoglobin; NA, not applicable; NR, not reported.

*Quality assessed using criteria from Quality Assessment of Diagnostic Accuracy Studies (QUADAS)20 and QUADAS2 21 instrument.

bResults reported for advanced neoplasia (composite of CRC and advanced adenoma) only.

c Median.
biopsy (References 45, 85, 120, 136, 158, 173, 178). Based on this limited subset of studies reporting adequate information, 36% (15/42) of perforations and 96% (49/51) of major bleeding events were from polypectomy.

All 18 SIG harms studies were conducted in general-risk screening populations (eTable 3 in the Supplement). Based on the results of 16 studies (n = 329,698) ( References 38, 39, 43, 48, 50, 66, 143, 146, 151, 157, 176, 183, 185, 186, 191, 192 ) perforations from SIG in average-risk screening populations were relatively uncommon: the pooled point estimate was 1 in 10,000 procedures (95% CI, 0.4-1.4 in 10,000; $I^2$ = 18.4%). In 10 studies (n = 137,987), ( References 27, 38, 48, 50, 66, 143, 146, 157, 185, 186 ) major bleeding episodes from SIG were also relatively uncommon, with a pooled point estimate of 2 in 10,000 procedures (95% CI, 0.4-6.2 in 10,000; $I^2$ = 43.4%).

Nineteen SIG studies examined differential harms of colonoscopy by age group ( References 116, 119, 122, 123, 128, 129, 131, 136, 140, 154, 156, 159, 161, 170, 172, 174, 187, 190 ). These studies generally found increasing rates of serious adverse events with increasing age, including perforation and bleeding.

CTC Harms

Fifteen fair- or good-quality studies addressed serious adverse effects of screening CTC (eTable 4 in the Supplement) ( References 45, 81, 82, 85, 87, 89, 101, 115, 145, 150, 162, 169, 179 ). Evidence suggested little to no risk of serious adverse events, including perforation, from CTC based on 11 prospective studies (n = 102,722) performed in screening populations ( References 45, 81, 82, 85, 87, 89, 101, 115, 145, 150, 162, 169, 179 ).

CTC Harms

There were no episodes of serious bleeding or perforation in the study.
ditional recent CTC screening studies (2004-2008), the estimated radiation dose decreased to a range of 1 to less than 5 mSv.

CTC Extracolonic Findings

Incidental extracolonic findings detected on CTC can be beneficial or harmful depending on the finding. Twenty-one studies (n = 38293) in 22 articles (7 studies with overlapping populations reported different types extracolonic findings) reported on extracolonic findings in asymptomatic persons (eTable 5 in the Supplement). In general, these studies varied greatly in their ability to accurately assess follow-up and the duration of follow-up.

Overall, extracolonic findings were common, occurring in 27% to 69% of examinations. Similarly, the studies suggested a very wide range of findings needing additional workup: 5% to 37% had E3 or E4 findings, and 1.7% to 12% had E4 findings. Among the studies that also reported medical follow-up of extracolonic findings, 1.4% to 11% went on to diagnostic evaluation, which is similar to the prevalence of E4 category findings. Among studies that adequately reported subsequent treatment, only up to 3% required definitive medical or surgical treatment. Extracolonic cancers were not common, occurring in 0.5% of persons undergoing CTC examinations. In the largest series of examinations (n = 10 286), which had about 4 years of follow-up, 0.35% of examinations revealed an extracolonic malignancy, 32 of which received definitive treatment. Abdominal aortic aneurysms were identified in 1.4% of persons or fewer. In our active surveillance of the literature, we identified 1 additional study evaluating extracolonic findings in screening CTC (n = 7952). This study's population overlapped with several already included studies and reported that 2.5% of examinations had E4 category findings, consistent with findings from included studies.

Discussion

Colorectal cancer screening continues to be a necessary and active field of research. Since the 2008 USPSTF recommendation was published, 95 new studies were identified, including more evidence on (1) the effectiveness of SIG for reducing CRC mortality, (2) the test performance of screening CTC and decreasing radiation exposure from CTC, and (3) the test performance of a number of FDA-approved FITs (including 1FIT plus sDNA test). Colonoscopy, SIG, CTC, and stool testing (gFOBT, FIT, and FIT plus sDNA test) each have differing levels of evidence to support their use, ability to detect cancer and precursor lesions, and risk of serious adverse events in screening average-risk adults for CRC (Table 4).

To date, no CRC screening modality has been shown to reduce all-cause mortality. Robust data from well-conducted population-based screening RCTs have demonstrated that both Hemoccult II and SIG can reduce CRC mortality, although neither of these tests is widely used for screening in the United States. Therefore, the empirical data on the performance of CRC screening programs using modalities used in clinical practice today are limited. Expen-
sive, large population-based RCTs of newer stool tests may not always be necessary, as evidence-based reasoning supports that screening with stool tests with sensitivity and specificity that are both as good as, or better than, Hemoccult II would result in CRC mortality reductions similar to better than reductions shown in existing trials. Based on this review, stool tests that meet those requirements are available, including specific single-stool sample FITs. However, FITs are not homogenous: they use different assays and have different diagnostic performance levels. The FDA-approved OC-Light and OC FIT-CHEK tests have the most evidence to support their use. Stool tests that maximize sensitivity (eg, FIT plus sDNA test, multiple sample FITs, or quantitative FIT using lower cutoffs) have lower specificity and therefore need new trials or modeling exercises to understand the tradeoff of higher false-positive findings. In addition, stool tests vary in cost; for example, the Centers for Medicare & Medicaid Services reimbursement is $23 per FIT vs $493 per FIT plus sDNA test.

Even though its superiority in a program of screening has not been empirically established, colonoscopy remains the criterion standard for assessing the test performance of other CRC screening tests. Moreover, colonoscopy is significantly more invasive than other available tests and thus carries a greater possibility of procedural complications, as well as harms of overdiagnosis and overtreatment of smaller lesions (ie, <10 mm). Three large RCTs of screening colonoscopy in average-risk adults are under way and will provide information about the long-term CRC incidence and mortality outcomes: the...
Table 4. Summary of Evidence by Key Question and Screening Test

<table>
<thead>
<tr>
<th>Test Name</th>
<th>Study Design</th>
<th>No. of Studies</th>
<th>No. of Participants</th>
<th>Summary of Findings (Includes Consistency, Precision)</th>
<th>Applicabilitya</th>
<th>Limitations (Includes Reporting Bias)</th>
<th>Overall Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIG</td>
<td>RCT</td>
<td>4</td>
<td>458 002</td>
<td>SIG consistently decreased CRC-specific mortality compared with no screening at 11-12 y of follow-up (IRR, 0.73; 95% CI, 0.66-0.82). Mortality benefit was limited to distal CRC.</td>
<td>Fair to poor. No longer widely used in the United States.</td>
<td>Only 1 trial evaluated more than a single round of screening. Variation in referral criteria led to differing rates of follow-up colonoscopy.</td>
<td>Fair to good</td>
</tr>
<tr>
<td>gFOBT, Hemoccult II</td>
<td>RCT</td>
<td>5</td>
<td>419 966</td>
<td>Biennial screening with Hemoccult II compared with no screening (404 396) consistently resulted in reduction of CRC-specific mortality (ranging 9%-22% after 2-9 rounds of screening with 11-30 y of follow-up).</td>
<td>Poor. No longer widely used.</td>
<td>Variation in number of screening rounds, use of rehydrated samples, definition of “test positive,” and recommended diagnostic follow-up.</td>
<td>Fair to good</td>
</tr>
<tr>
<td>Colonoscopy</td>
<td>Prospective</td>
<td>4</td>
<td>4821</td>
<td>Comparing colonoscopy with CTC or CTC plus colonoscopy, per-person (or per-lesion) sensitivity for adenomas ≥10 mm was 89%-98%, and per-person sensitivity for adenomas ≥6 mm was 75%-93%.</td>
<td>Fair. Colonoscopies were conducted or supervised by “experienced” specialists.</td>
<td>Studies were not designed to assess diagnostic accuracy to detect cancers. Limited studies with large number of endoscopists that were applicable to community practice.</td>
<td>Fair to good</td>
</tr>
<tr>
<td>CTC</td>
<td>Prospective</td>
<td>9</td>
<td>6497</td>
<td>The per-person sensitivity and specificity of CTC using bowel preparation to detect adenomas ≥10 mm ranged 67%-94% and 86%-98%, respectively. The per-person sensitivity and specificity to detect adenomas ≥6 mm ranged 73%-98% and 80%-93%, respectively. In 2 studies, sensitivity without bowel preparation to detect adenomas was lower than that of CTC protocols using bowel preparation.</td>
<td>Fair. Mostly single-center studies, with ≤3 highly trained radiologists. Current practice may use different technologies and protocols.</td>
<td>Studies were not designed to assess diagnostic accuracy to detect cancers. Unclear if the variation of test performance was due to differences in study design, populations, bowel preparation, CTC technology, reader experience, or reading protocols.</td>
<td>Fair to good</td>
</tr>
<tr>
<td>FIT</td>
<td>Prospective</td>
<td>6 Qualitative</td>
<td>36 808</td>
<td>In studies with colonoscopy follow-up for all, FIT sensitivity varied considerably across assays for each outcome. OC-Light had the highest sensitivity and specificity for CRC, from 88% and 91%, respectively, to 79% and 93%, respectively. OC FIT-CHEK had the best sensitivity and specificity for CRC, from 73% and 96%, respectively, to 92% and 87%.</td>
<td>Fair to good. There is a wide range in costs for specific tests (OC-Light, OC FIT-CHEK, Cologuard). Quantitative FITs included some that are older and now discontinued.</td>
<td>Variation in test performance resulted from the use of 18 different FITs (FIT families), different numbers of stool samples, and to some extent different assay cutoff values. Sparse data on most individual tests limited comparisons.</td>
<td>Fair to good</td>
</tr>
<tr>
<td>FIT plus sDNA</td>
<td>Qualitative</td>
<td>7 Quantitative</td>
<td>40 134</td>
<td>A FIT plus sDNA assay (Cologuard) had better sensitivity but lower specificity, 92% (95% CI, 84-97) and 84% (95% CI, 84-85), respectively, compared with OC FIT-CHEK.</td>
<td>Fair to good.</td>
<td>FIT plus sDNA was limited to a single study with 6% inadequate stool samples.</td>
<td>Fair to good</td>
</tr>
</tbody>
</table>

(continued)
Table 4. Summary of Evidence by Key Question and Screening Test (continued)

<table>
<thead>
<tr>
<th>Test Name</th>
<th>Study Design</th>
<th>No. of Studies</th>
<th>No. of Participants</th>
<th>Summary of Findings (Includes Consistency, Precision)</th>
<th>Applicability*</th>
<th>Limitations (Includes Reporting Bias)</th>
<th>Overall Quality</th>
</tr>
</thead>
</table>
| Key questions 3a, 3b: Harms of Screening
| Endoscopy | Prospective and retrospective studies | 18 SIG | 331 181 | Harms from screening SIG were estimated at 1 perforation/10 000 procedures (95% CI, 0.4–1.4/10 000) (No. of studies = 16) and 2 major bleeds/10 000 procedures (95% CI, 0.7–4/10 000) (No. of studies = 10). Good. Reflects community practice. | Only 2 studies reported serious adverse events in persons without colonoscopy (no difference in serious harms other than perforation and bleeding). Likely reporting bias of serious harms other than perforation and bleeding. | Fair |
| 55 Colonoscopy | 10 398 876 | Harms from screening colonoscopy or colonoscopy in asymptomatic persons was estimated at 4 perforations/10 000 procedures (95% CI, 2.5–10 000) (No. of studies = 26) and 8 major bleeds/10 000 procedures (95% CI, 5.1–14/10 000) (No. of studies = 22). Risk of perforations, bleeding, and other serious harms from colonoscopy increased with age. | 
| CTC | Prospective and retrospective studies | 15 | 75 354 | Harms from CTC in asymptomatic persons were uncommon. Risk of perforation for screening CTC was <2/10 000 examinations. The range of low-dose ionizing radiation per examination was 1–7 mSv. Fair to good. Radiation exposure per examination may be decreasing over time. | No studies reported serious adverse events in persons without CTC. Limited evidence in true average-risk screening populations. Likely reporting bias of serious harms other than perforation. No studies reported differential harms by age group. No studies were able to quantify net benefits and harms of ECF findings. Varying levels of follow-up and few studies with final disposition of ECF. Very limited studies comparing CFC by age group. | Fair |
| 21 ECF | 38 193 | ECF was estimated to occur in up to 69% of examinations, and 5%–37% of examinations might necessitate diagnostic follow-up; however, ≤3% required any type of definitive treatment. Higher prevalence of ECF with increasing age. | 

Abbreviations: CRC, colorectal cancer; CTC, computer tomographic colonography; ECF, extracolonic findings; FIT, fecal immunochemical test; gFOBT, guaiac-based fecal occult blood test; IRR, incidence rate ratio; RCT, randomized clinical trial; sDNA, stool DNA; SIG, flexible sigmoidoscopy.

* Applicability or external validity to US practice.

Key question 1: What is the effectiveness of screening programs in adults at average risk for colorectal cancer, compared with no screening, in reducing the incidence of or mortality from colorectal cancer?

Key question 2: In adults at average risk for colorectal cancer, what are the test performance characteristics (eg, sensitivity and specificity) of a 1-time application of a screening test, compared with an adequate reference standard, for detecting colorectal cancers, advanced adenomas, or adenomas based on size?

Key question 3a: What are the serious adverse effects of colorectal cancer screening tests in asymptomatic adults? Key question 3b: Do adverse effects vary by important subpopulations (eg, age)?

Northern European Initiative on Colorectal Cancer (NordICC) trial, comparing screening colonoscopy with usual care (estimated primary completion date, June 2026)\(^\text{203}\), COLONPREV, comparing colonoscopy with biennial FIT in Spain (estimated primary completion date, November 2021)\(^\text{27,204,205}\); and CONFIRM, comparing colonoscopy vs annual FIT in the United States (estimated primary completion date, September 2027)\(^\text{206}\).

Evidence continues to accrue that CTC adequately detects CRC and large potential precursor lesions. Although the risk of immediate harms from screening CTC (eg, bowel perforation from insufflation) is very low, it is unclear what (if any) true harm is posed by cumulative exposure to low-dose radiation or detection of extracolonic findings. Although the radiation dose appears to be decreasing over time due to technological and protocol advancements, it still ranges as high as 7 mSv per examination (dual positioning). Given that the average amount of radiation one is exposed to from background sources in the United States is about 3 mSv per year\(^\text{207}\), ionizing radiation from a single CTC examination is low. However, current expert recommendations are to repeat CTC every 5 years, and even low doses of ionizing radiation could cumulatively convey a small...
excess risk of cancer.\textsuperscript{208,209} From empirical evidence to date, it remains unclear whether detection of extracolonic findings represents a net benefit or harm.

This evidence report and systematic review did not address several important issues: screening in high-risk adults (ie, those with known family history of CRC), risk assessment to tailor screening, test acceptability, availability of or access to screening tests, methods to increase screening adherence, potential harms of overdiagnosis or unnecessary polypectomy, overdue or misuse of screening, and surveillance after adenoma detection. This review was commissioned along with a separate set of microsimulation decision models from the Cancer Intervention and Surveillance Modeling Network (CISNET) that addressed other important gaps in evidence, including ages to start and stop screening, screening intervals, and targeted or tailored screening.\textsuperscript{210} The review was limited to evidence conducted in countries with the highest applicability to US practice; in addition, only articles published in English were considered for inclusion.

Unlike other routinely recommended or conducted cancer screening, there are multiple viable options for CRC screening. These options have various levels of evidence to support their use, aims (eg, to detect cancers, potential precursor lesions, or both), test acceptability and adherence, intervals of time to repeat screening, need for follow-up testing (including surveillance incurred), associated serious harms, availability in practice, cost, and advocacy for their use. This complexity is compounded by testing whose quality is more operator-dependent (eg, colonoscopy, CTC), as well as rapid technological advancements in improving existing tests or developing new tests.

Empirical studies, trials, or well-designed cohort studies with average-risk populations are still needed to evaluate programs of screening using colonoscopy, the best-performing stool tests, and effect of CTC on cancer mortality and cancer incidence. Also needed are studies of diagnostic accuracy to confirm the screening test performance of promising stool tests based on high sensitivity to detect CRC or advanced adenomas with thus far limited reproducibility (ie, only 1 study). Diagnostic accuracy studies, particularly those evaluating new or more complex technologies, should report percentages of inadequate or indeterminate results. It is also important to understand the contribution of technological advancements to existing technology (eg, enhancements to optical colonoscopy or CTC) on test performance in average-risk adults as well as on reducing harms (eg, decreasing radiation exposure, less aggressive bowel preparation). More complete and consistent reporting regarding downstream benefits and harms from initial detection (ie, subsequent workup and definitive treatment) of C-RADS E3 and E4 findings need to be published in observational studies or trials with longer-term follow-up. Data are still needed on the differential uptake of and adherence to screening modalities and on continued adherence to repeated rounds of screening and diagnostic follow-up to screening over longer periods.

Conclusions

Colonoscopy, flexible sigmoidoscopy, CTC, and various stool tests have differing levels of evidence to support their use in CRC screening, ability to detect CRC and precursor lesions, and risk of serious adverse events in average-risk adults. Although CRC screening has a large body of supporting evidence, additional research is still needed to weigh the relative benefits and harms of each test in within a program of screening.

**Funding/Support:** This research was funded by the Agency for Healthcare Research and Quality (AHRQ) under a contract to support the USPSTF.

**Role of the Funder/Sponsor:** Investigators worked with USPSTF members and AHRQ staff to develop the scope, analytic framework, and key questions for this review. AHRQ staff provided project oversight; reviewed the report to ensure that the analysis met methodological standards; and distributed the draft for peer review. AHRQ reviewed and approved the manuscript before submission, but had no role in the design and conduct of the study including study selection, quality assessment, analysis, and interpretation of the data; preparation of the manuscript; and decision to submit the manuscript for publication.

**Additional Contributions:** We gratefully acknowledge the following individuals for their contributions to this project: Smyth Lai, MLS; Kevin Lutz, MFA; and Elizabeth Hess, MA, ELS(D), at the Kaiser Permanente Center for Health Research; Jennifer Crosswell, MD, MPH, at the Agency for Healthcare Research and Quality; and current and former members of the US Preventive Services Task Force who contributed to topic deliberations. USPSTF members, peer reviewers, and federal partner reviewers did not receive financial compensation for their contributions.

**Additional Information:** A draft version of this evidence report underwent external peer review from 6 content experts (James Allison, MD, University of California, San Francisco; Samir Gupta, MD, MScS, University of California, San Diego; Theodore R. Levin, MD, Kaiser Permanente; David Lieberman, MD, Oregon Health & Science University; Perry Pickhardt, MD, MPH, University of Wisconsin; David Ransohoff, MD, University of North Carolina at Chapel Hill) and 4 federal partners: Centers for Disease Control and Prevention, National Institutes of Health, US Department of Veterans Affairs, and Indian Health Service. Comments were presented to the USPSTF during its deliberation of the evidence and were considered in preparing the final evidence review.

**Editorial Disclaimer:** This evidence report is presented as a document in support of the accompanying USPSTF Recommendation Statement. It did not undergo additional peer review after submission to JAMA.

**REFERENCES**


3. US Preventive Services Task Force. Screening for colorectal cancer: US Preventive Services Task...


Evidence Report: Screening for Colorectal Cancer


Colorectal cancer screening is effective at reducing colorectal cancer mortality. However, its impact on overall mortality is uncertain. The net benefit of colorectal cancer screening is affected by many factors, including the prevalence of colorectal cancer in the screened population, the effectiveness of diagnostic and therapeutic interventions, and the prevalence of other diseases. This report reviews colorectal cancer screening strategies, including the prevalence of colorectal neoplasms (polyps and cancer) in the general population, the effectiveness of endoscopic and non-endoscopic screening tests, and the benefits and harms of colorectal cancer screening. The report also discusses the results of clinical trials, which have provided direct evidence on the impact of colorectal cancer screening on mortality in several countries. The evidence indicates that colorectal cancer screening is effective in reducing colorectal cancer mortality and that the benefits of screening outweigh the harms in most screened populations. The benefits of colorectal cancer screening are greatest in populations with relatively high colorectal cancer prevalence and in populations with a high prevalence of colorectal cancer risk factors. The harms of colorectal cancer screening are modest and include complications of colorectal cancer screening, such as perforation, bleeding, and procedure-related complications. Overall, the benefits of colorectal cancer screening outweigh the harms, and colorectal cancer screening is recommended for populations at risk for colorectal cancer. Evidence Report: Screening for Colorectal Cancer. JAMA. 2008;300(18):2136-2140.