View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Carolina Digital Repository

Incidental Findings Found In Typical Screening Populations Undergoing Colorectal Cancer Screening with Computed Tomography Colonography: A Systematic Review

By

Andrew M. Moon

A Master's Paper submitted to the faculty of
the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements for
the degree of Master of Public Health in
the Public Health Leadership Program

Chapel Hill
2012

_____

Advisor Signature

_____

Advisor Printed Name ***Date***

_____

Second Reader Signature

_____

Second Reader Printed Name ***Date***

Abstract:

Background: CT colonography (CTC) is a noninvasive technology used to screen for colorectal cancer. Unlike other screening modalities, CTC provides a view of the abdomen and pelvis allowing radiologists to detect lesions in extracolonic organs. There is much debate on the balance of potential benefits versus potential harms of discovering, working up and treating these extracolonic findings. This debate might be especially relevant for asymptomatic populations receiving screening with CTC.

Purpose: This systematic review aims to determine the frequency and clinical implications of finding incidental, extracolonic lesions during CT colonography (CT) in asymptomatic, screening populations. In addition, this review reports the frequency and clinical outcomes of clinically important lesions. Lastly, this review summarizes the various methods studies used to define the clinical significance of incidental findings.

Data Sources: I carried out a systematic search of MEDLINE, Embase, the Cochrane Clinical Trials databases and published reviews up to March 2012.

Study Selection: Two investigators independently reviewed 282 abstracts and 53 full text articles using a set of predefined inclusion and exclusion criteria. Both reviewers carried out independent critical appraisals of each study using criteria developed by the United States Preventive Services Task Force.

Data Extraction: One reviewer extracted information on study samples, designs, populations, interventions and outcomes from six studies. A second reviewer verified this information for accuracy.

Data Synthesis: The frequency of extracolonic findings (ECFs) ranged from 27.2% to 68.9% (mean 49.3%). Included studies used similar classification systems of clinical importance, which were primarily based on the likelihood of clinical workup. Studies reported that 5.6% of the reported ECFs were of high clinical importance and 15.5% of lesions were either moderate- or high-importance. A minority of these findings represented lesions that could have benefitted from early diagnosis and intervention. Studies reported that 0.09% to 1.2% of subjects were diagnosed with AAAs and 0.23% to 0.88% were diagnosed with extracolonic cancers. Studies used widely varying lengths and methods of following ECFs, making it difficult to estimate the true clinical implications of incidental findings. However, the range of moderate/high to high-importance findings (5.6% to 15.5%) provides a good estimate of the number of subjects requiring some clinical workup.

Limitations: I identified several weaknesses of the available literature on ECFs from screening CTC. For instance, many included studies suffered from poor follow-up and incomplete reporting of outcomes. In addition, no studies properly addressed the potential physical and psychological harms of being diagnosed, worked up and treated for extracolonic findings. Lastly,

the included literature does not address how ECFs are handled in non-academic settings. This systematic review also had several weaknesses. The decision to limit the review to screening populations might reduce the strength of my findings. I attempted to compensate by including populations at high risk of CRC and studies conducted outside the US, but this might have reduced the generalizability of my findings. Furthermore, I were unable to adjust for different follow-up time periods, making it difficult to compare the clinical outcomes of ECFs among included studies. Lastly, I attempted to develop an outcomes table for ECFs from screening CTC, but were unable to do so because of the imprecision of results, variable periods of follow-up and gaps in reported outcomes.

Conclusions: Based on these results, a large proportion of individuals receiving CT colonography for colorectal cancer screening will have an extracolonic lesion discovered. Roughly one-fifth to one-third of these findings will receive some clinical workup and the majority of these will ultimately be diagnosed as benign. Since a small percentage of potentially important findings will result in clinical benefit, it is possible that the classification systems are overly sensitive. In addition, the reporting of all extracolonic findings might result in unnecessary testing and patient anxiety. Unfortunately, the existing data does not provide enough certainty to know which lesions can go unreported without putting the patient at harm. However, based on the evidence, it appears that most radiologists and primary care physicians err on the side of reporting findings, which also results in unnecessary harms to patients. Another source of unnecessary and potentially harmful care is the large variability in radiologist interpretation of extracolonic findings. Based on this review, there are no indications that the development of a standardized classification system of ECFs has successfully reduced this variation. There are two primary ways to improve this practice variability in the future. First, classification systems could be improved to provide more guidance, especially for findings that have an uncertain balance of benefits and harms. More primary research might be required before this is possible. Second, training programs for CTC should require specific training for interpreting ECFs, including the proper follow-up of specific findings.

Introduction

Statement of Purpose:

CT colonography (CTC) is a noninvasive technology used to screen for colorectal cancer.[1] Unlike alternative screening modalities such as colonoscopy, CTC provides a view of the extracolonic abdomen and pelvis, which allows radiologists to detect lesions in several abdominal organs including the kidneys, liver and pancreas. For some, such incidental findings might be viewed as serendipitous discoveries that allow earlier intervention and improved outcomes. Alternatively, detecting such lesions could lead to harms such as unnecessary diagnostic workups, interventions and patient anxiety. In addition, the radiologic surveillance or intervention for extracolonic findings could carry significant financial costs that might affect the cost-effectiveness of CTC as a screening tool.[2]

In this systematic review, I examined the frequency and clinical implications of incidental lesions detected during screening CT colonography (CTC). Characteristics of study populations and CTC techniques are reported to indicate which factors might influence the frequency of incidental findings. In addition to reporting the overall frequency of these findings, I provide estimates of the reported clinical importance of these findings. Many studies on incidental findings have categorized lesions as high, moderate and low clinical importance and the Working Group on Virtual Colonoscopy have proposed a similar classification system.[3] These categories are designed to inform clinical care with higher importance findings requiring surveillance or immediate intervention. In this review, I systematically assessed the methods study authors used to classify findings into categories of clinical importance. The methods used for categorizing incidental findings provides important context for assessing the reported frequency of findings deemed to be clinically important. This review assesses these classification

systems to determine how they might affect the frequency and workup of extracolonic findings and whether they are likely to reduce variability in clinical practice.

This systematic review aims to inform radiologists and primary care physicians of the likelihood of finding incidental lesions during screening CTC, the clinical outcomes of detecting these lesions and the usefulness of classifying these findings by clinical importance. The last systematic review on the frequency and implications of incidental lesions was published in 2005 and included 3 studies focusing on screening populations, although some of these studies included a large percentage of symptomatic patients.[4] This previous review was thus unable to make a conclusion on the frequency and clinical implications of extracolonic findings in pure screening populations. My review, on the other hand, focuses specifically on asymptomatic populations that more closely resemble a general screening population for CTC. I thought it was particularly useful to focus on an asymptomatic population since it is possible that patients' symptoms could be attributable to incidental lesions. In addition, there are many more ethical issues raised when diagnosing asymptomatic, healthy patients with findings detected incidentally. My review should also provide a unique view on classification systems for judging the clinical importance of extracolonic findings. Only one review, published in the Journal of Law, Medicine, and Ethics,[5] has systematically assessed how study authors judge clinical significance of incidental findings. Therefore, this systematic review will provide an updated view of the frequency and clinical outcomes of incidental findings from CTC in a screening population and report how study authors classify the clinical importance of extracolonic lesions. I believe that such a review is important as lower radiation doses are used for CTC,[6] as recommendation statements on interpreting and managing extracolonic findings are disseminated,[7] and as more insurance plans cover CTC for colorectal cancer screening.[8]

Problems with Current Screening Technologies for Colorectal Cancer:

Colorectal cancer (CRC) is the third most common cancer and the second leading cause of cancer deaths in the United States.[9] CRC has a well-characterized preclinical period during which most tumors develop from precursor lesions. In addition, early detection and treatment of CRC reduces its mortality. All of these characteristics make it an appropriate candidate for screening.

Screening for CRC has been a significant factor in its declining incidence and mortality in recent years.[10,11] Consequently, the United States Preventive Services Task Force, the American College of Radiology and the US Multi-Society Task Force on Colorectal Cancer (USMSTF) all recommend colorectal cancer screening for individuals starting at age 50 or earlier for those with certain inherited syndromes or inflammatory bowel disease.[12,13] Despite its demonstrated benefit and supporting recommendations, CRC screening is being underutilized by huge numbers of American adults aged 50 and older.[14] There are many reasons for this, including the drawbacks associated with each individual screening technique.

Colonoscopy is an increasingly preferred screening technology,[14] potentially because of negative coverage decision by the Centers for Medicare and Medicaid Services[15] and similar decisions by many other private insurers to cover colonoscopy.[16] Colonoscopy also has high public perceptions of its accuracy[17] and the demonstrated benefits of high sensitivity, high specificity and ability for immediate polyp removal.[13] A long-term follow-up study of the National Polyp Study cohort reported a 53% reduction in mortality from colonoscopy and polypectomy.[18] However, colonoscopy has several drawbacks that might reduce compliance including laborious bowel cleansing requirements, the need for sedation, patient discomfort

during the procedure and the risk of serious complications such as bowel perforation and serious bleeding.[19] In a recent trial,[20] patients who were recommended colonoscopy for CRC screening had a significantly lower rate of adherence compared to those offered FOBT or given a choice between FOBT and colonoscopy. In addition, colonoscopy might not be available for some due to its expense and the limited number of trained endoscopists.[21]

Sigmoidoscopy might limit patient discomfort and the need for sedation, but it shares many of colonoscopy's drawbacks. In addition, some have raised concerns that sigmoidoscopy misses important lesions in the proximal colon.[22,23] One way to compensate for this lost sensitivity is pairing sigmoidoscopy with fecal occult blood tests (FOBT). Sigmoidoscopy every five years combined with yearly FOBT screening has been recommended as an alternative to colonoscopy.[12] But FOBT has its own shortcomings, including its high risk of false positives, resulting in unnecessary additional endoscopies.[24]

Double contrast enhanced barium enema (DCBE) is a CRC screening technique with good safety profile and a moderate cost[25] but concerns about DCBE's poor sensitivity to detect polyps have contributed to its decreased use in recent years.[26]

Techniques of Screening CTC:

After first being described in 1994,[27] CT colonography or virtual colonoscopy has emerged as a new, non-invasive technique for CRC screening.[28] CTC involves a helical, thin-section CT of a distended and cleansed colon, providing data that can be reconstructed into two- and three-dimensional images.[29] Bowel cleansing for CTC consists of patients maintaining a clear liquid diet for 24 hours and cathartic cleansing with laxatives, similar to conventional colonoscopy. Tagging liquid and fecal material is possible with the oral administration of barium

and/or iodine contrast.[29] For CTC it is necessary to distend the bowel by blowing room air or carbon dioxide through a catheter into the rectum.[30] Bowel distension helps avoid missing lesions hidden by undistended or collapsed segments of bowel. After preparation, CT scanning is performed with the patient in supine and prone positions to help differentiate between polyps and stool.

There are several CT characteristics that must be considered by radiologists carrying out CTC, including slice thickness, radiation dose and the use of IV contrast. Recommendations on slice thickness state that it should not exceed 3 mm when using multi-detector computed tomography (MDCT) on screening populations.[31] There are several ways to change the radiation dose in CTC including increasing pitch and slice collimation, decreasing the voltage (kVp) or current (mAs) or employing automatic exposure control in smaller patients.[32] The appropriate radiation dose has been a moving target, as researchers attempt to reduce the potential for iatrogenic injury. Investigators have taken advantage of the high contrast between the colonic mucosa and air in order to reduce radiation doses.[6,33] A recent study by Macari et al. reported excellent detection of polyps larger than 10 mm with significantly lower radiation doses.[34] Ultra-low dose protocols, which set the current to the lowest setting possible (10 effective mAs or 40 electric mAs) and cut the dose used by Macari et al. by 80%, have also proven effective for detecting polyps.[35] While these studies show that ultra-low doses of radiation can be used to detect polyps, such doses might not be sufficient for detecting extracolonic lesions. Using low tube currents decreases the number of photons that reach detectors and thus increases image noise, which is less of an issue for detecting polyps due to the high contrast between the intracolonic air and the colonic mucosa. In addition, low radiation doses might provide adequate penetration for obese patients.[29] Another consideration is the use of IV iodine contrast, which

might improve the diagnostic quality of CTC. But due to its extra cost, requirement for IV access and the risk of anaphylactic allergic reactions, IV contrast is not currently recommended for screening CTC.[31]

Acceptance of CTC in the United States:

CTC for colorectal cancer screening had delayed early acceptance, likely resulting from conflicting reports on its sensitivity for detecting polyps.[8] Some concerns were addressed by the National CTC Trial,[36] a trial with 2,531 individuals in 15 centers, reported 90% sensitivity and 86% specificity for detecting large adenomas. Nevertheless, this study involved only specially trained radiologists and did not report detection of lesions < 5 mm. In addition, reported sensitivities and specificities of CTC for screening have been more heterogeneous and less encouraging in low-risk populations.[29]

These weaknesses, in addition to concerns of the implications of extracolonic findings, contributed to the USPSTF's negative recommendation for CTC as a cancer screening technique in 2008.[19] The same conclusion was reached by the American College of Physicians in their updated recommendations in 2012.[37] In contrast, in 2008 CTC was endorsed as an appropriate CRC screening technique by the American Cancer Society, American College of Radiology (ACR) and the U.S. Multisociety Task Force on Colorectal Cancer.[38]

There are indications that CTC has not won over primary care physicians, who play a large role in recommending screening tests for colorectal cancer. Only a minority (23%) of surveyed primary care physicians in the U.S. felt that CT colonography was very effective at reducing colorectal cancer mortality. Less than 5 percent of these surveyed physicians said they routinely recommend CT colonography for CRC screening. [39] It is unclear whether these views

would be dramatically affected by a positive recommendation for CTC by the USPSTF or a change in Medicare reimbursement for CTC.

Indications and Advantages of Screening CTC:

There are several reported indications for screening CTC. For instance, CTC is a useful option after a failed optical colonoscopy.[40,41] Failed colonoscopy occurs in roughly 5% of patients as a result of patient discomfort, colon tortuosity, adhesions from previous surgeries or hernias.[28] CTC can also be used to evaluate the colon proximal to an obstructing colon cancer, although bowel preparation can be challenging if obstruction is near-complete.[42,43] CTC might also be used for patients with contraindications to endoscopy. Common contraindications for colonoscopy include advanced patient age, severe comorbidities, predisposition to severe bleeding or prior adverse reaction to sedation.[28] Due to its need for colonic distension, CTC also has several contraindications including acute abdominal pain, recent abdominal surgery, entrapment of colonic loops from abdominal wall hernia or acute inflammatory conditions (acute diverticulitis, active Crohn's disease or ulcerative colitis and toxic megacolon).[44-46] CTC can also be used for patients who refuse other CRC screening options. Lastly, it's possible that CTC could be considered a primary option for colorectal cancer screening in the near future.[47]

There are several advantages of CTC that might make it more acceptable to some patients and providers. First, CTC shows sensitivity and specificity for adenomas >10 mm comparable to colonoscopy, which is currently the gold standard.[19] This might make CTC more acceptable to those who worry about the variable sensitivity and specificity of FOBT and sigmoidoscopy. There is the promise of advances in technology that could lead to improved computer-aided polyp detection techniques in the near future.[48,49]

CTC might also be more acceptable to patients concerned about the discomfort of colonoscopy, although there are some inconsistencies in studies looking at this issue. Some studies have shown that patients receiving CTC report less pain than during colonoscopy,[26,50,51] although other studies report just the opposite.[52-54] These differences are likely explained by whether investigators used spasmodic bowel agents or sedation.[55] If fecal and fluid tagging procedures improve, the pendulum of patient acceptance might swing in favor of CTC, especially since bowel preparation is viewed as one of the most onerous features of CTC.[56] Another advantage to CTC is the decreased risk of serious complications such as bowel perforation, serious bleeding and adverse effects of sedation.[28] Since sedation is not required for CTC, patients are not required to secure a ride from the procedure and, compared to colonoscopy, can return to work sooner.

Concerns with CTC:

In addition to the need for full bowel preparation and the need for colonic insufflation, there are several concerns with using CT colonography for CRC screening. First, it is very difficult, if not impossible, to detect flat or small (< 5 mm) colonic lesions with CTC.[57] CTC also involves radiation exposure and the associated risks of iatrogenic malignancy,[58] although these risks will probably be reduced with low-dose protocols for screening CTC.[29] Furthermore, past cost-effectiveness analyses of CTC report that it is the most expensive modality for detecting an adenoma.[59]

Incidental findings are one feature of CTC that has been described as both an advantage[60] and a flaw[61] of this technology. Early detection of lesions such as abdominal aortic aneurysms or extracolonic malignancies might allow for early intervention and improved prognosis. But

detection of some AAAs or cancers might lead to unnecessary patient anxiety from being labeled with a serious condition, overdiagnosis and overtreatment. It is difficult to weigh these potential benefits and harms of detecting such life-threatening lesions. CTC might also detect benign lesions that could be misidentified as being clinically important. It is possible that a patient could be subjected to significant inconvenience, radiation exposure and possibly surgery for what turns out to be a benign finding. In the face of these uncertainties, it is difficult for radiologists and primary care physicians to know how to properly address incidental findings detected during screening CTC.

Important Factors in the Frequency of Incidental Findings:

There are several factors that might influence the frequency of extracolonic findings including CT technique, patient features and radiologists' detection thresholds. These factors are important when comparing the relative frequencies of incidental findings reported in the literature. In addition, these factors might affect the generalizability of findings for different screening populations in the United States.

Radiation dose is one aspect of the CT technique that might influence the ability to detect incidental lesions. In light of reductions in radiation for CTC, the frequency of extracolonic findings might decrease with emphasis on reduction of radiation doses for screening CTC. Low-radiation protocols might also limit the specificity of initial diagnoses of incidental lesions, potentially making it difficult to determine the clinical significance of an extracolonic finding. Another important CT factor is the use of IV contrast. Since IV contrast is also not recommended for screening CTC, its ability to correctly diagnosis incidental findings may be reduced. One study comparing the incidence of extracolonic findings in symptomatic patients receiving CTC

reported that lesions were found in 71% of those who received IV contrast compared to 29% of those who underwent an unenhanced scan.[62] Lastly, CT slice thickness might influence the detection of extracolonic findings, with thinner slices leading to increased detection of lesions.

There are several patient factors that might affect the detection of incidental findings. It is important to take patient age into consideration, especially since CRC screening is generally recommended for patients age 50 or older. In addition, the presence of symptoms might influence the detection of incidental findings or the likelihood of these lesions being reported. Therefore, reported frequencies from cohorts of symptomatic patients might be less applicable to screening populations. Lastly, the a-priori risk of colorectal cancer could affect the frequency of extracolonic findings. The accuracy of various screening techniques for detecting polyps might also vary for low- and high-risk patients. Lastly, the ACR White Paper on CTC states that screening CTC is contraindicated for certain high-risk patients (e.g. hereditary polyposis or nonpolyposis cancer syndromes).[32]

The radiologist's level of training might impact the likelihood of following up on an incidental finding. In a retrospective analysis of radiologist reports in the U.S., the odds of recommending additional imaging decreased by 15% with each decade of radiologist experience.[63] In addition, experience in community settings have shown that CTC experience does not substitute for proper training.[64] In light of these findings, frequencies of extracolonic findings from studies performed at academic medical centers might be less applicable to community hospitals or outpatient endoscopy suites. There might also be a temporal trend towards increased detection of incidental findings. After adjusting for potential confounders, radiologists in 2008 were 2.16 times as likely to recommend additional imaging than radiologists in 1995.[63] Lastly, the use of multiple radiologists to corroborate results might influence the

detection of incidental findings. Supporting this, large, prospective studies have reported large interobserver variability among radiologists interpreting CTC.[65-67]

Pressures to Address Incidental Findings:

The American College of Radiology's Incidental Findings Committee wrote that many physicians are unwilling to accept diagnostic uncertainty in the face of incomplete data, a lack of clear diagnostic and treatment algorithms, fear of litigation and a desire to adhere to the "better safe than sorry" philosophy.[7] To help address these concerns, the ACR developed a set of rules with the aim of "optimizing" utilization of imaging when addressing incidental lesions.[7] They have released specific diagnostic and treatment guidelines for incidental findings in the kidneys, liver, adrenal glands and pancreas. The recommendations include separate considerations for low-dose, unenhanced CT examinations, like CTC. While the recommendations attempt to provide straightforward guidance to radiologists in order to reduce unnecessary further workup, there are several weaknesses with their coverage of low-dose unenhanced CT. First, there were many gaps in the evidence, making it difficult for them to develop truly evidence-based recommendations. For instance, they found no studies addressing the management of lesions found during unenhanced CT including lesions in the kidneys, liver, adrenal glands and pancreas. Furthermore, the White Paper did not address how to handle lesions of the lungs, stomach, small bowel, ovaries, gallbladder, retroperitoneum, uterus, prostate, urinary bladder and bone. In addition, the ACR recommendations concede that formulaic recommendations are not always appropriate since patient factors, such as age or comorbidities, might change a physician's approach to workup and treatment. This is especially true for screening CTC, since screening populations are generally older and thus may have a different frequency of incidental findings or

require a more conservative surveillance and treatment strategies.[39] In light of these concerns, the recommendations provided by the ACR might be less helpful for guiding clinicians who discover incidental findings during CRC screening with CT colonography.

The Working Group on Virtual Colonoscopy has published a classification system of incidental findings (Table 1) as part of the CT Colonography Reporting and Data System (C-RADS).[3] This effort was designed to provide a standardized method for characterizing the clinical importance of incidental lesions on CTC in order to minimize excessive costs and unnecessary patient anxiety. Incidental lesions are classified as E0-E4, in a way similar to the BI-RADS classification system for screening mammography.[68] As indicated in Table 1, E2 findings should not receive workup, by definition. E3 findings are incompletely characterized lesions with work-up that is subject to local physician practice and patient preference. E4 findings are potentially important findings that should be communicated to the referring physician (e.g. primary care provider or gastroenterologist). These are likely to require further workup or immediate treatment.

While the C-RADS system provides a helpful framework, it is unclear how often these definitions are used in clinical practice or research protocols. Since this standardized system has only been introduced fairly recently, it is likely that many of the studies reporting incidental findings from CTC have used different definitions for the clinical importance of findings.

Table 1. C-RADS Classification of Extracolonic Findings on CTC (Adapted from Zalis et al.[3])

	Description	Examples
E0	**Limited Exam** Compromised by artifact; evaluation of extracolonic soft tissues is severely limited	N/A

E1	**Normal Exam or Anatomic Variant** No extracolonic abnormalities visible.	Retroaortic left renal vein
E2	**Clinically Unimportant Finding** No workup indicated	a. Liver, Kidney: simple cysts b. Gallbladder: cholelithiasis without cholecystitis c. Vertebra: hemangioma d. Hiatal hernia
E3	**Likely Unimportant Finding, Incompletely Characterized**	a. Kidney: minimally complex or homogenously hyperattenuating cyst b. Small lung nodule
E4	**Potentially Important Finding**	a. Kidney: solid renal mass b. Ovarian mass c. Lymphadenopathy d. Vasculature: aortic aneurysm e. Lung: non-uniformly calcified parenchymal nodule ≥ 1 cm

It is unclear how these classification systems were developed and how they might impact clinical practice. For instance, non-uniformly calcified, large (≥ 1 cm) lung nodules are considered potentially important, but it is unclear how to classify small, calcified nodules or large, uniformly calcified nodules. In addition, it is unclear which classifications require workup or intervention. Two of the most serious clinical findings, extracolonic malignancies and abdominal aortic aneurysms (AAAs), have significant clinical heterogeneity that might require a more nuanced workup and treatment. Lastly, the reliability or validity of this classification system has not been independently assessed. As a result, this classification system has not been endorsed or widely applied in practice. A survey of 1600 radiologists asking about the approach to incidental findings at chest CT demonstrated a wide variability in practices and a substantial deviation from recommended medical practice.[69] It is therefore possible that even the best classification systems for addressing incidental findings has a limited effectiveness on clinical practice. Furthermore, there are indications that non-radiologist physicians have a poor grasp of recommendations for CTC.[70] This is important since primary care physicians, not radiologists, are responsible for the workup that ensues from detecting extracolonic lesions on CTC.

Summary:

To address the low uptake of colorectal cancer screening, CT colonography might be used increasingly to screen average risk adults. In fact, President Obama underwent a CTC for colorectal cancer screening during his most recent presidential physical.[71] Third-party payers might soon reimburse colorectal cancer screening of average-risk adults with CTC, leading to its widespread adoption. Such a development would increase the number of incidental, extracolonic lesions detected during screening CTC. Currently, radiologists and primary care physicians face conflicting information regarding the proper approach to extracolonic findings. In the face of this uncertainty, this systematic review reports the frequency of extracolonic findings in CTC used for screening for colorectal cancer, including how many of these findings were considered to be clinically important. This review also reports how study authors classified the clinical importance of lesions and the potential effects these determinations have on the workup and treatment of incidental findings. It would be informative for both radiologists and primary care providers to know the extent of testing required to achieve diagnostic certainty. Standardized classification systems and algorithms will not necessarily improve care, even if they standardize care, since they have not been validated. In addition, large inter-rater variability might limit the usefulness of classification systems and reduce the generalizability of reported frequencies of clinically important incidental findings.

Key Questions:

This review aims to address the following questions:

(1) What is the overall frequency of incidental lesions detected during screening CT colonography?

(2) What is the frequency and clinical outcomes of incidental lesions deemed to be of high, moderate and low clinical importance?

(3) How do the authors of studies determine how to classify the clinical importance of incidental lesions?

Methods

In addition to reporting the frequency and clinical implications of extracolonic findings on CTC, I aimed to determine whether each study's distribution of high, moderate and low importance findings and the rationale provided for these classifications. In addition, I explored the potential benefits and harms of detecting extracolonic findings. While it is impossible to determine definitively whether detecting incidental findings will result in a net benefit,[72] in this review I have attempted to use information on the natural history of individual lesions to develop probabilistic estimates of potential harms and benefits to patients. A search of systematic reviews was conducted and revealed that no recent reviews have focused specifically on the frequency and clinical implications of incidental findings specifically for screening CT colonography.

Eligibility criteria:

I limited the review to randomized controlled trials, cohort studies and high-quality case series. I reviewed systematic reviews and meta-analyses for references and commented on any their unique conclusions, but their findings were not included in my results to avoid duplicate counting of studies. Studies were required to identify the incidence of extracolonic lesions detected from screening CTC.

I based my inclusion/exclusion criteria for study populations to find patients that would best represent a screening population. As such, I excluded studies with 15% or more symptomatic patients. I also limited my review to studies with at least 75% patients between the age of 50 and 74 years old. When not enough information was provided to determine whether the study met these age criteria, I estimated the percentage of subjects aged 50 to 74 years old assuming uniform distribution. This age range was chosen to match the USPSTF recommendations for colorectal cancer screening.[73] I also chose to include only studies looking at CT colonography without IV contrast, in accordance with guidelines for screening CTC.[32] I considered applying a cut-off value for CTC radiation doses, but opted against doing so since radiation doses for CTC continue to change as technologies improve. I excluded studies including patients with a personal history of colorectal cancer. However, I included studies with subjects at higher risk for colorectal cancer, including those with a personal history of polyps, polyposis syndrome or family history of cancer. The studies that included high-risk subjects were considered separately from studies with low-risk screening populations. I considering limiting this review to studies looking at U.S. practices, given the heterogeneity of patient populations and the variability in radiologist experience and training.[74] However, the implications of incidental findings affect CTC programs beyond the U.S. In addition, many of these sources of variability, including radiologist training, experience and type of institution, are present within the United States. I therefore included studies from all settings and countries. To account for the potential variability introduced by including non-U.S. studies, I collected information on the setting of the study intervention and training of the radiologists. I did not consider any comparators for CT colonography since extracolonic findings are not detected during any other screening tests for colorectal cancer. In addition, I excluded studies that did not follow up

incidental findings and only reported frequency of extracolonic lesions from initial CTC. Furthermore, I only considered RCTs and cohort studies for inclusion in the final review. The full PICOTTS inclusion and exclusion criteria are listed below in Table 2.

I excluded studies with inadequate reporting of methods, population, interventions or outcomes. The full reporting criteria used to judge studies are included in Appendix B.

Two reviewers independently reviewed the titles and abstracts of studies identified during the searches and included studies based on the criteria listed above. When the reviewers disagreed or the abstract did not contain enough information to apply the criteria, the full article was reviewed. Disagreements were settled by consulting with a third investigator.

Table 2. PICOTTS framework for review of incidental findings in CTC

Category	Inclusion Criteria	Exclusion Criteria
Population	Predominantly asymptomatic patients age 50-74 being screened for colorectal cancer with CT colonography	- Studies with less than 75% of patients age 50-74 years old - Studies including greater or equal to 15% of symptomatic patients - Studies including patients with personal history of colorectal cancer
Intervention	Standard technique screening CT colonography (i.e. low-dose, non-contrast enhanced CTC)	- Diagnostic CTC (e.g. IV contrast) - MRI colonography - PET/CT
Comparators	N/A	N/A
Outcomes	- Primary outcomes: o Frequency of extracolonic lesions Frequency of high, moderate and low clinical significance extracolonic findings o Frequency of N0M0 extracolonic cancers and AAAs o Number of follow-up procedures (e.g. imaging, invasive procedures, surgeries) o Unnecessary procedures (i.e. for benign findings) o Patient anxiety - Secondary outcomes: o Reported methods of determining clinical importance of clinical findings (high, moderate, low)	Failure to report the frequency of extracolonic findings
Timing of Effect	Incidental findings and their resulting clinical implications reported at any time post-screening CTC	Studies that report no follow-up of patients with extracolonic findings.
Timing of Search	All studies published before March 4, 2012	
Setting	Any setting for screening CTC including hospitals, outpatient radiology suites and primary care clinics	
Study Designs	Randomized controlled trials and cohort studies Systematic reviews and meta-analyses can be used if PICOTTS are deemed to be equivalent to this review	Case-control studies, case reports, high-quality case series

Search strategy:

To identify original research on this topic, I conducted a systematic search of MEDLINE, Embase and the Cochrane Clinical Trials databases. I searched PubMed on March 4, 2012 using the following search: "(colonography[tw] OR virtual colonoscopy[tw] OR colography[tw] OR CT colonoscopy[tw] OR virtual endoscopy[tw]) AND (extracolonic[tw] OR incidental*[tw] OR incidentaloma*[tw] OR serendipitous[tw])." The search was adapted for Embase, which was accessed via Elsevier. A research librarian was consulted for the development of the search terms, which can be found below in Appendix A. I also performed manual searches of systematic reviews, included studies and background articles to find additional studies missed by my search strategy.

I placed no date or language limits on the search to avoid missing studies that had not yet been indexed. I performed an updated search, using the same search parameters, three months following the initial search to identify any studies published since the initial search. I imported citations into an Endnote (Thompson Reuters, New York, NY) electronic database.

Data extraction:

One reviewer extracted data on study samples, designs, populations, interventions and outcomes using a standardized spreadsheet. These data were verified by a second reviewer and discrepancies were resolved by consensus.

I extracted information on study methodologies such as the method of sampling (e.g. whether subjects were selectively studied or were consecutive cases), whether data collection was prospective or retrospective and the funding sources or potential conflicts of interest. I determined whether studies excluded previously diagnosed extracolonic findings. In addition, I

extracted the study's methods for determining classifications of clinical importance. I recorded the number of observers (radiologists), the training level of radiologists, whether the study was set in a community or academic setting and where the study was based. Lastly, I recorded the clinical specialty of study authors and any potential conflicts of interest.

I also collected information on study populations including age, ethnicity, presence of symptoms, risk of colorectal cancer. I obtained information on interventions including CT factors such as slice thickness, radiation dose, use of supine and/or prone exam and whether IV contrast was used. I extracted information on outcomes including the frequency of extracolonic findings from selected studies and the number of lesions with high, moderate and low clinical importance. Since these definitions varied among authors, I collected information on findings that might be considered life-extending: the number of "early" (N0M0) cancers and AAAs.[4] For AAAs, I used the definition from the USPSTF, which defines AAAs as expansions of the aorta below the renal arteries to a diameter greater or equal to 3 cm.[75] This is likely a conservative definition for potential clinical importance since the only two conditions with evidence supporting survival benefit from screening with abdominal imaging are AAA[76] and renal cell carcinoma.[77] In addition, many of these diagnoses will represent aneurysms and cancers that would have never have otherwise become clinically apparent (i.e. are "overdiagnosed"). Lastly, I gathered information on the number of recommendations for additional imaging (RAIs), surgeries, biochemical investigations, medical treatments and clinical appointments.

Assessing the Internal Validity of Studies:

I assessed the quality, or internal validity, of studies by applying the rules used by the United States Preventive Services Task Force (USPSTF).[78] I rated the internal validity/risk of

bias of studies as good, fair or poor quality. These ratings were based on my assessment of the recruitment of patients (i.e. consecutive sampling), maintenance of groups (i.e. attrition, cross-overs, adherence, contamination), the measurement of outcomes (i.e. equal, valid and reliable), consideration of all important outcomes, description of study populations, description of intervention, consideration of potential confounders and sample size. In addition, I developed quality criteria specific to my topic. The full quality criteria are included in Appendix C. Poor quality studies were defined as those with a fatal methodological flaw, such as more than 40% patient dropout. Fair quality studies had a few methodological flaws but no fatal flaws. Good quality studies had one to no methodological flaws. I hypothesized that the methodological quality of studies contributed to the likely heterogeneity of reported outcomes. I explored this hypothesis after grading each study's internal validity. I also assessed the power of studies as part of my quality assessment. I excluded studies with inadequate power ($n < 50$), since these results would be more susceptible to random variation.

Two reviewers independently made judgments on internal validity and analysis/power of studies. Disagreements were settled through consultation with a third investigator.

Assessing the Generalizability of Studies:

To assess the external validity of studies, I used the guidance provided in the USPSTF procedure manual with some minor adjustments.[78] I used information gathered for each study's populations, settings and interventions to assess how closely each resembled an asymptomatic population receiving colorectal cancer screening with CTC. I also used the GRADE Working Group's definition of directness to help guide this assessment of generalizability.[79] I judged each study's directness based on characteristics of the study population such as age, presence of

symptoms and risk of colorectal cancer. I also considered each study's setting, training and number of radiologists and CT colonography technique. Two reviewers independently graded the external validity of studies as good, fair or poor (full descriptions in Appendix D). Disagreements were settled through consultation with a third investigator.

Data Synthesis and Analysis:

For outcomes data, I collected the reported frequencies of extracolonic findings from CTC reported in each study, including the classification of findings into high, moderate and low clinical importance. I reported the number of recommendations for additional imaging, surgeries, medical treatments, confirmed diagnoses and unknown diagnoses for findings. Lastly, I attempted to estimate the potential harms of the workup, treatment and surgeries.

I also analyzed patient and CT factors to judge whether there were any clear variables that might affect the frequency of extracolonic findings or the likelihood of recommending additional imaging or treatment. When reported, the prevalence of extracolonic findings in those determined to have colorectal cancer was considered separately, since these patients require an abdominal CT scan for staging purposes.

Given the subjectivity of interpreting CT colonography scans for extracolonic findings, I decided that the heterogeneity of findings would preclude a quantitative data synthesis. The heterogeneity of included studies was not formally assessed but I did collect information on the various methodologies and their risks of bias to provide some indication on the potential sources of variability. For completeness, I used PRISMA reporting guidelines for systematic reviews.[80,81]

I did not carry out a formal assessment of potential publication bias. I did, however, assess published studies for selective reporting of results by assessing the outcomes reported in the methods section compared to the results section.

Strength of Evidence:

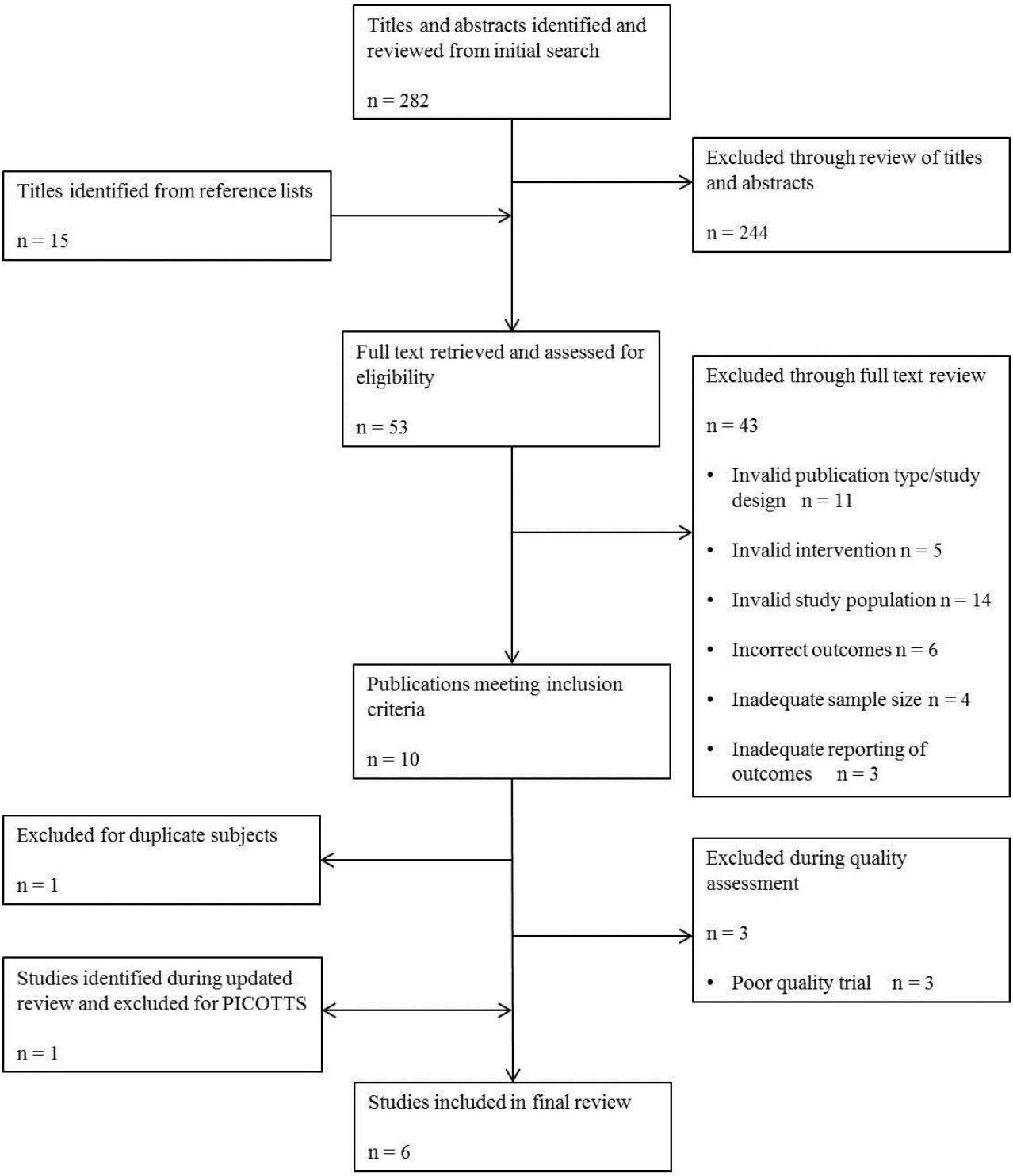
To assess the strength of evidence, I applied the approach used by the USPSTF.[78] When applicable, I assessed each of the six critical appraisal questions used by the USPSTF (see Appendix E). Since this review considered no comparators, I considered well-done, large, prospective cohort studies to be of near-equal value to randomized controlled trials. By evaluating the internal and external validity grades of individual studies, I assessed the aggregate internal and external validity of the body of evidence for each key question. I judged the consistency of evidence by looking at the variability in outcomes between studies and whether there were clear differences in study methodologies or populations that accounted for the differences. Lastly, I assessed directness by asking how well study populations, interventions and outcome measures fit my key questions. In other words, I evaluated how generalizable each study was to a typical screening population receiving CT colonography for colorectal cancer screening. Based on these elements, I graded the quality of evidence as high, moderate, low or very low. The full description of these ratings can be found in Appendix F.

Results

Searches, performed between March 4, 2012 and March 8, 2012, identified 282 titles and abstracts, including 44 past reviews, commentaries or letters. An additional 15 studies were identified by hand-searching these previously published reviews, editorials and commentaries.

After both reviewers performed a title and abstract review, 53 titles remained for full text review. During the full text review, 43 studies were excluded for invalid publication type/study design (n = 11), wrong intervention (n = 5), invalid study population (n = 14), incorrect outcomes (n = 6), poor quality (n = 4) and inadequate reporting (n = 3). After an extensive review of reporting and quality criteria, an additional 3 studies were excluded for inadequate follow-up of patients following CTC. Disagreements between reviewers on inclusion/exclusion of 6 studies were settled by consensus. Two studies[82,83] performed at the University of Wisconsin had overlapping dates of enrollment. Study authors concluded that these studies had overlapping datasets, so I decided to exclude the study published in 2010[83] since the reporting of outcomes was more complete for my population of interest in the 2008 study.[82] After removing this duplicate study, 6 cohort studies (2 prospective and 4 retrospective) met my final inclusion criteria and were included in the systematic review. An updated search performed on May 28, 2012 identified 4 new studies, which were all excluded for incorrect PICOTTS. A flow diagram of my search results and exclusion of studies is displayed in Figure 1 below.

Figure 1. Flow diagram of search strategy



Included Studies:

The six studies included a total of 6,316 subjects with a mean age of 59.9 years old (range 26-90). All of these six studies were cohort studies (4 retrospective, 2 prospective), five were conducted in the United States and one was based in Australia. The studies included subjects with varying levels of risk for colorectal cancer. Table 3 contains summaries of these studies' characteristics.

Internal Validity

I assessed the quality of included studies using the USPSTF criteria for internal validity including measurement bias, confounding bias and selection bias.

Included studies had varying potentials for measurement bias resulting from the CTC technique and method of reading scans. Equality of measurement was judged based on the use of a single CTC technique and standardized system of clinical importance. Half of studies included CTC scans performed on different scanners with varying radiation dose and slice thickness, increasing the risk of measurement bias. In addition, only half of studies used a standardized system for judging the clinical importance of ECFs.

I judged the validity of measurements based on the use of a classification system based on the system of judging clinical importance, the exclusion of previously-diagnosed ECFs and the method of following ECFs. Almost no studies used valid measurement criteria for reading and interpreting CTC scans. Only one study[60] assessed each lesion's likelihood of benefiting from diagnosis or treatment. Furthermore, several studies failed to exclude previously diagnosed ECFs, which are likely to be treated differently than newly diagnosed findings. Furthermore, only a few studies had a method of follow-up that was likely to capture all resulting workup.

Most studies relied on their institution's electronic medical record without accounting for investigations, treatments or procedures delivered at other institutions.

The reliability of measurements was also highly variable. Most studies employed adequately trained radiologists. However, only three studies[60,84,85] used duplicate reading of CTC scans. Only one of these studies[86] specified that the two radiologists performed their interpretations of scans independently. Moreover, only two studies[33,85] explicitly stated that radiologists were blinded to past radiological scans and patient history.

Potential confounding bias was noted in two studies[82,84] which reported higher average age, more comorbidities and a higher likelihood of intracolonic findings among those with ECFs. These same studies noted that older, sicker populations were less likely to receive clinical workup for their ECFs. While the risk of confounding bias is likely to be small, there was not enough information provided to determine its magnitude in most studies. Three out of the six studies[33,85,86] failed to report any information on potential confounders and the remaining studies reported only a few relevant variables. Patient enrollment was unlikely to contribute bias, as most studies consecutively enrolled individuals referred to their institutions for screening CTC. No studies provided information on the subjects who were lost to follow-up (i.e. did not receive all or part of their clinical workup at the same institution), making it difficult to assess the potential for selection bias introduced by differential loss to follow-up. However, all studies were able to follow at least 70% of their population.

The full quality assessment and final grades of internal validity can be found in Appendix G below.

Populations

Almost all studies included a group of consecutive patients who took part in their institution's CTC screening program. The only study that did not enroll consecutive patients[84] invited a randomly-selected group of patients from the community to participate. One study[60] required that patients be referred from a gastroenterology clinic in the local area and, similarly, another recruited patients referred to their institution for CRC screening, guaiac-positive stool or incomplete colonoscopy.

Settings and Description of Radiologists

These studies were predominantly conducted in the United States, with the exception of one study conducted in Australia.[84] All studies were conducted at academic medical centers, including one[60] conducted at a military medical center.

The degree of CTC-specific training varied greatly between studies. Two studies[60,86] enrolled participating radiologists in a training program in which they read 50 pathology-proven CTC cases, in accordance with the American College of Radiology's recommendations for CTC training.[87] Other studies[82,84] reported that participating radiologists had reached this recommended threshold through their clinical experience. The remaining two studies[33,85] stated that their radiologists were experienced and board-certified. However, these studies did not report the radiologists' CTC-specific training.

Half of the six included studies[60,84,85] had multiple radiologists review the CTCs and this review was conducted independently in only one of these three studies.[85] This review was performed retrospectively on the initial reads of CTC scans. Those subjects with "clinically important" (C-RADS E3 or E4) extracolonic findings discovered retrospectively were reassessed to confirm the finding and determine why it was missed initially. Two studies[33,85] required that

radiologists were blinded to patient's previous imaging and medical history. In addition, none of the included studies reported whether CTC readers were blinded to the purpose of their project. In fact, several studies[82,84,86] employed study authors to read and interpret CTCs for the study.

CTC Parameters

All studies used multidetector CT scanners (4-, 8-, 16-, 64- and/or 124-detector rows) and some used single-slice detector CT scanners as well.[33,85] Each study scanned patients in both the supine and prone positions and none used IV contrast material. Studies showed wide variations in slice thickness, reconstruction interval and pitch. Radiation levels also varied, with currents ranging from 40-100 mA. Each study used a CT scanner with a peak voltage of 120 kV. Only one study[84] calculated the total radiation dose, reporting that each scan resulted in less than 5 mSv (total effective body dose)[2].

Three of the studies[60,82,84] used a standardized CTC technique for all subjects. Two of the studies without a standardized technique[85,86] were retrospective and had a range of CTC parameters corresponding to the ones used by different radiologists in their institution. In the remaining study,[33] subjects underwent multi- or single-detector scans. The currents of the multidetector scans were adjusted to match the image noise of the single-slice technology.

Definitions of Clinical Importance of ECFs

Most of the included studies[60,85,86] applied the C-RADS classifications published by the Working Group on Virtual Colonoscopy.[3] These were applied retrospectively in these studies, with the exception of one[60] which began applying it prospectively while their study was ongoing. Two of these studies classified lesions further – one classified them based on their need for

additional workup[85] and the other labeled lesions that were particularly high-risk.[60] Two of the studies that did not use C-RADS were conducted before the Working Group on Virtual Colonoscopy published this classification system. These studies[33,84] classified extracolonic lesions based on their need for further workup, similar to the C-RADS system. In the final study,[82] radiologists prospectively labeled findings as moderate or greater importance or minimal or no potential importance depending on the need for further workup or clinical importance of the findings themselves, if diagnostic. The full descriptions of clinical importance classification systems used are found in Table 4 below.

Despite the homogeneity in classification systems, similar clinical findings were often classified differently in different studies. For instance, lung nodules were described as both moderate- and high-importance. Most studies did not further characterize lung nodules by size or appearance (i.e. calcification or speculation). In addition, one study[60] characterized cystic masses of the ovary as high importance but classified complex ovarian cysts as moderate importance. Similarly, cystic pancreatic lesions were classified as high-importance[33] while pancreatic cysts were labeled as moderate importance[60]. Osteoblastic and osteolytic bone lesions were classified as both moderate and high clinical importance. Conflicting classification was also applied to mesenteric lymph nodes and splenic artery aneurysms. As was the case with pulmonary nodules, lesions of the kidneys, liver, adrenal glands and ovaries were inadequately characterized by study authors. Lastly, studies rarely provided an explicit definition for AAAs. One study[84] reported at least 1 AAA with a diameter < 3 cm, others[85] only reported aneurysms ≥ 3 cm and another study[60] separately reported "high-risk" AAAs as those with a diameter ≥ 5 cm.

Method of Following ECFs

The studies had widely varying study lengths ranging from a mean follow-up time of 1 to 2 years. Only one study[84] set a defined follow-up duration (2 years) from the time of CTC. Other studies based their follow-up duration and clinical endpoints. For instance, one study[86] searched subjects' medical records solely to confirm completion of the imaging studies recommended during the initial CTC read. Another study[33] focused only on post-CTC imaging studies and did not report on subsequent clinical appointments, biochemical investigations, medical treatments, surgeries or other invasive procedures. On the other hand, four studies[33,60,82,84] reported on the surgeries performed subsequent to CTC. Some of the retrospective studies tried to collect information on all medical workup performed subsequent to the subject's CTC, but their methods for collecting this information were not always clear. Most studies failed to explicitly report a duration or specific endpoint that would end their follow-up of ECFs.

Some studies reported the follow-up investigations received by all study subjects. Others limited the scope to subjects with high clinical importance findings (i.e. C-RADS E3 and E4) lesions since, by definition, lesions below this threshold were not supposed to receive workup. One study[84] determined that subjects follow up with a radiologist, general practitioner or appropriate specialist depending on the clinical importance of their finding (e.g. specialist for high clinical importance lesions) and the recommended workup (e.g. radiologist for imaging). In another study,[85] information on ECFs was passed on to subjects' primary care physicians at the time of the initial CTC and all subsequent medical workup initiated by the PCP was collected from subjects' medical records.

The included studies used several different sources of information to determine what follow-up investigations occurred. For instance, most[33,60,82,85,86] employed a review of subjects' electronic medical records to gather information about workup of ECFs. Only one of these

studies[60] reported the percentage of subjects whose subsequent workup would be captured by this system. For the others, it was unclear how often subjects received follow-up that would not be captured by their electronic medical record. Conversely, one study[84] had all subjects receive their follow-up exams at a single institution thus ensuring that they could account for all subsequent imaging, procedures and treatments.

Frequency of Incidental Findings

The method of tallying extracolonic findings varied among studies. For instance, some studies included previously detected extracolonic findings while others excluded them or reported them separately. One study[84] included only extracolonic findings that had changed significantly since the last time they were imaged. In addition, other studies only reported extracolonic findings they deemed to be of high or moderate clinical significance. The full description of methods of tallying ECFs and number of ECFs detected is included in Table 5 below.

Four of the six studies[33,60,84,86] reported the frequency of subjects with extracolonic findings. Among these studies, the frequency of subjects with at least one ECF during their initial CTC ranged from 27.2% to 68.9% (mean 49.3%). Two of these studies[33,84] reported the total number of ECFs, ranging from 1.2 to 1.8 ECFs per subject with an extracolonic lesion detected. Three of the other studies[60,83,86] did not report the overall frequency of ECFs at all. The other[85] only reported the number of ECFs for the entire population, which included symptomatic and high-risk subjects. This study reported that among all subjects, 272 of 376 (72.5%) had at least one extracolonic finding. Among these 272 subjects, investigators found 520 incidental lesions (average 1.9 ECFs/subject).

Two studies reported the number of extracolonic findings found that were missed on the initial CTC. One study[85] reviewed all initial CTC scans and reports, stating that 144 E2, 29 E3 and 4 E4 lesions were missed by the initial reader. Based on these numbers, they calculated an 8.8% miss rate for E3 and E4 findings. Another study[33] reported the ECFs not reported on the initial CTC that were discovered during the subsequent radiological workup. They noted 44 lesions (23 of medium- or high-importance) that were not reported on the initial CTC report. Their retrospective read of the initial scan was able to see 25% of these findings on the initial CTC. The remaining 75% could only be seen on scans taken as part of the radiological workup.

Frequency of Clinically Important Lesions

The method of determining clinical significance varied somewhat among studies. Study authors sometimes based clinical importance on the lesion's characteristics and some based it retrospectively on the final diagnosis. For studies using the C-RADSs classification system (Table 1), I considered E4 findings to be high significance and E3 findings to be moderate significance. By definition, E2 lesions do not require clinical workup[3] and therefore were not included in either of these groups. The methods of reporting ECFs and reported frequencies can be found in Table 5 below.

In the four studies reporting high-importance lesions separately,[33,60,84,85] their frequency ranged from 1.5% to 10.4% (mean 5.6%). In one of these studies (frequency of high-importance findings = 7.4%),[84] there was no separate category for moderate importance extracolonic findings. Other than this study, all others reported the combined frequency of moderate- and high-importance ECFs. The frequency of moderate- and high-importance findings ranged from 4.4% to 37.3% (mean 15.5%).

To address the heterogeneity between classification systems, I also looked at the number of abdominal aortic aneurysms, diagnosed early (N0M0) cancers, total cancers and lesions suspicious for cancer. In the three studies[33,60,84] that reported the number of abdominal aortic aneurysms diagnosed during the initial CTC the frequency of AAAs ranged from 0.088% to 1.2%. The study[84] reporting a frequency of 1.2% (5 of 432 subjects) stated that they found 6 AAAs with diameters of 2.8 to 4.5 cm. Since I defined AAAs as ≥ 3 cm and this study did not report the individual diameters of each aneurysm, only 5 of these AAAs were counted although there might have been fewer that were ≥ 3 cm. In the study including symptomatic patients, there was no separate reporting of outcomes for their average-risk population. However, they reported a newly diagnosed AAA in 1 subject (0.19%). The remaining two studies either stated total number of aneurysms in the aortoiliac system[82] or did not report individual outcomes.[86]

Four studies[33,60,82,84] reported the number of newly discovered cancers, which were diagnosed in 0.23% to 0.88% of subjects (mean 0.45%). The frequency of N0M0 cancers was discussed in two studies,[60,84] which reported early stage cancers among 0.13% and 0.23% of subjects. The frequency of lesions suspicious for malignancy had wide variation and was often unclear due to poor characterization of lesions. The complete list of diagnosed AAAs, malignancies and suspicious lesions are included in Table 7 below.

Chin et al.[84] reported that 5 subjects (1.2%) had AAAs, 1 subject (0.23%) had a newly discovered cancer and no additional subjects had lesions suspicious for malignancy. This newly discovered cancer was a noninvasive renal cell carcinoma. Veerappan et al.[60] reported AAAs in 2 subjects (0.088%) and early-stage cancers in 3 subjects (0.13%), including one stage 1a lung adenocarcinoma and two stage 1 renal cell carcinomas. In addition, this study reported subjects with stage IIIb nodular lymphoma, recurrent bronchoalveolar carcinoma and stage II renal cell

carcinoma, totaling 6 subjects (0.26%) receiving cancer diagnoses. Gluecker et al.[33] reported 4 subjects (0.59%) with AAAs and 6 subjects (0.88%) with cancer diagnoses, including 1 squamous cell carcinoma of the lung, 1 renal adenocarcinoma, 1 renal oncocytoma and 3 ovarian serous cystadenomas. In addition, they reported 77 lesions suspicious for malignancy, representing 8.9% of all ECFs. This study did not report the number of N0M0 cancers.

Clinical Implications of Incidental Findings

The duration, methodology and reporting of follow-up varied greatly among studies. The number of subjects requiring some workup ranged from 2.0% to 8.7% (mean 5.4%). After excluding studies that did not report surgeries or invasive procedures,[85,86] the mean increases to 7.4%. One study[33] did not state how many subjects required workup but did report the number of individual tests and procedures required.

Most studies reported primarily the imaging workup required. CTs and ultrasound scans encompassed the bulk of recommended or performed imaging tests, making up 52.8% and 35.0% of all imaging tests, respectively. The complete list of imaging studies can be found in Table 8 below. In addition to the imaging reported in Table 8, Veerappan et al.[60] noted that subjects received 6 bone scans, 3 upper endoscopies and 1 bronchoscopy. There were also some scans reported in Pickhardt et al.[82] that are not listed in Table 8, including 1 skeletal scintigraphy scan, 2 renal scintigraphy scans and 3 small-bowel capsule endoscopies.

The four studies that listed surgeries reported that, of the 370 subjects requiring workup, 48 surgical procedures were performed. The majority of reported surgeries were for AAA repair or treatment of suspected malignancy. The full summary of these findings can be found in Table 6 below.

There were also several procedures for lesions that were ultimately found to be benign. One study[84] specifically reported this number, stating that 75% (24 of 32) of findings receiving workup were ultimately diagnosed as benign. Another study[82] reported that while adnexal lesions accounted for 45% (10 out of 22) of all surgeries, all lesions proved to be benign and all 10 liver lesions receiving contrast-enhanced CT were eventually diagnosed as benign. The same study stated that pulmonary nodules discovered on CTC led to chest CT in 6 subjects, CT-guided lung biopsy in 4 subjects and thorascopic resection in 2 subjects. One of these pulmonary nodules was found to be malignant. In a separate study,[60] none of the 8 patients receiving surgery for a pelvic mass ultimately had a malignancy and one of the 2 AAA repairs were on aneurysms they considered to be "low risk." However, the majority of studies did not explicitly report the number of potentially unnecessary procedures performed for benign findings.

There were many other clinical outcomes that were omitted from the majority of these studies. For instance, only two studies[33,84] reported the number of medical treatments required. One of these studies[84] reported that no subjects required medical treatment. The other study[33] reported that one subject required chemotherapy for thyroid cancer metastases to the lungs and another received antihelmintic treatment for ileal ascaris.

An especially glaring omission was discussion of harms from this additional workup, especially considering the description of its potential benefits. For instance, only one study[82] discussed the complications resulting directly from workup, stating that there were no reported complications of surgeries or invasive procedures. In addition, no studies calculated subjects' exposure to ionizing radiation during the subsequent radiological workup or discussed the potential harms of surgeries for benign findings. These surgeries carry the possibility of complications but also the inconvenience and cost of a surgical procedure and the resulting

recovery. Lastly, the studies failed to collect data on the potential psychological harms from extracolonic findings, such as the anxiety of a potentially serious diagnosis.

Funding Sources and Potential Conflicts of Interest

None of the six studies reported their funding sources and only one included study[86] reported conflicts of interest. This particular study reported that authors had no potential conflicts of interest. However, the study that was excluded for duplicate data[83] reported relevant conflicts of interest for some of the authors from the study I included.[82] These authors were consultants for companies that develop computer software for CTC and the cofounders for a company that provides educational materials and trainings on CT colonography.

Generalizability of Studies:

I assessed the external validity of each study using criteria developed from the USPSTF[78] and the GRADE working group.[79] I judged the generalizability of study populations, settings and interventions to a typical screening population.

I defined a typical population as one that was primary asymptomatic, at average-risk of colorectal cancer and within the recommended screening ages of 50-74. Three studies[60,82,84] were rated as having good population external validity. The one study with a fair rating[86] primarily enrolled screening patients but also included those with an incomplete colonoscopy. I gave the remaining two studies[33,85] poor generalizability ratings due to their inclusion of symptomatic and high-risk patients without any separate reporting of outcomes.

I defined a typical setting as some mixture of academic and non-academic institutions. All included studies were conducted in academic settings with no inclusion of community

hospitals. Therefore, these results of these studies are likely more generalizable to academic settings alone.

I defined a typical intervention as a non-contrast-enhanced, low-dose CTC interpreted by radiologists with varied levels of experience, but with the recommended 50 endoscopically-confirmed cases (if skilled at abdominal CT) or 75 cases (if unskilled at abdominal CTC).[32] None of the included studies used IV-contrast for their CTCs. The voltage of 120 kVp was consistent with those reported in a recent survey of screening CTC programs.[88] The effective tube current levels, however, were slightly higher than the surveyed programs, which all reported currents less than 50 mAs. In addition, almost all included studies in my review used highly trained radiologists to read and interpret CTCs. While this might increase the internal validity of their results, it might also make them less generalizable to most CRC screening programs, especially those in community settings. Furthermore, a few studies[60,86] retrospectively applied clinical importance classifications without blinding radiologists to subsequent medical or radiological history. This practice is not representative of the technique of reviewing and interpreting CTCs in clinical practice and thus might reduce the generalizability of these findings. As a result, the two studies that interpreted CTCs retrospectively, without masking, were given poor/fair generalizability scores for intervention.

Overall, I judged all the studies to have fair external validity, meaning that these studies differed a few ways from the standard CT colonography screening program. I conclude that it is moderately probable (50%-89%) that the experience with CTC described in this study would be attained in a typical screening population. All the generalizability grades can be found in Appendix H below.

Overall Results and Strength of Evidence

The first key question addressed by this review was the overall frequency of incidental lesions detected during screening CTC colonography. Based on the reported studies, there is evidence of low strength that at least 30% of subjects will have at least one extracolonic finding found on CT colonography. In addition, there is evidence of moderate strength that this proportion is higher among asymptomatic patients at high-risk of colorectal cancer.

The second key question assessed the frequency and clinical implications of moderate- and high-clinical importance extracolonic findings. There is evidence of low strength that 10% or lower of individuals receiving screening CTC will have high clinical importance findings. There is also evidence of low strength that the frequency of moderate or high frequency findings ranges from 10% to 30%. The majority of high importance findings and the vast majority of moderate/ high importance findings will eventually be diagnosed as benign. Moderate strength evidence supports the fact that roughly 2% of subjects receiving CTC will have an incidentally discovered AAA or extracolonic cancer. There is also moderate strength evidence that almost all findings of high importance will receive some clinical workup. The evidence strength for workup required for moderate/high importance findings or all ECFs is very low.

Lastly, I set out to find out how studies were determining the clinical significance of extracolonic findings. Since the development of the C-RADS criteria in 2005,[3] most studies have used this classification system. However, despite using a uniform system, studies show variability in the classification of some lesions.

Table 3. Description of studies

Author (Year)	Study Design	Country	Number of Subjects (N)	Age of Subjects (range &/or S.D.)	Risk of CRC	Percent Symptomatic	Description of CTC Readers
Chin et al. (2005)[84]	Prospective cohort	Australia	432 (230 M, 202 F)	Mean 59 yrs. (range 50-69)	Average risk	0%	2 readers (concurrent review) At least 100 CTCs reviewed previously Academic center
Pickhardt et al. (2008)[82]	Retrospective cohort	US	2195 (996 M, 1199 F)	Mean 58 yrs. (S.D. 8.1)	NR[a]	0%	1 reader Range of 1-5 years interpreting CTC studies Academic center
Flicker et al. (2008)[85]	Retrospective cohort	US	527 (210 avg. risk) (genders not reported[b])	Mean 61 yrs. (range 26-89)	Both average and high-risk[c]	0%[d]	2 readers (reinterpretation of original read specifically for this study) Abdominal imaging fellow and experienced abdominal imaging attending Academic center

NR: Not reported
[a] Stated that the subjects were "representative of typical U.S. screening population" but did not have exclusion criteria for high-risk
[b] For the 272 subjects with E2-E4 findings there were 101 males and 171 females
[c] Separate analysis for average risk patients
[d] Patients within low risk group were not symptomatic

Table 3 (continued). Description of studies

Author (Year)	Study Design	Country	Number of Subjects (N)	Age of Subjects (range &/or S.D.)	Risk of CRC	Percent Symptomatic	Description of CTC Readers
Gluecker et al. (2003)[33]	Prospective cohort	US	681 (426 M, 255 F)	Median 64 yrs. (range 26-89)	High risk[a]	0%	1 reader Each with 10 years of practice Academic center
Macari et al. (2011)[86] [b]	Retrospective cohort	US	204 (110 M, 94 F)	Mean 53 yrs. (range 41-64)	Mixed risk[c]	10.9% (all with guaiac+ stool)	1 reader Readers had 8, 14, 30 and 32 years of experience and either completed CTC course or had read > 100 CTCs Academic center
Veerappan et al. (2010)[60]	Retrospective cohort	US	2277 (1207 M,1070 F)	Mean 61 yrs. (S.D. 11)	NR[e]	NR[f]	2 readers (each scan with significant findings reinterpreted by radiologist) All had routinely read CTCs, reviewer had read more than 5000 CTCs Academic center

NR: Not reported
[a] First-degree relative with colorectal cancer or polyps, prior personal history of polyps or colorectal cancer, or new onset of asymptomatic anemia
[b] Only included younger cohort since this group fit into the inclusion criteria for age distribution
[c] Included subjects with guaiac-positive stool
[d] Subjects referred for screening but no exclusion criteria for high risk
[e] Likely to be primarily average-risk subjects since authors state that high-risk patients are generally excluded from this screening program
[f] Assumed to be asymptomatic because authors stated that these were screening patients and hematochezia part of exclusion criteria

Table 4. Definitions of clinical importance of extracolonic findings

Author (Year)	Definition for Clinical Importance of ECFs
Gluecker et al. (2003)	- **High importance:** requiring surgical treatment, medical intervention and/or further investigation during that patient care visit - **Medium importance:** did not require immediate treatment but would likely require investigation, recognition or treatment at a later time - **Low importance**: benign and unlikely to require further medical treatment or additional workup
Chin et al. (2005)	- **Clinically relevant**: "required medical or surgical attention, or further hematological, biochemical, and/or radiological investigation after assessment of several factors..." including "the patient's medical history and prior investigations taken, the radiological appearance of the CT findings, and the relevance of these findings in the current clinical context of the patient." - **Non-clinically relevant**: those judged not clinically relevant did not have further diagnostic testing or new treatments initiated
Flicker et al. (2008)	- C-RADs classification[a] - E3 and E4 lesions further classified into three categories: (1) previously imaged, (2) additional imaging (performed to evaluate CTC finding) and (3) received no subsequent workup
Pickhardt et al. (2008)	- **Moderate or greater potential importance:** need for further workup or, when CTC is diagnostic, findings considered moderate or greater clinical importance (authors stated that these corresponded to E3 or E4 C-RADS findings) - **Minimal or no potential importance:** no need for further workup or, when CTC is diagnostic, findings considered low clinical importance
Veerappan et al. (2010)	- C-RADs classification[b] - Classified high-risk lesions as those discovered to be a malignancy on the basis of pathologic findings or a large abdominal aortic aneurysm (≥ 5 cm) confirmed in the operating room
Macari et al. (2011)	- C-RADs classification[a]

[a] C-RADS classification system (provided in Table 1) applied retrospectively
[b] C-RADS classification system (provided in Table 1) applied retrospectively for scans performed before January 1, 2006 and prospectively after January 1, 2006

Table 5. Primary outcomes

Author (Year)	Follow-up Duration	Excluded Previously Detected ECFs?	Frequency of Incidental Findings and total # of ECFs	Frequency of High Clinical Importance Findings[a]	Frequency of Moderate or High Clinical Importance Findings[b]	Imaging Performed or Recommended (%ECFs & Total)	%ECFs Needing Surgery
Chin et al. (2005)	2 years after CTC	Yes[c]	**27.2%** (118 of 432 subjects) 146 ECFs in 118 subjects (avg. 1.2 ECFs/subject)	**7.4%** (32 of 432 subjects)[d]	NR[d]	All ECFs: **21.2%** (31 of 146) High imp: **96.9%** (31 of 32)	**0%**
Pickhardt et al. (2008)	Avg. 18 mos. (range 3 mos.-24 mos.)	Yes	**NR** # ECFs not reported	NR	**8.6%** (189 of 2195 subjects)	All ECFs: NR[e]	NR[f]
Flicker et al. (2008)	E3 lesions: avg. 26 mos. (range 9-76 mos.) E4 lesions: avg. 18 mos. (range 20 days-48 mos.)	No (but reported separately)	**NR**[g] All subjects combined: 72.5% (272 of 376 subjects) All subjects combined: 520 ECFs in 272 subjects (avg. 1.9 ECFs/ subject)	**2.9%** (6 of 210 subjects)	**17.1%** (36 of 210 subjects)	All ECFs: NR[g] E3 & E4: **35.5%** (11 of 31) E4: **100%** (5 of 5)	NR[h]

NR: Not reported

[a] Clinical importance defined by study author, some prospectively based on the lesion's characteristics and some retrospectively based on the final diagnosis. For studies using C-RADS, I reported frequency of patients with E4 findings. See Table 4 above for definitions.

[b] For studies using C-RADS, I reported combined frequency of patients with E3 or E4 findings.

[c] Included previously detected ECFs with significant change

[d] No separate category for moderate significance findings (see Table 4 above)

[e] Number of ECFs not reported. 7.2% (157 of 2195) had workup recommended by radiologist and 5.2% (115 of 2195) completed recommended workup

[f] Total of 22 surgeries reported

[g] Not reported for the separate analysis for asymptomatic patients with average risk of colorectal cancer

[h] Only reported surgeries that might have resulted in benefit, which included 1 AAA repair and 2 partial nephrectomies for renal cell carcinoma

Table 5. Primary outcomes (continued)

Author (Year)	Follow-up Duration	Excluded Previously Detected ECFs?	Frequency of Incidental Findings and total # of ECFs	Frequency of High Clinical Importance Findings[a]	Frequency of Moderate or High Clinical Importance Findings[b]	% ECFs with Imaging Performed or Recommended	% ECFs Needing Surgery
Gluecker et al. (2003)	At least 12 mos. (range 401-1513 days)	No	**68.9%** (469 of 681 subjects) 858 ECFs in 469 subjects (avg. 1.8 ECFs/subject)	**10.4%** (71 of 681 subjects)	**37.3%** (254 of 681 subjects)	All ECFs: **12.8%** (110 of 858) High/Mod Imp: **56.1%** (110 of 196) High Imp: **98.9%** (87 of 88)	All ECFs: **1.3%** (9 of 858) High/Mod Imp: **4.6%** (9 of 196) High Imp: **9.1%** (8 of 88)
Macari et al. (2011)	NR	NR	**55.4%** (113 of 204 subjects)[c] # ECFs: NR	NR	**4.4%** (9 of 204 subjects)	NR	NR
Veerappan et al. (2010)	Mean 19.5 mos. (range 6 mos.-4 yrs.)	No (but reported separately)	**45.5%** (1037 of 2277 subjects) # ECFs: NR	**1.5%** (35 of 2277 subjects)[d]	**10.1%** (230 of 2277 subjects)[d]	All ECFs: NR E3 & E4:**86.1%** (198 of 230) E4: **88.6%** (31 of 35)	All ECFs: NR E3 & E4: **8.3%** (19 of 230) E4: **34.3%** (12 of 35)

NR: Not reported

[a] Clinical importance defined by study author, some prospectively based on the lesion's characteristics and some retrospectively based on the final diagnosis. For studies using C-RADS, I reported frequency of patients with E4 findings. See Table 4 above for definitions.

[b] For studies using C-RADS, I reported combined frequency of patients with E3 or E4 findings

[c] Separate analysis for younger cohort since this group fit into the inclusion criteria for age distribution

d This is excluding previously detected extracolonic findings

Table 6. Summary of high clinical importance extracolonic findings

				Early stage cancers (All cancers)								
Study Author	**N**	**Total ECFs/High clinical importance ECFs (%)**	**AAAs**	**Ovary**	**Lymphoma**	**Lung**	**Kidney**	**Liver**	**Pancreas**	**Gallbladder**	**Bladder**	**Other**
Chin et al. (2005)	432	27.2% / 7.4%	5[a]	0 (0)	0 (0)	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Pickhardt et al. (2008)	2195	NR / NR	?[b]	? (0)	? (3)	? (1)	? (3)	? (0)	? (0)	? (0)	? (0)	? (2)[c]
Flicker et al. (2008)	527 (210 avg. risk)	NR / 2.9%	?[d]	? [e]	? [e]	? [e]	? [e]	? [e]	? [e]	? [e]	? [e]	?[e]
Gluecker et al. (2003)	681	68.9% / 10.4%	4	? (3)	? (0)	? (1)	? (2)	? (0)	? (0)	? (0)	? (0)	? (0)
Macari et al. (2011)	204	55.4% / NR	?	?	?	?	?	?	?	?	?	?
Veerappan et al. (2010)	2277	45.5% / 1.5%	2	0 (0)	0 (1)	1 (2)	2 (3)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

NR: Not reported

[a] Reported 6 AAAs were diameter 2.8-4.5 cm. I defined AAAs as greater or equal to 3 cm but the study did not report each individual size meaning the actual figure could be 5 or less AAAs.

[b] Reported 12 aneurysms in the aortoiliac system

[c] Other: 2 abdominal metastatic cancer

[d] Among all (average- and high-risk) subjects, they reported 1 new AAA

[e] Among all (average- and high-risk) subjects, they reported 3 confirmed cancers (metastatic ovarian and 2 RCC)

Table 7. Summary of potentially clinically important extracolonic findings

Study Author	N	Total ECFs/High clinical importance ECFs (%)	Liver solid mass	Lung nodule	Kidney solid mass	Adrenal nodule	Pancreatic solid mass	Pancreatic cystic mass	Ovarian cystic or complex mass	Enlarged lymph nodes	Other
Chin et al. (2005)	432	27.2% / 7.4%	0	0	0	0	0	0	0	0	0
Pickhardt et al. (2008)	2195	NR / NR	?	?	?	?	?	?	?	?	?
Flicker et al. (2008)	527 (210 avg. risk)	NR / 2.9%	?[a]	?[a]	?[a]	?	?	?	?[a]	?	?[b]
Gluecker et al. (2003)	681	68.9% / 10.4%	2	26	?[c]	0	0	1	6	2	6[d]
Macari et al. (2011)	204	55.4% / NR	?	?	?	?	?	?	?	?	?
Veerappan et al. (2010)	2277	45.5% / 1.5%	0	4	15	1	1[e]	1[e]	7	3	1[f]

[a] Among all (average- and high-risk) subjects, reported 6 lesions suspicious for malignancy (2 solid renal lesions, 1 solid liver lesion, 1 lytic bone lesion, 1 ovarian mass and 1 lung nodule > 2 cm)
[b] Other: 1 lytic bone lesion
[c] Did not report solid kidney masses specifically, but did report 34 kidney masses
[d] Other: 1 low-attenuation liver lesion and 5 cystic liver lesions
[e] Only one pancreatic mass – did not report if it is solid or cystic
[f] Other: 1 retroperitoneal mass

Table 8. Required Workup for ECFs

Study Author	N	Total ECFs/High clinical importance ECFs (%)	Total N[a] (%)	US	CTs	MRIs	X-rays	PET scans	Surgeries	Non-surgical Invasive Procedures	Medical Tx
Chin et al. (2005)	432	27.2% / 7.4%	32 (7.4%)	26	10	0	1	1	0	0	0
Pickhardt et al. (2008)	2195	NR / NR	133 (6.1%)	64	59	11	10	2	22	19	?
Flicker et al. (2008)	527 (210 avg. risk)	NR / 2.9%	6 (2.9%)	?[b]	?[b]	?[b]	?[b]	?[b]	?[b]	?[b]	?[b]
Gluecker et al. (2003)	681	68.9% / 10.4%	?	46	41	0	12	0	8	?[c]	2
Macari et al. (2011)	204	55.4% / NR	4 (2.0%)	4	5	0	0	0	?	?	?
Veerappan et al. (2010)	2277	45.5% / 1.5%	199 (8.7%)	57	182	13	9	10	18	2	?

[a] The total number of patients receiving some type of workup
[b] Not reported separately for asymptomatic, low-risk subjects. Workup for all subjects included: 8 US, 9 contrast-enhanced CTs, 1 abdominal X-ray and 3 surgeries (only selected surgeries reported)
[c] Specific number not reported, but authors stated that many patients received treatment for renal and bladder calculi

Discussion:

Frequency of Extracolonic Findings

The included studies reported overall frequencies of extracolonic findings (ECFs) ranging from 27.2% to 68.9% (mean 49.3%). Combining study populations, 1737 of 3594 subjects (48.3%) had at least one ECF. Studies including patients with a personal history of polyps, a polyposis syndrome or a family history of colorectal cancer reported higher rates of ECFs. These results provide further support for the finding that risk of colorectal cancer is related to the frequency of ECFs, which was previously noted by Pickhardt et al.[1] This finding could have important implications for clinical practice, as the balance of benefits to harms and the cost-effectiveness might be different for screening and surveillance populations.

There are several other important factors in the reported frequencies of extracolonic findings. For instance, one excluded study[83] reported that their older cohort (mean age = 69.2 years) had a significantly higher percentage of extracolonic findings than the remaining screening population (mean age = 56.9 years). Subjects with comorbid diseases also had a higher frequency of ECFs.[84] These findings might be related, since older patients are more likely to have comorbidities. There are a few potential explanations for these findings. It is possible that older individuals are more likely to develop lesions that can be discovered on CT scans. Alternatively, radiologists might be more prone to identify or report such lesions for older patients.

Differences in radiologist practice and experience might also contribute to differences in ECF frequency. For instance, Chin et al.,[84] who reported the lowest frequency of ECFs, employed two CTC readers. The study with the next lowest frequency[60] used an experienced radiologist to review scans with significant findings. Conversely, only one reader interpreted the

CTC scans in the two studies with the highest reported frequencies of ECFs.[33,86] These findings support the fact that employing a second radiologist reduces the rates of false-positives. However, it is also possible that employing two radiologists could increase false negatives. Two studies in this review[33,85] collected information on lesions missed during the initial CTC. One of these studies[85] found 144 E2, 29 E3 and 4 E4 lesions that were previously unreported, resulting in an 8.8% miss rate per patient. The other study[33] reported that 44 lesions (3 high importance, 20 medium importance, 21 low importance) were missed on the initial radiological exam (25%) or found on subsequent radiological follow-up (75%).

Radiologist experience could also played a role, as studies report that more experienced radiologists are less likely to recommend additional imaging.[63] This review's results did little to strengthen or weaken this claim. In addition, unimportant (i.e. E2 and below) lesions might have higher variation since some radiologists choose not to report these findings. A previously-published review stated that 58% of surveyed research practices reported all extracolonic findings while the remaining 42% only reported significant lesions.[5] Another factor that could greatly affect ECF frequency is whether previously diagnosed ECFs are included. This is an important consideration since most physicians are only interested in finding previously undiagnosed lesions with CTC. One study[33] combined previously detected and newly diagnosed extracolonic findings. Not surprisingly, this study reported the highest frequency of ECFs.

My results provide no clear indication that CTC parameters, such as varying radiation dose or slice thickness, played a role in the observed differences in ECF frequency. Previous studies[2] have raised the possibility that smaller slice thickness and higher radiation doses might increase the frequency of ECFs. However, there are several confounding factors that make it difficult to determine the importance of slice thickness and radiation dose.

Classification of Clinical Importance

The six included studies either employed the C-RADS classification system or a similar system based on the likelihood of additional workup. Despite the similarities of these systems, there was considerable variability in the number of clinically important findings and the recommended workup. Based on the results of this review, it appears that these classification systems do little to reduce clinical variability, as they set out to accomplish. Therefore, poor dissemination of these classification systems among radiologists might not be the primary source of clinical variability. In addition, these systems do not appear to classify lesions based on their likelihood of improving patient outcomes. The classification systems might benefit from validation studies that could properly assess their likelihood of providing accurate prognostic information to patients.

Frequency of Clinically Important ECFs

Studies reported that 1.5% to 10.4% (mean 5.6%) of subjects had at least one finding of high clinical importance. Combining the populations, a total of 144 out of 3600 subjects (4.0%) had at least one high-importance finding. Studies reported that the frequency of subjects with at least one moderate or high importance lesion ranged from 4.4% to 37.3% (mean 15.5%). In total, 230 out of 2277 subjects (10.1%) had a moderate- or high-importance lesion.

Many of the same factors that contributed to the variability in the overall frequency of ECFs likely influenced the number of high importance findings reported. However, one of the biggest contributors to variability is the inconsistency of radiologists reviewing CTCs. The C-

RADS classification system was developed to reduce the variability of these findings, but there are three major reasons that classification systems might not significantly reduce this variability.

First, there might be inadequate dissemination of guidelines. In fact, a national survey reported that many radiologists are unaware of the published recommendations on CTC.[69] However, most studies in my review employed similar classification systems and employed trained radiologists. Therefore, poor dissemination of recommendations does not explain the large variability seen in my review.

The second possibility is that classification systems do not provide enough guidance for some extracolonic lesions. As shown in Table 1, the C-RADS classification system provides little guidance for determining the clinical significance of a lung nodule and does not mention skeletal lesions at all. In addition, the C-RADS system might be more helpful when paired with additional clinical guidance for working up ECFs, such as the White Paper published by the ACR's Incidental Findings Committee.[89]

Third, even with clear classification systems, inter-rater variability might remain. This is the most likely explanation for the variability seen in my study. Radiologist inter-rater variability is already well described in the literature. One randomized controlled trial of 50 asymptomatic, average-risk patients assigned half these patients to receive total-body scanning with CT and the other half to be followed clinically.[90] Two experienced specialty radiologists, including 2 abdominal radiologists, were provided guidelines on the clinical classification and workup for specific lesions. Of the subjects receiving CT scans, 64% (16 out of 25) had incidental findings. The radiologists disagreed on 9 of these 16 findings ($\kappa = 0.52$), corresponding to moderate agreement. While the full-body CT is not entirely comparable with CT colonography, this study identifies the potential for large practice variation even when clear guidelines are provided to

radiologists. It is possible that improved training, specifically on interpreting ECFs, could reduce some of this variability. In fact, experience in community CTC screening programs shows that experience alone is not a proper substitute for CTC-specific training.[64]

Reported AAAs and Extracolonic Malignancies

I also collected information on hard outcomes such as AAAs and extracolonic cancers, in order to circumvent the variability between classification systems and determine how many how many highly-important findings might benefit from early intervention. Studies reported that 0.09-1.2% of subjects were diagnosed with AAAs. Among the three studies reporting these outcomes, AAAs represented 8.6% of high importance findings. Of these 12 newly diagnosed AAAs, 4 (33.3%) required surgical repair during the reported follow-up period.

It is difficult to estimate the balance of benefits and harms from detecting AAAs from screening CTC, but previously published analyses provide some clues. According to estimates from the USPSTF, approximately 500 men aged 65 to 74 who are current or former smokers would need to be screened to prevent 1 AAA-related death. Among never-smokers of the same age range, approximately 1,800 would need to be screened to prevent 1 AAA-related death.[91] CT colonography is unlikely to provide additional diagnostic accuracy, as ultrasonography has a reported sensitivity of 95% and specificity of 100% for AAAs.

The potential benefits of screening for AAAs are more limited for women age 65 to 75, since their risk of AAA rupture is much smaller. In addition, the size of the screen-detected AAA has an influence on the balance of benefits and harms. For instance, little evidence supports benefits of surveillance or surgery for small AAAs (diameter 3-3.9 cm). The benefits of surgical repair are more pronounced for older patients with large AAAs (diameter $\geq$ 5.5 cm). One study

reported that surgery for older patients with large AAAs resulted in an estimated 43% reduction in AAA-related mortality, although no improvement in all-cause mortality.[75] Only one study in my review[60] included a separate category for high-risk AAAs (≥ 5 cm), reporting that 1 out of the 2 newly-diagnosed AAAs were high-risk. Therefore, I cannot conclude how many AAAs detected from screening CTCs are likely to have benefitted from detection and intervention.

Another important factor to consider is the frequency of screening for AAAs. Several studies have reported that the incidence of new AAAs in a period of ten years ranges from 0-4%. Furthermore, none of these incidental AAAs were larger than 4 cm in diameter, meaning they were less likely to benefit from early intervention.[92-95] Based on this evidence, it is unlikely that subsequent CTC exams would increase the number of high-risk AAAs detected. There are also potential harms of finding AAAs from screening, which were not addressed in the included studies. Open surgical repair of AAAs carries a 4-5% mortality and results in complications (e.g. myocardial infarction, respiratory complications, spinal cord ischemia and graft infection) for approximately one-third of surgical patients. Endovascular repair of AAAs, which has become increasingly popular, might have a lower risk of complications but also has less certain long-term effectiveness.[75] There are also potential psychological harms of screening for AAAs. For instance, one study reported significant decreases in quality of life scores for patients with AAAs receiving follow-up tests.[96] The same decreases in quality of life were not seen in a control group without AAAs. Unfortunately, no studies included in my review explored the possibility of psychological outcomes of being diagnosed with or receiving surveillance for an abdominal aortic aneurysm.

My included studies report that 0.23-0.88% of subjects were diagnosed with cancers and roughly a quarter of these lesions were early-stage cancers. Renal cell carcinoma (RCC) is the

cancer with the most robust evidence suggesting potential benefit from its incidental discovery. The incidence of renal cell carcinoma has increased dramatically with the rise of cross-sectional imaging and approximately 60% of RCC cases are detected incidentally. However, despite the drastic rise in its detection, the mortality rates for RCC have been steady over the past few decades.[89] One explanation for these trends is that many of these cancers would not have otherwise caused important symptoms (i.e. they represent overdiagnosis).[97,98] Some evidence supports the potential for overdiagnosis of RCC. For instance, while only 0.5% of individuals die from RCC, this cancer is detected in up to 2% of autopsies.[89] Another study investigating the progression of renal tumors reported that 14% regressed in size and 40% grew at such a slow rate that it would take more than six years for them to double in size.[99] They also reported that slow growing tumors were more common in elderly patients. Based on this evidence, of the 9 renal cancers reported in this review, it is likely that some would not have presented clinically during a patient's lifetime. As a result, some of these patients might have experienced surgery unnecessarily. In addition, even those that would have otherwise presented clinically might not have benefitted from early treatment. On the other hand, it is possible that some of these patients received benefits from early detection and intervention. Unfortunately, there is no way to determine which incidentally-detected RCCs are indolent and which are aggressive.

The screening CTC studies included in my review also reported 3 confirmed ovarian carcinomas and several lesions suspicious for ovarian cancer. Again, it is impossible to determine the exact balance of benefits and harms for women with incidentally-detected ovarian cancers. However, current evidence does not indicate that screening for ovarian cancer improves patient outcomes. The largest RCT to-date on this topic is the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. In this study, investigators randomly assigned 78,216

average-risk women to screening for ovarian cancer with CA-125 and transvaginal ultrasound or usual care.[100] After 13 years of follow-up, the two groups showed no differences in the stage of diagnosed ovarian cancer or the death rate from ovarian cancer. Conversely, there is more evidence on the harms of screening for ovarian cancer, including the risks of false-positives and complications from the resulting workup. For instance, one-third of the women with false-positives in the PCLO trial underwent oophorectomy.[100] In the entire screening group, there were 20 surgeries performed for every screen-detected ovarian cancer and approximately 20% of these surgeries resulted in some complication. The potential for false-positive ovarian lesions was also demonstrated in my review. One study[82] included in this review reported that all 10 surgeries for adnexal lesions revealed benign findings. Another study[60] reported that CTC revealed incidental ovarian lesions that led to 8 surgeries, all of which revealed benign masses.

My included CTC studies only reported a few diagnoses of lung cancer. Based on the recent randomized controlled trial,[101] which reported a 20% reduction in mortality, it is possible that some benefit might arise from discovering incidental lung lesions. However, the reported benefits of screening for lung cancer were found among a highly-selected group of individuals aged 55 to 74 with at least a 30 pack-year history of smoking. In addition, these benefits were coupled with severe potential harms, including the deaths of 16 patients (6 of whom did not have lung cancer) within 60 days of the invasive diagnostic procedure. Therefore, it would be harder to justify the large harms of this screening in the broad population of adults age 50-74 receiving screening CTC. Screening for other cancers, including those of the pancreas, bladder and adrenal glands, is likely to provide minimal benefit at best, if not result in some net harm to patients.

Despite the paucity of supporting evidence for population-wide screening for AAAs and extracolonic cancers, it is likely that the vast majority of benefit from detecting ECFs during

screening CTC come from finding these two groups of lesions. However, these two lesions represent a minority of findings deemed to be of high clinical importance. This suggests that my classification systems are too sensitive, picking up too many benign findings that lead to unnecessary and potentially harmful workup. Sliding down the receiver operating characteristic (ROC) curve to decrease sensitivity and increase specificity might improve the cost-effectiveness of CTC and improve patient care. This could be accomplished by limiting the reporting of extracolonic lesions to AAAs and high-yield extracolonic cancers. Hassan et al.[102] suggested a similar approach in their cost-effectiveness analysis of CTC by only looking for intracolonic lesions, AAAs and extracolonic cancers. But, in this analysis, detecting extracolonic cancers contributed only 2% to life-years gained (LYG) but 55% of the costs of CTC. This equates to $1.5 million per LYG, likely well above the expense considered to be cost-effective.[103] The cost-effectiveness of AAAs was much better, accounting for 16% of LYG while contributing only 6% to the overall costs. However, many organizations, including the USPSTF, already recommend one-time screening for AAAs with ultrasound for older men who have ever been smokers.[91] Based on the evidence, the survival benefits of AAA screening are greatest for older males who are former or current smokers. In addition, the evidence suggests limited effectiveness of repeat screening for AAAs. Thus, the majority of life years will be gained during the first CTC, with large drop-offs in benefit for subsequent CTCs. Furthermore, the psychological harms of screening (e.g. anxiety), however small, are likely to occur for each round of screening. Therefore, while there are likely mortality benefits from detecting AAAs with screening CTC, this approach might not provide an optimal ratio of benefits to harms or maximize cost-effectiveness. In fact, the cost-effectiveness analysis by Hassan et al.[102] compared screening CTC

to optical colonoscopy alone, without considering the alternative of one-time screening with ultrasonography.

To account for the number of findings that might eventually be diagnosed as cancer, I reported the number of lesions suspicious for cancer. Unfortunately, only three studies reported the numbers of such lesions and these studies had huge variability in follow-up and reporting. Chin et al.[84] only reported the final diagnoses and thus had no lesions suspicious for malignancy. Veerappan et al.,[60] on the other hand, listed all initial CTC interpretations of ECFs in addition to their final diagnoses, reporting 32 lesions suspicious for malignancy among 2277 subjects. Another reported 43 suspicious lesions among 681 subjects. The majority of reported lesions were found in either the lungs or kidneys. The imprecision in these numbers makes it difficult to estimate the percentage of screening CTC patients who might develop cancer after the study follow-up period. In addition, most of these lesions will be false positives meaning a minority are likely to represent truly clinically important findings. For instance, a recent study stated that approximately 24 of 25 lung nodules detected during screening with low-dose CT were false-positives.[101]

Clinical Workup of ECFs

Among all studies of screening CTC, approximately 1 in 20 subjects required some clinical workup. This number likely underestimates the true frequency of clinical workup as many studies focused solely on radiological workup or workup performed at their institutions. In the three studies that reported surgeries and invasive procedures in addition to radiological workup, roughly 1 out of every 14 subjects required some clinical workup.

Since many studies did not report clinical workup or only reported selective outcomes, a more accurate range might be provided by the frequency of moderate/high and high-clinical importance findings. According to the C-RADS classification system, E4 (high importance) lesions require immediate workup or intervention while E3 (moderate importance) findings sometimes require workup, depending on local practice and patient preference. E2 findings and below require no workup or intervention, by definition. Using these assumptions, the frequency of subjects requiring some clinical workup falls somewhere between 5.5% and 15.5%. In other words, 1 out of every 6 to 18 subjects receiving screening CTC will require some clinical work-up.

The relative frequency of final diagnoses and suspicious lesions in these studies is likely to be influenced by the length of follow-up and whether the study was retrospective or prospective (reported in Table 5 above). For instance, Veerappan et al.[60] performed an extensive retrospective review of patient's records to determine a final diagnosis for the majority of suspicious lesions found on screening CTC. Conversely, Macari et al.[86] did not report the length of follow-up and were unable to locate several electronic medical records, which was their lone source of information for clinical workup.

Paradoxically, some of the factors that increase the frequency of ECFs might reduce the chance of receiving some clinical workup. For instance, in the study looking at the effect of age on CTC results,[83] only 50.6% of the older cohort with an ECF received workup compared with 57.3% in the general screening population. Similarly, Chin et al.[84] reported that while patients with comorbidities were more likely to have an ECF discovered on the initial CTC, they were much less likely to receive clinical follow-up. Physicians might be less willing to perform further testing on older populations that are more likely to have comorbidities. This may be because

some of these patients would be unable to tolerate an invasive procedure or surgery. Moreover, physicians could be aware that an older and sicker population is more likely to die from causes other than their ECF, even if it represents a serious diagnosis like an extracolonic cancer. In other words, older and sicker populations, which have higher frequencies of ECFs, are more likely to be overtreated. This claim is strengthened by the finding that slow-growing cancers are more likely to be found in older individuals.[99]

Comparison to Previous Publications on this Topic

My findings are similar to the existing literature on extracolonic findings from CTC. The last systematic review on this topic,[4] which looked at both screening an diagnostic CTC, reported that 58% of subjects (range 12%-90%) had at least one incidental lesion and 13.8% of subjects received some clinical workup. This review noted that only 0.8% of patients required immediate treatment and 3.7% were diagnosed with an extracolonic cancer or AAA. Nearly half (42%) of the reported cancers in their review were early-stage (N0M0).

The total number of patients with extracolonic findings was likely greater than my systematic review because the authors combined symptomatic and asymptomatic populations. Their review only included three studies with "screening" populations and one of these three[104] included patients with a personal history of colorectal cancer. Furthermore, the review was published the same year as the C-RADS classification system, making it impossible to assess how the system might affect clinical practice. Thus, my review adds to their findings by suggesting a slightly lower frequency of ECFs in true screening populations and providing evidence that a standardized classification system might not be particularly effective at reducing variability in the frequency and clinical importance of ECFs.

Implications of Results for Clinical Practice

There is little evidence in this review supporting the claims that detecting extracolonic findings during screening CTC improves patient outcomes. Some organizations have come to a similar conclusion and have not endorsed CTC for colorectal cancer.[73] In regards to ECFs specifically, screening with abdominal imaging is not currently supported by the evidence and is not recommended by the American College of Radiology.[105] Furthermore, if such screening did have clear net benefits, it would be most effective with IV-contrast and higher radiation doses. Therefore, the claimed benefits of detecting extracolonic findings from screening CTC, which uses low radiation doses and no IV contrast, should be interpreted with caution.

Currently, with the medicolegal concerns and reporting requirements for reimbursement,[106] the debate is not whether to report extracolonic findings but how to report extracolonic findings. One option would be selective reporting of incidental lesions, focusing on lesions with the greatest potential benefit such as high-risk AAAs and select extracolonic cancers. Still, the current medical culture in the United States might prohibit such withholding of information. At very least, radiologist might consider a move towards not reporting findings that do not require any workup (i.e. lesions of C-RADS E2 or below), as suggested by the Working Group on Virtual Colonoscopy.[3] In addition, primary care physicians should try to limit clinical workup to findings that might benefit from early intervention and report only potentially important findings to patients.

In addition, physicians should inform patients who opt for screening CTC of the likelihood of extracolonic findings, the potential for additional workup and the possible benefits and harms of such a workup. Given the complexity of concepts such as overdiagnosis and the

prevailing belief in the benefits of early detection, the discussion of harms might warrant more time and attention. Patients must be able to comprehend the benefits and harms of finding ECFs before they can make truly informed decisions.

One change that might reduce the variability of ECFs and decrease the number of unnecessary workups is a requirement for ECF-specific training for radiologists. Currently, the recommended training focuses on correctly interpreting a specific number of pathology-confirmed colorectal cancers,[87] with no specific training requirements for extracolonic lesions. Several studies support the benefits of existing CTC training requirements on the sensitivity and specificity of detecting colorectal cancers.[107] It is possible that similar training for ECFs would reduce unnecessary surveillance and treatment.

Moreover, classification systems and recommendation statements do not appear to provide ample guidance on the type and method of working up specific extracolonic findings. These studies showed huge variability in how extracolonic findings were worked up. Currently, the C-RADS classification system provides no guidance on how to address these findings other than stating which ones require some workup. The ACR's Committee on Incidental Findings provides a bit more information on how to address specific extracolonic findings by suggesting radiographic follow-up for specific findings. Unfortunately, their recommendations still fall short in many ways. For instance, in their section on low-dose, non-contrast CT they do not address lesions of the lungs, ovaries or bone.

Lastly, classification systems do not categorize lesions based on the likelihood of net benefit from workup and treatment. For instance, findings with some evidence supporting benefits for their screening, such as AAAs and renal cell carcinomas, are grouped together with incidental lesions that are less likely to benefit from workup and treatment. In addition, there is

reason to question the ability for classification labels to provide accurate prognostic information. The guidelines from the ACR and Virtual Colonoscopy Working Group might therefore benefit from proper validation. Lastly, the reported variability among included studies that employed C-RADS raises questions about its reliability.

Limitations of the Literature

The literature on extracolonic findings from screening CTC suffers from several weaknesses.

First, many studies had poor follow-up of ECFs and incomplete reporting of outcomes. Several studies decided not to report on the clinical workup for E1 or E2 findings since C-RADS dictates that they should not require follow-up. However, such an approach risks introducing bias into estimations. First, lesions determined to be of low clinical importance by the study authors might be followed in other settings, such as community hospitals. Therefore, the reported workup required for ECFs might be an underestimate. In fact, Pickhardt et al.,[82] who followed all ECFs, reported that 18 subjects received unnecessary workup for findings deemed to be less than moderate importance. These findings show that unnecessary workup might occur in the more controlled settings of a clinical study. It is possible that these rates would be even higher in community settings, where less experienced radiologists might be unfamiliar with the guidelines for addressing ECFs. In addition, excluding the workup for E1 and E2 findings assumes that the radiologists correctly classified these findings during the initial CTC. It is possible that findings initially overlooked during the first CTC could require workup in the future.

The poor follow-up of ECFs made it difficult to compare outcomes among studies, thus increasing the uncertainty of my conclusions. In order to reduce this variability in the future,

studies looking at screening CTC should use a set methodology for following ECFs. For instance, studies should report all workup including radiological exams, clinical visits, surgical procedures, invasive tests and medical treatments. Studies should also investigate the complications and psychological effects of this workup. The psychological effects could be beneficial (e.g. relief from a workup that reveals a benign lesion) or harmful (e.g. anxiety and uncertainty from a workup that might reveal a life-threatening diagnosis). Lastly, studies should report who initiated specific aspects of the workup for incidental lesions. For instance, it would be helpful to know what aspects of the workup are managed by radiologists or primary care physicians so that practice guidelines can be directed to relevant groups.

These studies collected little information on the potential harms of being diagnosed with incidental findings. For instance, there is the considerable risk of false positives. False positives come with potential psychological harms (e.g. anxiety) and physical harms of the resulting workup (e.g. ionizing radiation from CT or complications from invasive diagnostic procedures). These same psychological and physical harms are also experienced by those with true positives, although these patients have the possibility of clinical benefit. However, some of true positives likely represent overdiagnosis meaning that the resulting psychological and physical harms were experienced with no corresponding benefits. While the psychological harms could be more difficult to measure, they might be especially relevant for incidental findings that could result in years of diagnostic surveillance. There is sparse evidence on these harms, but some evidence suggests that they can be significant. One systematic review on the psychological impact of predicting individuals' risk of illness reported that receiving a positive diagnosis is associated with a greater risk of depression, anxiety, poorer perceptions of health and psychological distress.[108]

Another weakness of the current body of literature is that no studies address how ECFs are addressed in community settings. All six of my included studies were performed in academic or military medical centers with highly experienced radiologists. In order to truly gauge the effects of widespread CTC, it will be important to determine the frequency and clinical implications of extracolonic findings in non-academic settings. It is possible that these areas might employ less experienced radiologists who are less informed of the guidelines on ECFs, leading to an increased frequency of incidental lesions and their subsequent workup.

Limitations of Review

This systematic review suffered from a few limitations. For instance, the decision to limit the review to screening populations reduced the number of included studies and might have weakened the strength of evidence. It was thought that focusing on a homogenous population might increase the precision of previous estimates on the frequency of ECFs, but my results were also widely variability. Nonetheless, I thought that it was important to focus on screening populations in order to provide better estimates for this group. This is especially important since the implications of ECFs are different for a screening population compared with symptomatic patients. In addition, my wide range of ECF frequencies among screening CTC studies strengthens the evidence that other factors (i.e. radiologist inter-rater variability) are primarily responsible for the existing variability. These sources of practice variability should be further explored before CT colonography becomes a first-line screening test for colorectal cancer.

In order to increase the number of studies on asymptomatic populations in this review, I included studies that enrolled individuals at high risk of colorectal cancer. As a result, my study population might be less representative of a true screening population than a mixed

screening/surveillance population. This is especially relevant because my results indicated that patients at high risk of colorectal cancer had a higher frequency of ECFs. In addition, my review included studies from countries outside the United States. The only included non-US study[84] was performed in Australia and did not differ greatly from other studies in terms of reported interventions or outcomes.

Another weakness of my review is the inability to adjust for uneven follow-up periods, making it difficult to compare studies. The way I reported final diagnoses and workup set up an uneven comparison of studies, which have different times and methods of follow-up. Some studies performed an extensive, retrospective search of subjects' medical records to determine diagnoses while others reported on selected outcomes. For instance, Flicker et al.[85] followed only radiological workup, which might explain why they did not report the frequency of specific important diagnoses (i.e. AAAs, extracolonic cancers, etc.). I attempted to minimize the effects of these inconsistencies by excluding such studies from my calculated frequencies of specific diagnoses. However, it is possible that some of the studies that did report these outcomes also had poor follow-up. As a result, they might have underestimated the number of potentially serious diagnoses. This might not be an issue for AAAs, which can be diagnosed by screening CTC. But for extracolonic cancers, the length of follow-up might have a significant influence on the number of diagnosed cancers. To account for these differences, I also reported the number of lesions suspicious for malignancy.

I had originally planned to develop an outcomes table following a hypothetical cohort of 1,000 individuals receiving screening CTC. This table would report the expected number of ECFs, including moderate/high and high-importance findings, imaging tests, surgeries and medical treatments. In addition, it would lay out those findings that resulted in potential benefits,

potential harms or uncertain benefits/harms. But given the imprecision of results and the gaps in many reported outcomes, I were unable to develop such a table.

Conclusions

In sum, about one-third of patients receiving screening CT colonography will have at least one extracolonic finding discovered. The likelihood of discovering an ECF is higher for patients at higher risk of colorectal cancer. Roughly 10% of subjects who receive screening CTC will receive a diagnosis of high clinical importance and nearly all of these will result in some clinical workup. Approximately 2% of all subjects will have an abdominal aortic aneurysm or extracolonic cancer, which might potentially benefit from early treatment. However, some of these diagnoses will not result in net benefits and the majority of subjects receiving clinical workup will be diagnosed with a benign lesion.

The certainty in these findings is reduced by the large variability between studies, which does not seem to be reduced by a universal classification system for clinical importance. Since 2005, when the C-RADS system for classifying extracolonic findings was published, there has not been a reduction in the variability of reported ECFs or lesions of high-importance. Inter-rater variability between radiologists and inconsistent strategies for working up ECFs are likely the greatest contributors to the differences between studies. Increased radiologist training on correctly triaging ECFs might reduce some of the variability and unnecessary workup. In addition, selective reporting of outcomes that are more likely to benefit from workup (i.e. high-risk AAAs and renal cell carcinomas) might improve the ratio of benefit to harm for patients receiving screening CTC.

Funding Sources and Conflicts of Interest

There are no funding sources to report and no relevant conflicts of interest to disclose.

Appendices: Methods

Appendix A: Search Strategy

Search Database	Search Terms
PubMed	"(colonography[tw] OR virtual colonoscopy[tw] OR colography[tw] OR CT colonoscopy[tw] OR virtual endoscopy[tw]) AND (extracolonic[tw] OR incidental*[tw] OR incidentaloma*[tw] OR serendipitous[tw])"
EMBASE	"(Colonography:ti,ab,de OR "virtual colonoscopy":ti,ab,de OR colography:ti,ab,de OR CT colonoscopy:ti,ab,de OR "virtual endoscopy":ti,ab,de) AND (Extracolonic:ti,ab,de OR incidental*:ti,ab,de OR incidentaloma*:ti,ab,de OR serendipitous:ti,ab,de)"
Cochrane Libraries	"(Colonography OR "virtual colonoscopy" OR colography OR CT colonoscopy OR "virtual endoscopy") AND (Extracolonic OR incidental* OR incidentaloma* OR serendipitous)"

Appendix B: Reporting Criteria

Category	Criteria
Description of Population	Were the eligibility criteria well defined? Was recruitment well described, including: - Were patients enrolled consecutively? - How did investigators recruit patients (e.g. referral or self-selection)?

	- From where were patients recruited? Was the population well-defined including information on subjects': - Gender? - Ethnicity? - Previous history of polyps? - Risk factors for colon cancer (e.g. family history, polyposis syndrome)? - Present symptoms?
Description of Interventions	Were the interventions adequately described including: - Radiation dose - Slice thickness - Positioning of patient (i.e. supine and/or prone exam) - Use of IV contrast
Description of Outcomes	Were primary and secondary outcomes defined a-priori and clearly reported?

Appendix C: Internal Validity Criteria

Category	Criteria
Measurement Bias	Equal: - Was there a standardized CTC technique used for all patients (i.e. radiation dose, slice thickness, use of IV contrast)? - Was the clinical importance of findings judged using a standardized classification system? Valid: - Were extracolonic findings clearly defined? - Were patients' previously diagnosed extracolonic lesions excluded from the cohort? - Was the clinical importance of findings based on a classification system that was related to likelihood of gaining benefit from follow-up and/or treatment? Reliable:

	- Did radiologists have adequate training to interpret CT colonography, according to the training guidelines published by the American College of Radiology (ACR)?[87] - Did another trained radiologist review scans? - If there were more than one observer (radiologist), did they independently assess scans? - Were radiologists blinded to past scan results during follow-up examinations?
Selection Bias	Are there any clear sources of selection bias? - Is there something that distinguishes those with extracolonic findings from those that don't? - Are those who are lost to follow-up different from those who are left in the cohort?
Confounding	Is there something that distinguishes those with extracolonic findings from those that don't?
Lost to Follow-up/Dropouts	Was more than 30% of the cohort lost to follow-up?
Power	Was the sample size < 50 subjects?

Appendix D: External Validity Ratings – Adapted from the USPSTF[78]

Rating	Description
Good	The study differs minimally from the standard CT colonography screening population/situation/providers. It is highly probable (>90%) that the experience with CTC described in this study would be attained in a typical screening population.
Fair	The study differs in a few ways from the standard CT colonography screening population/situation/providers that has the potential to affect the clinical outcomes. It is moderately probable (50%-89%%) that the experience with CTC described in this study would be attained in a typical screening population.

Poor	The study differs in many ways from the standard CT colonography screening population/situation/providers that has a high likelihood of affecting the clinical outcomes. The probability is low (<50%) that the experience with CTC described in this study would be attained in a typical screening population.

Appendix E: Critical Appraisal Questions – Adapted from the USPSTF[78]

1. Do the studies have the appropriate research design to answer the key question(s)?
2. To what extent are the existing studies of high quality? (i.e., what is the internal validity?)
3. To what extent are the results of the studies generalizable to the general US primary care population and situation? (i.e., what is the external validity?)
4. How many studies have been conducted that address the key question(s)? How large are the studies? (i.e., what is the precision of the evidence?)
5. How consistent are the results of the studies?
6. Are there additional factors that assist us in drawing conclusions (e.g., presence or absence of dose-response effects; fit within a biologic model)?

Appendix F: Strength of Evidence Grades

Grade	Description
High	Future research is unlikely to change the confidence of the evidence
Moderate	Future research will likely have an important effect on the confidence of this evidence and might change these estimates
Low	Future research will very likely have an important effect on the confidence of this evidence and is

	likely change these estimates
Very low	The estimate provided is very uncertain

Appendices: Results

Appendix G. Internal Validity Ratings

Aspects of Measurement Bias in Included Studies

Author (Year)	Equal Measurements		Valid Measurements			Reliable Measurements			
	Standardized CT technique for all?	Standardized system for clinical importance?	Clinical importance based on dx/tx benefit?	Previously diagnosed ECFs excluded or reported separately?	Method of follow-up complete and accurate?	Radiologists adequately trained for CTC?	Duplicate reading of CTC scans?	Review of CTC scans performed independently?	Radiologists blinded to patient history?
Chin et al. (2005)	Yes	No	No	Yes	Yes	Yes	Yes	No	?
Pickhardt et al. (2008)	Yes	No	No	Yes	Yes	Yes	No	N/A	No
Flicker et al. (2008)	No	Yes	No	Yes[a]	No	?[b]	Yes	Yes	Yes
Gluecker et al. (2003)	No	No	No	No	No	?[b]	No	N/A	Yes
Macari et al. (2011)	No	Yes	No	?	No	Yes	No	N/A	?
Veerappan et al. (2010)	Yes	Yes	Yes	Yes[a]	No	Yes	Yes	?[c]	?

[a] Previously diagnosed ECFs were reported separately

[b] Radiologists were board-certified and experienced but CTC-related experience not reported
[c] Stated that senior radiologist reviewed all scans with significant findings, so likely that it was done with the other radiologist's report

Aspects of Selection Bias in Included Studies

Author (Year)	Confounding	Selection Bias		
	Something distinguishing those with ECFs from those without?	Consecutive enrollment?	Do dropouts differ from those properly followed up?	More than 30% of cohort lost to follow-up?
Chin et al. (2005)	Yes	No[a]	N/A[b]	No
Pickhardt et al. (2008)	Yes	Yes	?	No
Flicker et al. (2008)	?	Yes	N/A[b]	No
Gluecker et al. (2003)	?	Yes	?	No
Macari et al. (2011)	?	Yes	?	No
Veerappan et al. (2010)	No	Yes	?	No

[a] Randomly selected cohort representing a small percentage of eligible cohort
[b] No subjects were lost to follow-up

Overall Assessment of Internal Validity

Author (Year)	Risk of Measurement Bias	Risk of Confounding	Risk of Selection Bias	Overall Internal Validity
Chin et al. (2005)	Low/Medium	Medium	Low/Medium	Fair
Pickhardt et al. (2008)	Medium	Medium	Low/Medium	Fair
Flicker et al. (2008)	Medium	Medium	Low/Medium	Fair
Gluecker et al. (2003)	High	Medium	Medium	Fair
Macari et al. (2011)	High	Medium	High	Fair/Poor
Veerappan et al. (2010)	Medium/High	Low	Low	Fair

Appendix H. External Validity Ratings

External Validity of Population, Setting and Interventions and Overall Generalizability

Author (Year)	Population[a]	Setting[b]	Intervention[c]	Overall External Validity
Chin et al. (2005)	Good	Fair	Fair	Fair
Pickhardt et al. (2008)	Good	Fair	Fair	Fair
Flicker et al. (2008)	Poor	Fair	Fair	Fair
Gluecker et al. (2003)	Poor	Fair	Fair	Fair
Macari et al. (2011)	Fair	Fair	Poor/Fair	Fair
Veerappan et al. (2010)	Good	Fair	Poor/Fair	Fair

[a] How closely the study population matches an asymptomatic, average-risk population in the recommended screening age of 50-74
[b] How well did study settings replicate the mixture of academic and community settings
[c] How close were study interventions (i.e. CTC modality, radiologist experience, technique of reading CTC) to the typical CTC screening programs

Acknowledgements

I would like to thank Kate McGraw of the UNC Health Sciences Library for her assistance with refining the search for studies.

References:

1. Pickhardt PJ, Choi JR, Hwang I, et al. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. *N Engl J Med.* Dec 4 2003;349(23):2191-2200.
2. Hara AK. Extracolonic findings at CT colonography. *Semin Ultrasound CT MR.* Feb 2005;26(1):24-27.
3. Zalis M, Barish M, Choi JR. CT colonography reporting and data system: a consensus proposal. *Radiology.* 2005;236(1):3-9.
4. Xiong TR, M.; Woodroffe, R.; Halligan, S.; Morton, D.; Lilford, R. J. Incidental lesions found on CT colonography: their nature and frequency. *Br J Radiol.* Jan 2005;78(925):22-29.
5. Siddiki HF, J. G.; McFarland, B.; Dajani, N.; Orme, N.; Koenig, B.; Strobel, M.; Wolf, S. M. Incidental findings in CT colonography: literature review and survey of current research practice. *J Law Med Ethics.* Summer 2008;36(2):320-331, 213.
6. Iannaccone R, Laghi A, Catalano C, et al. Detection of colorectal lesions: lower-dose multi-detector row helical CT colonography compared with conventional colonoscopy. *Radiology.* Dec 2003;229(3):775-781.
7. Berland LL, Silverman SG, Gore RM, et al. Managing incidental findings on abdominal CT: white paper of the ACR incidental findings committee. *J Am Coll Radiol.* Oct 2010;7(10):754-773.
8. Johnson CD. CT colonography: coming of age. *AJR Am J Roentgenol.* Nov 2009;193(5):1239-1242.
9. Ransohoff DF, Sandler RS. Clinical practice. Screening for colorectal cancer. *N Engl J Med.* Jan 3 2002;346(1):40-44.
10. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. *CA Cancer J Clin.* Jul-Aug 2009;59(4):225-249.
11. Espey DK, Wu XC, Swan J, et al. Annual report to the nation on the status of cancer, 1975-2004, featuring cancer in American Indians and Alaska Natives. *Cancer.* Nov 15 2007;110(10):2119-2152.
12. United States Preventive Services Task Force. Screening for colorectal cancer: recommendation and rationale. *Am Fam Physician.* Dec 15 2002;66(12):2287-2290.
13. Winawer SJ, Zauber AG, Fletcher RH, et al. Guidelines for colonoscopy surveillance after polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer and the American Cancer Society. *CA Cancer J Clin.* May-Jun 2006;56(3):143-159; quiz 184-145.
14. Smith RA, Cokkinides V, Eyre HJ. Cancer screening in the United States, 2007: a review of current guidelines, practices, and prospects. *CA Cancer J Clin.* Mar-Apr 2007;57(2):90-104.
15. Centers for Medicare and Medicaid Services (CMS). Decision Memo for Screening Computed Tomography Colonography (CTC) for Colorectal Cancer (CAG-00396N). 2009; http://www.cms.gov/medicare-coverage-database/details/nca-decision-memo.aspx?NCAId=220&NcaName=Screening+Computed+Tomography+Colonography+%28CTC%29+for+Colorectal+Cancer&TAId=58&IsPopup=y&bc=AAAAAAAAIAAA&. Accessed May 25, 2012.
16. Klabunde C, Breen N, Meissner H, Subramanian S. Use of colonoscopy for colorectal cancer screening. *Cancer Epidemiol Biomarkers Prev.* Sep 2005;14(9):2279-2280; author reply 2280-2271.
17. Meissner HI, Breen N, Klabunde CN, Vernon SW. Patterns of colorectal cancer screening uptake among men and women in the United States. *Cancer Epidemiol Biomarkers Prev.* Feb 2006;15(2):389-394.
18. Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N Engl J Med.* Feb 23 2012;366(8):687-696.

19. Whitlock EP, Lin JS, Liles E, Beil TL, Fu R. Screening for colorectal cancer: a targeted, updated systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med.* Nov 4 2008;149(9):638-658.
20. Inadomi JM, Vijan S, Janz NK, et al. Adherence to colorectal cancer screening: a randomized clinical trial of competing strategies. *Arch Intern Med.* Apr 9 2012;172(7):575-582.
21. Rex DK. Virtual colonoscopy: time for some tough questions for radiologists and gastroenterologists. *Endoscopy.* Mar 2000;32(3):260-263.
22. Lieberman DA, Weiss DG, Bond JH, Ahnen DJ, Garewal H, Chejfec G. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. Veterans Affairs Cooperative Study Group 380. *N Engl J Med.* Jul 20 2000;343(3):162-168.
23. Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, Ransohoff DF. Risk of advanced proximal neoplasms in asymptomatic adults according to the distal colorectal findings. *N Engl J Med.* Jul 20 2000;343(3):169-174.
24. Rockey DC, Koch J, Cello JP, Sanders LL, McQuaid K. Relative frequency of upper gastrointestinal and colonic lesions in patients with positive fecal occult-blood tests. *N Engl J Med.* Jul 16 1998;339(3):153-159.
25. Winawer SJ, Stewart ET, Zauber AG, et al. A comparison of colonoscopy and double-contrast barium enema for surveillance after polypectomy. National Polyp Study Work Group. *N Engl J Med.* Jun 15 2000;342(24):1766-1772.
26. Gluecker TM, Johnson CD, Harmsen WS, et al. Colorectal cancer screening with CT colonography, colonoscopy, and double-contrast barium enema examination: prospective assessment of patient perceptions and preferences. *Radiology.* May 2003;227(2):378-384.
27. Vining D, Gelfand D. Noninvasive colonoscopy using helical CT scanning, 3D reconstruction, and virtual reality. *23rd Annual Meeting and Postgraduate Course of the Society of Gastrointestinal Radiologists*. Maui, Hawaii1994.
28. Macari M, Bini EJ. CT colonography: where have we been and where are we going? *Radiology.* Dec 2005;237(3):819-833.
29. Mang T, Graser A, Schima W, Maier A. CT colonography: techniques, indications, findings. *Eur J Radiol.* Mar 2007;61(3):388-399.
30. Dachman AH. Advice for optimizing colonic distention and minimizing risk of perforation during CT colonography. *Radiology.* May 2006;239(2):317-321.
31. Barish MA, Soto JA, Ferrucci JT. Consensus on current clinical practice of virtual colonoscopy. *AJR Am J Roentgenol.* Mar 2005;184(3):786-792.
32. McFarland EGF, J. G.; Pickhardt, P.; Dachman, A.; Yee, J.; McCollough, C. H.; Macari, M.; Knechtges, P.; Zalis, M.; Barish, M.; Kim, D. H.; Keysor, K. J.; Johnson, C. D. ACR Colon Cancer Committee white paper: status of CT colonography 2009. *J Am Coll Radiol.* Nov 2009;6(11):756-772 e754.
33. Gluecker TM, Johnson CD, Wilson LA, et al. Extracolonic findings at CT colonography: evaluation of prevalence and cost in a screening population. *Gastroenterology.* Apr 2003;124(4):911-916.
34. Macari M, Bini EJ, Xue X, et al. Colorectal neoplasms: prospective comparison of thin-section low-dose multi-detector row CT colonography and conventional colonoscopy for detection. *Radiology.* Aug 2002;224(2):383-392.
35. Iannaccone R, Laghi A, Catalano C, Mangiapane F, Piacentini F, Passariello R. Feasibility of ultra-low-dose multislice CT colonography for the detection of colorectal lesions: preliminary experience. *Eur Radiol.* Jun 2003;13(6):1297-1302.
36. Johnson CD, Chen MH, Toledano AY, et al. Accuracy of CT colonography for detection of large adenomas and cancers. *N Engl J Med.* Sep 18 2008;359(12):1207-1217.

37. Qaseem A, Denberg TD, Hopkins RH, Jr., et al. Screening for Colorectal Cancer: A Guidance Statement From the American College of Physicians. *Ann Intern Med.* Mar 6 2012;156(5):378-386.

38. Levin B, Lieberman DA, McFarland B, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology.* May 2008;134(5):1570-1595.

39. Klabunde CN, Lanier D, Nadel MR, McLeod C, Yuan G, Vernon SW. Colorectal cancer screening by primary care physicians: recommendations and practices, 2006-2007. *Am J Prev Med.* Jul 2009;37(1):8-16.

40. Macari M, Berman P, Dicker M, Milano A, Megibow AJ. Usefulness of CT colonography in patients with incomplete colonoscopy. *AJR Am J Roentgenol.* Sep 1999;173(3):561-564.

41. Morrin MM, Kruskal JB, Farrell RJ, Goldberg SN, McGee JB, Raptopoulos V. Endoluminal CT colonography after an incomplete endoscopic colonoscopy. *AJR Am J Roentgenol.* Apr 1999;172(4):913-918.

42. Fenlon HM, McAneny DB, Nunes DP, Clarke PD, Ferrucci JT. Occlusive colon carcinoma: virtual colonoscopy in the preoperative evaluation of the proximal colon. *Radiology.* Feb 1999;210(2):423-428.

43. Gollub MJ, Ginsberg MS, Cooper C, Thaler HT. Quality of virtual colonoscopy in patients who have undergone radiation therapy or surgery: how successful are we? *AJR Am J Roentgenol.* May 2002;178(5):1109-1116.

44. Burling D, Halligan S, Slater A, Noakes MJ, Taylor SA. Potentially serious adverse events at CT colonography in symptomatic patients: national survey of the United Kingdom. *Radiology.* May 2006;239(2):464-471.

45. Pickhardt PJ. Incidence of colonic perforation at CT colonography: review of existing data and implications for screening of asymptomatic adults. *Radiology.* May 2006;239(2):313-316.

46. Sosna J, Blachar A, Amitai M, et al. Colonic perforation at CT colonography: assessment of risk in a multicenter large cohort. *Radiology.* May 2006;239(2):457-463.

47. Rockey DC. Computed tomographic colonography: current perspectives and future directions. *Gastroenterology.* Jul 2009;137(1):7-14.

48. Summers RM, Johnson CD, Pusanik LM, Malley JD, Youssef AM, Reed JE. Automated polyp detection at CT colonography: feasibility assessment in a human population. *Radiology.* Apr 2001;219(1):51-59.

49. Yoshida H, Nappi J, MacEneaney P, Rubin DT, Dachman AH. Computer-aided diagnosis scheme for detection of polyps at CT colonography. *Radiographics.* Jul-Aug 2002;22(4):963-979.

50. Svensson MH, Svensson E, Lasson A, Hellstrom M. Patient acceptance of CT colonography and conventional colonoscopy: prospective comparative study in patients with or suspected of having colorectal disease. *Radiology.* Feb 2002;222(2):337-345.

51. Thomeer M, Bielen D, Vanbeckevoort D, et al. Patient acceptance for CT colonography: what is the real issue? *Eur Radiol.* Jun 2002;12(6):1410-1415.

52. Ristvedt SL, McFarland EG, Weinstock LB, Thyssen EP. Patient preferences for CT colonography, conventional colonoscopy, and bowel preparation. *Am J Gastroenterol.* Mar 2003;98(3):578-585.

53. Akerkar GA, Yee J, Hung R, McQuaid K. Patient experience and preferences toward colon cancer screening: a comparison of virtual colonoscopy and conventional colonoscopy. *Gastrointest Endosc.* Sep 2001;54(3):310-315.

54. Forbes GM, Mendelson RM. Patient acceptance of virtual colonoscopy. *Endoscopy.* Mar 2000;32(3):274-275.

55. Florie J, van Gelder RE, Stoker J. Colonography by computed tomography. *Eur J Gastroenterol Hepatol.* Aug 2005;17(8):809-813.
56. Lefere P, Gryspeerdt S, Baekelandt M, Van Holsbeeck B. Laxative-free CT colonography. *AJR Am J Roentgenol.* Oct 2004;183(4):945-948.
57. Rex DK, Vining D, Kopecky KK. An initial experience with screening for colon polyps using spiral CT with and without CT colography (virtual colonoscopy). *Gastrointest Endosc.* Sep 1999;50(3):309-313.
58. Brenner DJ, Hall EJ. Computed tomography--an increasing source of radiation exposure. *N Engl J Med.* Nov 29 2007;357(22):2277-2284.
59. McGrath JS, Ponich TP, Gregor JC. Screening for colorectal cancer: the cost to find an advanced adenoma. *Am J Gastroenterol.* Nov 2002;97(11):2902-2907.
60. Veerappan GR, Ally MR, Choi JH, Pak JS, Maydonovitch C, Wong RK. Extracolonic findings on CT colonography increases yield of colorectal cancer screening. *AJR Am J Roentgenol.* Sep 2010;195(3):677-686.
61. Iannaccone R, Laghi A, Catalano C, Mangiapane F, Marin D, Passariello R. Computed tomography colonography: colon examination or Pandora's box. *Gut.* Jun 2004;53(6):915.
62. Spreng AN, P.; Mattich, J.; Dinkel, H. P.; Vock, P.; Hoppe, H. Importance of extracolonic findings at IV contrast medium-enhanced CT colonography versus those at non-enhanced CT colonography. *Eur Radiol.* Oct 2005;15(10):2088-2095.
63. Sistrom C, Dreyer K, Dang P. Recommendations for additional imaging in radiology reports: multifactorial analysis of 5.9 million examinations. *Radiology.* 2009;253(2):453-461.
64. Burling D, Halligan S, Atchley J, et al. CT colonography: interpretative performance in a non-academic environment. *Clin Radiol.* May 2007;62(5):424-429; discussion 430-421.
65. Johnson CD, Toledano AY, Herman BA, et al. Computerized tomographic colonography: performance evaluation in a retrospective multicenter setting. *Gastroenterology.* Sep 2003;125(3):688-695.
66. Rockey DC, Paulson E, Niedzwiecki D, et al. Analysis of air contrast barium enema, computed tomographic colonography, and colonoscopy: prospective comparison. *Lancet.* Jan 22-28 2005;365(9456):305-311.
67. Cotton PB, Durkalski VL, Pineau BC, et al. Computed tomographic colonography (virtual colonoscopy): a multicenter comparison with standard colonoscopy for detection of colorectal neoplasia. *JAMA.* Apr 14 2004;291(14):1713-1719.
68. American College of Radiology (ACR). *Breast Imaging Reporting and Data System Atlas (BI-RADS Atlas).* Reston, VA: American College of Radiology;2003.
69. Quint LE, Watcharotone K, Myles JD, Gruppen LD, Mullan PB. Incidental findings at chest CT: a needs assessment survey of radiologists' knowledge. *Acad Radiol.* Dec 2011;18(12):1500-1506.
70. Chang MS, Shah JP, Amin S, et al. Physician knowledge and appropriate utilization of computed tomographic colonography in colorectal cancer screening. *Abdom Imaging.* Oct 2011;36(5):524-531.
71. American College of Radiology (ACR). President Obama Gets Virtual Colonoscopy (CT Colonography) But Medicare Denies CTC Coverage to Seniors. 2010; http://www.acr.org/MainMenuCategories/media_room/FeaturedCategories/PressReleases/President-Gets-CTC.aspx. Accessed March 31, 2012.
72. Westbrook JI, Braithwaite J, McIntosh JH. The outcomes for patients with incidental lesions: serendipitous or iatrogenic? *AJR Am J Roentgenol.* Nov 1998;171(5):1193-1196.
73. Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med.* Nov 4 2008;149(9):627-637.

74. Elmore JG, Nakano CY, Koepsell TD, Desnick LM, D'Orsi CJ, Ransohoff DF. International variation in screening mammography interpretations in community-based programs. *J Natl Cancer Inst.* Sep 17 2003;95(18):1384-1393.
75. Fleming C, Whitlock EP, Beil TL, Lederle FA. Screening for abdominal aortic aneurysm: a best-evidence systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med.* Feb 1 2005;142(3):203-211.
76. Ashton HA, Buxton MJ, Day NE, et al. The Multicentre Aneurysm Screening Study (MASS) into the effect of abdominal aortic aneurysm screening on mortality in men: a randomised controlled trial. *Lancet.* Nov 16 2002;360(9345):1531-1539.
77. Tsui KH, Shvarts O, Smith RB, Figlin R, de Kernion JB, Belldegrun A. Renal cell carcinoma: prognostic significance of incidentally detected tumors. *J Urol.* Feb 2000;163(2):426-430.
78. U.S. Preventive Services Task Force (USPSTF). U.S. Preventive Services Task Force Procedure Manual. 2011; http://www.uspreventiveservicestaskforce.org/uspstf08/methods/procmanual.pdf. Accessed April 21, 2012.
79. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *BMJ.* Jun 19 2004;328(7454):1490.
80. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* Oct 2009;62(10):e1-34.
81. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol.* Oct 2009;62(10):1006-1012.
82. Pickhardt PH, M. E.; Vanness, D. J.; Lo, J. Y.; Kim, D. H.; Taylor, A. J.; Winter, T. C.; Hinshaw, J. L. Unsuspected extracolonic findings at screening CT colonography: clinical and economic impact. *Radiology.* 2008;249(1):151-159.
83. Kim DHP, P. J.; Hanson, M. E.; Hinshaw, J. L. CT colonography: performance and program outcome measures in an older screening population. *Radiology.* Feb 2010;254(2):493-500.
84. Chin MM, R.; Edwards, J.; Foster, N.; Forbes, G. Computed tomographic colonography: prevalence, nature, and clinical significance of extracolonic findings in a community screening program. *Am J Gastroenterol.* Dec 2005;100(12):2771-2776.
85. Flicker MST, A. T.; Hazra, A.; Dachman, A. H. Economic impact of extracolonic findings at computed tomographic colonography. *J Comput Assist Tomogr.* Jul-Aug 2008;32(4):497-503.
86. Macari M, Nevsky G, Bonavita J, Kim DC, Megibow AJ, Babb JS. CT colonography in senior versus nonsenior patients: extracolonic findings, recommendations for additional imaging, and polyp prevalence. *Radiology.* Jun 2011;259(3):767-774.
87. American College of Radiology (ACR). *ACR Practice Guideline for the Performance of Computed Tomography (CT) Colonography in Adults (Resolution 36).* Reston, VA: American College of Radiology (ACR),;2009.
88. Liedenbaum MH, Venema HW, Stoker J. Radiation dose in CT colonography--trends in time and differences between daily practice and screening protocols. *Eur Radiol.* Oct 2008;18(10):2222-2230.
89. Berland LL. The American College of Radiology strategy for managing incidental findings on abdominal computed tomography. *Radiol Clin North Am.* Mar 2011;49(2):237-243.
90. Obuchowski NA, Holden D, Modic MT, et al. Total-body screening: preliminary results of a pilot randomized controlled trial. *J Am Coll Radiol.* Sep 2007;4(9):604-611.
91. United States Preventive Services Task Force. Screening for abdominal aortic aneurysm: recommendation statement. *Ann Intern Med.* Feb 1 2005;142(3):198-202.

92. Scott RA, Vardulaki KA, Walker NM, Day NE, Duffy SW, Ashton HA. The long-term benefits of a single scan for abdominal aortic aneurysm (AAA) at age 65. *Eur J Vasc Endovasc Surg.* Jun 2001;21(6):535-540.
93. Crow P, Shaw E, Earnshaw JJ, Poskitt KR, Whyman MR, Heather BP. A single normal ultrasonographic scan at age 65 years rules out significant aneurysm disease for life in men. *Br J Surg.* Jul 2001;88(7):941-944.
94. Emerton ME, Shaw E, Poskitt K, Heather BP. Screening for abdominal aortic aneurysm: a single scan is enough. *Br J Surg.* Aug 1994;81(8):1112-1113.
95. Lederle FA, Johnson GR, Wilson SE, et al. Yield of repeated screening for abdominal aortic aneurysm after a 4-year interval. Aneurysm Detection and Management Veterans Affairs Cooperative Study Investigators. *Arch Intern Med.* Apr 24 2000;160(8):1117-1121.
96. Wanhainen A, Rosen C, Rutegard J, Bergqvist D, Bjorck M. Low quality of life prior to screening for abdominal aortic aneurysm: a possible risk factor for negative mental effects. *Ann Vasc Surg.* May 2004;18(3):287-293.
97. Fenton JJ, Weiss NS. Screening computed tomography: will it result in overdiagnosis of renal carcinoma? *Cancer.* Mar 1 2004;100(5):986-990.
98. Black WC. Overdiagnosis: An underrecognized cause of confusion and harm in cancer screening. *J Natl Cancer Inst.* Aug 16 2000;92(16):1280-1282.
99. Zhang J, Kang SK, Wang L, Touijer A, Hricak H. Distribution of renal tumor growth rates determined by using serial volumetric CT measurements. *Radiology.* Jan 2009;250(1):137-144.
100. Buys SS, Partridge E, Black A, et al. Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. *JAMA.* Jun 8 2011;305(22):2295-2303.
101. Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med.* Aug 4 2011;365(5):395-409.
102. Hassan CP, P. J.; Laghi, A.; Kim, D. H.; Zullo, A.; Iafrate, F.; Di Giulio, L.; Morini, S. Computed tomographic colonography to screen for colorectal cancer, extracolonic cancer, and aortic aneurysm: model simulation with cost-effectiveness analysis. *Arch Intern Med.* Apr 14 2008;168(7):696-705.
103. Grosse SD. Assessing cost-effectiveness in healthcare: history of the $50,000 per QALY threshold. *Expert Rev Pharmacoecon Outcomes Res.* Apr 2008;8(2):165-178.
104. Hara AK, Johnson CD, MacCarty RL, Welch TJ. Incidental extracolonic findings at CT colonography. *Radiology.* May 2000;215(2):353-357.
105. American college of Radiology (ACR). ACR Statement on Whole Body CT Screening. 2002; http://www.acr.org/About-Us/Media-Center/Position-Statements/Position-Statements-Folder/ACR-Statement-on-Whole-Body-CT-Screening. Accessed May 25, 2012.
106. Silva E, 3rd. Extracardiac and extracolonic findings: extra important. *J Am Coll Radiol.* Mar 2010;7(3):173-176.
107. Haycock A, Burling D, Wylie P, Muckian J, Ilangovan R, Thomas-Gibson S. CT colonography training for radiographers--a formal evaluation. *Clin Radiol.* Dec 2010;65(12):997-1004.
108. Shaw C, Abrams K, Marteau TM. Psychological impact of predicting individuals' risks of illness: a systematic review. *Soc Sci Med.* Dec 1999;49(12):1571-1598.