CT colonography in the detection of colorectal polyps and cancer: systematic review, meta-analysis, and proposed minimum data set for study level reporting


CRD summary
This review addressed the diagnostic performance of computed tomographic (CT) colonography for the detection of colorectal polyps and cancer. The authors concluded that CT colonography was adequate for detecting polyps larger than 5 mm and particularly sensitive for detecting symptomatic cancer. These conclusions are likely to be generally reliable, however, estimates for cancer detection were based on a small data set and suboptimal method of analysis.

Authors’ objectives
To assess the methodological quality of published data on computed tomographic (CT) colonography. The authors also reported on an assessment of the diagnostic performance of CT colonography.

Searching
Two reviewers independently searched MEDLINE from January 1994 (first description of CT colonography) to December 2003, in addition to EMBASE, the Cochrane CENTRAL Register and the Science Citation Index (January 2004 to May 2003); the search terms were reported. Key journals in radiology, gastroenterology, surgery and general medicine were handsearched and the reference lists of the included studies were checked. No language restrictions were applied.

Study selection
Study designs of evaluations included in the review
Study designs were not specified by the inclusion criteria. Studies that were not peer reviewed (e.g. abstracts) and those with less than 30 participants were excluded. The included studies were required to interpret CT findings before those of the reference standard, or to blind CT observers to the reference standard findings.

Specific interventions included in the review
Studies that used key methods for CT colonography (see Other Publications of Related Interest) were eligible for inclusion. In particular, full bowel preparation had to be administered, prone and supine images acquired, and helical scanners used. Studies that used intravenous iodinated contrast material in all participants were excluded, as this was unlikely to be routinely used in the context of a screening programme. In addition, the included studies had to use commercially available software for CT interpretation.

Reference standard test against which the new test was compared
Studies that verified CT colonography findings with a reference standard (conventional endoscopy or surgical findings) were eligible for inclusion. No details of the reference standards used in the individual included studies were reported.

Participants included in the review
Studies of the detection of colorectal polyps in humans were eligible for inclusion. Studies were excluded if the prevalence of abnormality could be estimated a priori by the CT observers (e.g. where an inclusion criterion was a prior positive test). Studies in which CT examination was conducted because of an incomplete colonoscopy due to an obstructing tumour were also excluded (unless these patients comprised less than 50% of the study population or formed a distinct, excludable subset). The prevalence of abnormality in participants in the included studies ranged from 15 to 72%.

Outcomes assessed in the review
The authors did not state any inclusion criteria for the outcome measures (measures of accuracy for detecting polyps and cancers), but stated that they attempted to extract 2x2 contingency tables of test characteristics for both per-patient and per-polyp data. Sensitivity and specificity with 95% confidence intervals (CIs) were calculated on a per-patient...
basis. The results were also assessed on a per-polyp basis, but only sensitivity was calculated as no denominator was available for calculating specificity.

How were decisions on the relevance of primary studies made?
Two reviewers independently assessed all retrieved abstracts for relevance and then assessed articles for inclusion. Any disagreements were resolved by consensus or at monthly project meetings.

Assessment of study quality
The quality assessment was based upon the STARD (Standards for the Reporting of Diagnostic Accuracy Studies) and QUADAS (Quality Assessment of Diagnostic Accuracy Studies) criteria. The items assessed related to: participant characteristics; the time interval between CT colonography and the reference standard (potential for disease progression between tests); adequacy of the description of CT colonography and reference standard and how these were applied; any influence of the CT result upon interpretation of the reference standard; the reporting of technical failures for both tests; details of the numbers of observers and how data from multiple observers were handled; and any learning effect for CT colonography over the course of a study. Two reviewers independently extracted data on methodological quality.

Data extraction
Two reviewers independently extracted data on sample size and diagnostic performance in the form of 2x2 data. Where possible, the extracted data were stratified by polyp size: less than 6 mm (small), 6 to 9 mm (medium), and 1 cm or larger (large). An attempt was also made to distinguish established cancers from polyps. Readings from multiple observers were averaged and rounded down to the nearest whole number, except in the case of comparisons of experienced and inexperienced observers where data from inexperienced observers were excluded. If data were incomplete or inconsistent, authors were contacted for clarification. Authors of multiple papers were also contacted to determine any overlap.

Methods of synthesis
How were the studies combined?
The per-patient analyses were based upon the largest polyp in each patient. The data were analysed in three categories: large polyp (1 cm or larger), large and medium polyps combined (larger than 5 mm), and all polyps combined (i.e. small, medium and large polyps; no minimum size). A hierarchical binary random-effects model was used to pool the sensitivities and specificities, treating them as paired outcomes. This was used to estimate summary receiver operating characteristic (ROC) curves, allowing for variation in threshold between studies. The average operating point on the summary ROC curve which represents sensitivity and specificity at the average threshold was calculated.

For the per-polyp analyses a random-effects model was used to pool sensitivities.

The number of cancers per study was too small to enable a meta-analysis. The sensitivity for cancer detection was therefore estimated by treating the data as if they were from a single study.

How were differences between studies investigated?
Between-study heterogeneity was visualised using ROC space plots (for per-patient analyses only) and forest plots. A chi-squared test was used to estimate between-study heterogeneity for sensitivity and specificity, although the results were not reported.

Results of the review
Twenty-four studies with 4,181 participants were included in the review.

Methodological quality.
Twenty-three of the 24 included studies were of symptomatic patients or patients with prior history. Studies used between one and four observers for CT colonography; 14 reported data for individual observers and 5 investigated learning effects. CT colonography and the reference test were performed on the same day in all but six patients from
studies. The results of CT were used to modify colonoscopy in 6 studies. The description of CT technique and application was adequate to permit replication in all studies, but colonoscopy was inadequately described in 11 studies. CT technical failures were reported in 17 studies and 4 stated that there were no failures. Eleven studies presented incomplete colonoscopy data and 6 stated that colonoscopy was complete in all patients.

Per-patient analyses.

For large polyps (2,610 patients in 7 studies), the average sensitivity was 93% (95% CI: 73, 98) and the average specificity was 97% (95% CI: 95, 99).

For medium and large polyps (1,834 patients in 7 studies), the average sensitivity was 86% (95% CI: 75, 93) and the average specificity was 86% (95% CI: 76, 93).

For all polyps (1,361 patients in 12 studies), a meta-analysis was deemed inappropriate on the basis of between-study heterogeneity in both sensitivity and specificity. The sensitivity ranged from 45 to 97% and the specificity from 26 to 97%.

Per-polyp analyses.

The average sensitivity was 77% (95% CI: 70, 83) for large polyps and 70% (95% CI: 63, 76) for medium and large polyps. No meta-analysis was undertaken for all polyps because of heterogeneity.

Cancer detection.

Of a total of 150 cancers, 144 were detected. When treating data from the included studies as though they were derived from a single study, the estimated sensitivity was 96% (95% CI: 91, 99).

Authors’ conclusions

The sensitivity and specificity of CT colonography appeared adequate for the detection of large and medium colorectal polyps; sensitivity was particularly good for the detection of cancer. However, the studies were poorly reported and a minimum data set for study reporting was proposed.

CRD commentary

This review addressed a complex application of CT technology. Consequently the inclusion criteria for the review were very detailed, but were clearly described and justified. The search strategy and dates searched were appropriate to the subject of the review and no language restrictions were applied. The restriction to published, peer-reviewed studies might have resulted in the loss of some relevant data, particularly given the rapidly evolving nature of the field. Appropriate measures were undertaken to minimise the introduction of error and bias during the review process. The methodological quality of the included studies was assessed using tools specific to diagnostic accuracy studies, and the potential impact of elements of methodological quality upon estimates of accuracy was discussed in detail in the text; the data were insufficient to support formal investigation by inclusion in the regression model.

The meta-analytic methods used were appropriate and their application and meaning were clearly described. Interpretation may have been aided by more detailed reporting of the test and reference standard methods and characteristics of the participants in the individual included studies. The authors highlighted that not all relevant data were available, thus some caution is needed when interpreting the meta-analysis results, particularly for cancer detection. This was a good review, using a new method of meta-analysis for diagnostic tests, and the authors’ conclusions are likely to be reliable.

Implications of the review for practice and research

Practice: The authors stated that CT colonography had a high sensitivity and specificity for the detection of polyps larger than 5 mm and a particularly high sensitivity for the detection of cancer.
Research: The authors stated that more studies in asymptomatic individuals are needed. The potential of CT colonography as a diagnostic tool for cancer (rather than a screening tool) merits further investigation. The authors proposed a minimum data set for the reporting of CT colonography studies; this was described in detail in the discussion section of the article.

**Funding**
European Society of Gastrointestinal and Abdominal Radiology.

**Bibliographic details**

**PubMedID**
16304111

**DOI**
10.1148/radiol.2373050176

**Original Paper URL**
http://radiology.rsajnl.org/

**Other publications of related interest**

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Colonic Polyps /radiography; Colonography, Computed Tomographic; Colorectal Neoplasms /radiography; Humans; Sensitivity and Specificity

**AccessionNumber**
12005002431

**Date bibliographic record published**
31/07/2007

**Date abstract record published**
31/07/2007

**Record Status**
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.