Authors:

Ethna McFerran, BSc (Corresponding Author)
Dr James F O’ Mahony
Richard Fallis
Prof Duncan Mc Vicar
Prof Ann G. Zauber
Prof Frank Kee

Title:

Evaluation of the effectiveness and cost-effectiveness of personalised surveillance after colorectal adenomatous polypectomy
Acknowledgements

Authors affiliations:

1. Ethna McFerran, BSc Hons, RN, PG Dip, PhD Candidate, Centre of Public Health, Queen’s University Belfast (Corresponding Author, Corresponding email address for Publication - emcferran01@qub.ac.uk c/o 1233 York Ave, Apt 2D, New York, NY 10065, Phone No: +16465832837 (USA)/ +0447966267586(UK)) *Corresponding Author

2. Dr James F. O’ Mahony, Centre for Health Policy and Management, Trinity College Dublin.

3. Richard Fallis, Medical Library, Queen’s University Belfast

4. Prof Duncan Mc Vicar, Queen’s Management School, Queen’s University Belfast

5. Prof Ann G. Zauber, Attending Biostatistician, Memorial Sloan Kettering Cancer Center

6. Prof Frank Kee, Director, UKCRC Centre of Excellence for Public Health Research (NI) and Deputy Director, Centre for Public Health Queen’s University, Belfast

Running Head:
Post-polypectomy surveillance cost-effectiveness

Grants

Ethna McFerran was supported by the award of Health and Social Care Northern Ireland and National Cancer Institute Health Economics Fellowship (Grant Number: CDV/4980/14). Prof Zauber acknowledges funding in part through the National Institutes of Health /National Cancer Institute Cancer Center Support Grant P30 CA008748.

Word count 6007, excluding abstract (Word count =198), tables and appendices.
Abstract

Lifetime risk of developing colorectal cancer is 5% and five-year survival at early-stage is 92%. Individuals with pre-cancerous lesions removed at primary screening are typically recommended surveillance colonoscopy. Since greater benefits are anticipated for those with higher risk of colorectal cancer, scope for risk-specific surveillance recommendations exists. This review assesses published cost-effectiveness estimates of post-polypectomy surveillance to consider the potential for personalised recommendations by risk-group.

Meta-analyses of incidence of advanced-neoplasia post-polypectomy for low-risk cases were comparable to those without adenoma; with both rates under the lifetime risk of 5%. This group may not benefit from intensive surveillance, which risks unnecessary harms and inefficient use of often scarce colonoscopy capacity. Therefore, greater personalisation through de-intensified strategies for low-risk individuals could be beneficial. The potential for non-invasive testing such as faecal immunochemical tests combined with primary prevention or chemoprevention may reserve colonoscopy for targeted use in personalised risk-stratified surveillance.

This review appraised evidence supporting a program of personalised surveillance in patients with colorectal adenoma according to risk-group and compared the effectiveness of surveillance colonoscopy with alternative prevention strategies. It assessed trade-offs between costs, benefits and adverse effects which must be considered in a decision to adopt or reject personalised surveillance.

Key Words:
Colorectal cancer, adenoma, cost-effectiveness, Precision Medicine, early detection, cancer prevention, surveillance
Background

Lifetime risk of developing colorectal cancer (CRC) is 5% for an average risk individual in the US\(^1\). CRC is the third most common cancer globally and imposes a significant burden of ill-health\(^2\). Worldwide, CRC deaths form 8.5% of total cancer deaths (694,000 annually)\(^3\). Many deaths could be avoided by early detection through screening\(^4,5\); as given five-year relative survival rates for CRC detected at a local stage are 92%\(^6\).

Screening programs have been widely implemented to manage CRC risk\(^7\). Such programs employ colonoscopy either as the primary test or as a diagnostic test following a positive finding on a non-invasive stool test, which detects blood or other markers suggestive of cancerous lesions. Colonoscopy offers direct visualisation and examination of the entire colon permitting the identification and removal of polyps leading, it is thought, to the prevention of CRC\(^5\).

There are concerns over claims that screening programs reduce mortality or improve survival\(^8\), based largely on arguments related to lead time bias. Lead time bias occurs when a diagnostic test merely identifies the disease earlier, thus increasing perceived survival without significant modification of the disease course\(^9\). Despite such concerns, a recent meta-analysis of randomised screening trials (which addressed the effect of lead-time bias,) showed that one CRC death is prevented for every 1000 people screened, with this benefit being manifest on average after 9.4 years\(^9,10\). Moreover, micro-simulation modeling is reported to show that declines in CRC death rates are consistent with a relatively large contribution from screening\(^11\). While there is considerable randomised control trial
evidence to support screening overall, the magnitude of benefit available for surveillance (in
terms of CRC deaths prevented) is uncertain.

Post-polypectomy surveillance by colonoscopy has become a common feature of CRC
prevention strategies, offering intensive monitoring to individuals with prior
precancerous findings at primary screening. In the case of colorectal screening,
appropriate surveillance after endoscopic diagnosis of an adenoma, is typically a strategy
of surveillance colonoscopy at intervals of between 3 and 10 years. Surveillance intensity
can be adjusted dependent on an individual’s estimated CRC risk, as predicted by the
number and grade of polyps removed at index colonoscopy. Despite being widely
recommended, the evidence that post-polyp surveillance reduces CRC incidence or
mortality is lacking and is rarely established for sub-groups.

Up to 85% of CRCs are thought to develop from conventional adenomas. Adenomas begin
in the glandular tissue lining the colon and while many are benign, some may have
malignant potential. Genetic changes in the colon’s lining can lead to malignancy as a result
of a complex multi-step process in which adenoma is an intermediate stage. A process
referred to as the adenoma-to-carcinoma sequence, taking an estimated 7 to 15 years.
The long preclinical sojourn time of many adenomas creates the opportunity for successful
early detection through screening. Reported adenoma prevalence is estimated at 20-53% in
persons over 50 years, with gender differences showing higher prevalence (40%) in men
than in women (29%).
Colonic polyps were conventionally classified as either hyperplastic or adenomatous, of which the latter were believed to have the potential to progress to carcinoma. Advances in genetic pathology are alleviating so called ‘variant classification’ which ‘obfuscated the correct classification’ of sessile serrated adenomas, which unfortunately, were not as readily detected by many screening tests. As new information emerges it is possible that sessile serrated lesions may be responsible for up to 30% of CRC. The implications of the different pathologies for clinical management warrant the vigilance of physicians who may consider follow-up colonoscopies in accordance with sessile serrated adenomas guidelines. Although sessile lesions may have greater contribution to CRC than previously thought, this review focuses on the evidence related to adenomatous polyp risk groups.

Although there is limited decisive evidence from colon-polyp surveillance, current guidelines for post-polypectomy surveillance employ explicit risk stratification by sub-groups, using the predictive features of adenomas detected at screening colonoscopy. The size, the number of polyps and their histology provide further qualification in differentiating those with tubular features from those with villous features, considered more likely to have cancers develop in them. For example, US guidelines recommend that individuals with 3–10 adenomas undergo a surveillance colonoscopy every 3 years, while those with 1–2 tubular adenomas <10mm receive a surveillance colonoscopy every 5-10 years. Surveillance colonoscopies account for approximately 25% of colonoscopies among people over 50 years in the US.
While the surveillance guidelines are clear, conflicting reporting might lead to a conclusion that these persons are at a significantly increased risk, whilst other reports contend that many of the lesions detected at screening are likely to be of low risk. It has been suggested that following initial detection and removal of adenomas, approximately half of people (51.4%) will have further adenomas within 3 years of initial colonoscopy, of which significant numbers may meet at least one criterion for advanced adenoma\textsuperscript{13,32–35}. However, 84% of all polyps removed at colonoscopy in a large screening study of 13,992 participants were less than 10mm. Within a subset of the study population, CRC was detected in 0.03% of participants whose largest polyp was 1-5mm, (1 patient amongst 3744 patients with polyps 1-5mm), moreover only 3 of the 74 cancers detected were found as a consequence of detecting advanced adenomas\textsuperscript{17,36}. Consequently, screening typically generates many ‘positive’ findings that ultimately may be of low-risk, accounting for a small portion of cancer cases, meaning that large numbers of patients will be referred to surveillance, the clinical utility of which can be debated\textsuperscript{17,36}.

Whilst one benefit of surveillance is the possibility to detect lesions of significance, it may expose patients to unnecessary risks as a result of overdiagnosis, that is, the inclusion of ‘pseudodisease,’ that would not become evident before the patient dies of other causes\textsuperscript{37}. For example, it was reported that CRC was diagnosed in 19 of 2915 patients, who were deemed free of remaining lesions at a baseline clearing colonoscopy, over a mean follow-up of 3.7 years (incidence, 1.74 cancers/ 1000 person-years) amongst those in close surveillance. Equating to 0.65% of atypical post-polypectomy surveillance participants developing CRC\textsuperscript{38}, this includes a considerable numbers of individuals who undergo a surveillance test who could therefore be considered subject to over-diagnosis.
Some regions are adopting resect and discard policies, whereby lesions judged by the clinician performing polypectomy not to be of high risk can be discarded without being evaluated by a pathologist, thus reducing the risk of procedural over diagnosis. Another obvious means to lower potential overdiagnosis and limit the harms of invasive testing might be to consider an alternative to colonoscopy and to personalise approaches to surveillance by exploring the role of faecal immunochemical testing (FIT). In a recent systematic review, FIT shows high diagnostic accuracy for detecting CRC and has shown the capability of quantifying and adjusting cut-off concentrations for positivity. Moreover, its acceptability to patients has also been demonstrated. Therefore, FIT could be an appropriate, acceptable and cost-effective surveillance test.

Decision making requires careful balancing to avoid either too little surveillance, which may jeopardise CRC prevention goals, or lead to overuse of surveillance, chancing unnecessary harms and inefficient use of colonoscopy resources. Health economic evaluations aim to impartially identify, measure and compare the cost and consequences of the different interventions being considered to manage particular clinical problems. Recent economic evaluation in the US estimated an inflection point between conferring benefit and risking harm in the use of colonoscopy in older adults, whereby the anticipated harms of false positives and unnecessary investigations outweighed the benefits of early detection.

The relevant resource utilisation relates not only to the financial costs of providing surveillance, but also to colonoscopy capacity, which is often constrained in many health systems. Therefore, decision makers need to consider how best to allocate the limited
number of colonoscopy examinations to those individuals with the greatest likelihood of benefit.

Consensus has not yet emerged on what personalised surveillance practice ought to involve, with variation in current guideline recommendations shown in Table 1. For example, Japan does not differentiate its surveillance guidance by risk category; recommending colonoscopy every three years, whereas the UK recommendations vary between annual colonoscopy for high-risk patients and five year colonoscopy (or return to screening) for low-risk patients. Concerns over how best to balance surveillance intensity will be increasingly pressing given anticipated growth in numbers of people being directed into surveillance colonoscopy, in part due to demographic aging and changes in the primary screening technology employed.

Current data suggest that screening colonoscopy may identify patients at low risk of death from colorectal cancer or who may derive greatest value from a single screening test, but who may not benefit from subsequent intensive surveillance. Although meta-analysis of incidence of advanced neoplasia after polypectomy for a low-risk individual is comparable to persons without findings of an adenoma at colonoscopy, absolute risk in both groups was under the average persons' lifetime risk of 5% (low-risk 3.6% vs without adenoma 1.6%). This indicates that the low-risk group may indeed have a CRC risk that is broadly comparable to the average risk population eligible for primary screening. For that reason, there may be arguments for de-intensifying surveillance towards the types of screening frequencies and non-invasive testing technologies used in primary screening, which in turn would lead to greater personalisation of colonoscopy use.
Existing approaches to the adjustment of surveillance intensity rely on the frequency of testing, that is, through changes to the interval of use of the current technology (colonoscopy), offering, for example, 3 and 5-10 year colonoscopy. The ability to vary surveillance has been limited to this interval approach. Newer, more effective stool tests may offer the ability to change the type of test offered, which may add flexibility to surveillance programs and as a result reduce the number of colonoscopies required during surveillance.

Accordingly, this systematic review has three aims:

1. To assess if there is sufficient evidence to evaluate a program of personalised surveillance in patients with colorectal adenoma according to risk sub-group.
2. To compare the effectiveness of surveillance colonoscopy with alternative prevention strategies.
3. To assess trade-off between costs (resource use), benefits and adverse effects that need to be considered in a decision to adopt or reject personalised surveillance.

Methods

Data Sources and Search Strategy

The systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidance recommendations and the Centre for Reviews and Dissemination guidance. The review has been registered with PROSPERO – reference: CRD42016033509.
An initial check for previous reviews on the topic was conducted, as recommended\textsuperscript{46,54}. The search for the key words ‘adenoma’ AND ‘cost’ in ANY FIELD (September 2015), was carried out within the Centre for Reviews and Dissemination database including all databases (DARE, NHS EED and HTA; those most specific to economic evaluations of health and social care interventions)\textsuperscript{55}. This search indicated no existing systematic reviews addressing cost-effectiveness within colorectal adenoma surveillance and prevention programs.

The systematic review search strategy was optimised with help from a Specialist Medical Librarian (RF), informing the choice of available databases and developing the search to meet the needs of the review. The search strategy was run in MEDLINE, MEDLINE in-process and EMBASE. These databases were searched from their inception to February 2016, for key words, medical subject heading terms and synonyms of:

(a) Colorectal neoplasms OR adenoma.
(b) Costs-benefit analysis OR synonyms.
(c) Early detection of cancer OR surveillance.

Searches a, b and c were then combined with AND, as shown in Web Appendix 1.

In order to optimise the resultant yield of studies, we expanded the medical subject heading terms, used a modified strategy in each database (MEDLINE / EMBASE) to identify the literature under relevant terms and included techniques for word proximity and suffixes, which optimised database search tools to find relevant papers.
The titles and abstracts of the studies returned by the database searches were then screened for inclusion by eligibility criteria according to the patient population or the disease being addressed (P) the interventions or exposure (I) the comparator group (C) the outcome or endpoint (O) the study date / time frame (T) and the study design chosen (S) – ‘PICOTS’ criteria, as shown in Table 2. The reference lists of the retrieved studies were searched to find studies not captured by our database searches.

Study selection was conducted in three stages, as shown in Figure 1, including removal of duplicates (n=264), title and abstract screening against the PICOTS criteria (n=1009) and independent screening of all full text articles (n=32) to confirm their eligibility, by two reviewers (EMF / JFOM); conducted according to the selection criteria detailed in Web Appendix 2. In order to minimize bias, studies were retained in situations where both reviewers were not in agreement on exclusion, with discrepancies resolved by adjudication with a third reviewer. All excluded papers were codified by ineligibility of PICOTS category. This process resulted in n=7 papers that were fully evaluated for the review.

Data Extraction and Identification of Cost-Effectiveness Analyses

We extracted the initial data from each study using the Consolidated Health Economic Evaluation Reporting Standards statement checklist tool. We have not conducted a meta-analysis as the outcomes of economic evaluations are typically not commensurate for comparison. Some studies reported incremental cost-effectiveness ratios (ICERs) that differ from the conventional interpretation, as the ratio of incremental costs to incremental health effects, relative to the next most effective strategy, whereby strategies that are more costly and less effective are ruled out by simple or extended dominance. In these
instances, the ICERs were recalculated from the reported costs and effects and replicated cost-effectiveness estimations were used to re-examine the comparisons and analyses made by the studies, as carried out in another recent review\textsuperscript{61}. The recalculated results are presented alongside the originally published results in Web Table 1.

**Results**

**Study Descriptions**

The systematic review returned 7 papers that were fully evaluated. An overview of the key quality attributes of each paper as assessed in this review is given in Web Table 2, following the Consolidated Health Economic Evaluation Reporting Standards quality indicators\textsuperscript{48,50,52–57}. The studies were published between 1991 and 2011; no studies from more recent years were identified.

In brief, the search returned a small number of studies and the prevention strategies compared in the studies varied such that not all compared the same alternative interventions. Thus, the potential for cross-comparison of the effectiveness and cost-effectiveness of particular strategies was limited. Whilst some papers compare surveillance by colonoscopy to natural history, others model compared surveillance by colonoscopy to a screening colonoscopy a 10 year interval\textsuperscript{62}, or for performing an early 1 year colonoscopy\textsuperscript{63}, whilst other models compared surveillance colonoscopy combined with chemo-prevention\textsuperscript{64,66} and chemo-prevention alone compared to natural history\textsuperscript{65}.

Strategies considered include:

- one year surveillance by colonoscopy\textsuperscript{63},
• a three-year high-risk and five-year low-risk colonoscopy\textsuperscript{62},
• a three-year high-risk and ten-year low-risk colonoscopy\textsuperscript{62},
• a three-year high-risk and three-year low-risk colonoscopy\textsuperscript{62},
• aspirin as chemoprevention alone\textsuperscript{66},
• aspirin therapy combined with colonoscopy\textsuperscript{66},
• celecoxib as chemoprevention alone\textsuperscript{65},
• a three-year high-risk colonoscopy \textsuperscript{65},
• calcium as chemoprevention alone\textsuperscript{64},
• calcium therapy combined with colonoscopy\textsuperscript{64},
• fixed interval / modified interval colonoscopy surveillance\textsuperscript{67}.

To address the primary aims of the review in a systematic way, the following sections critically address how respective papers’ methods, assumptions and outputs support or prohibit clear evidence for each objective.

**Evidence to support personalised surveillance by sub-group at index colonoscopy**

No papers reported cost-effectiveness results disaggregated by high-risk/ low-risk sub-groups. While two studies described clear elements of stratification, identifying high-risk and low-risk subgroups of patients with adenoma, neither reported a comparison of outcomes by these subgroups; rather they reported results as combined group data\textsuperscript{59,64}. Accordingly, this limited what our review was able to determine regarding risk-optimised surveillance strategies.

The reporting in one paper did permit a step wise comparison of interval change for surveillance by colonoscopy in high-risk and low-risk groups\textsuperscript{62}. The ICERs presented
indicated that it is beneficial to change from a 10 year interval colonoscopy to a 3 year interval for high-risk individuals, as this strategy was more effective and its ICER of $5743/QALY indicated that it is a cost-effective policy change, within conventional thresholds thought to be at least $50,000/QALY. Whilst it is also beneficial to move from 10 year interval colonoscopy to 5 year in low-risk individuals, the ICER of $296,266/QALY is greater than conventionally accepted thresholds for the US. Importantly, these results also indicated that more intensive surveillance by a change from a 5 year to 3 year interval for low-risk individuals resulted in reduced quality adjusted life years, (-0.0023 QALYs). This ‘disutility of colonoscopy’, shows that it becomes more harmful for low-risk individuals to receive a more intensive surveillance strategy of a 3 year colonoscopy.

**Effectiveness of colonoscopy compared to alternative prevention strategies**

An important purpose of this review was to find studies that compared alternatives to colonoscopy-based surveillance. The review found no studies that considered other clinical test strategies in post-polypectomy surveillance other than colonoscopy. All papers retrieved assumed that the default test for surveillance was colonoscopy. There are, however, comparisons of colonoscopy to three types of chemoprevention drugs, all of which compared chemoprevention benefit to no intervention, or compared colonoscopy combined with chemoprevention to no intervention. A summary of the results from the strategies evaluated for surveillance is shown in Web Table 1.

In addressing clinical variations in colonoscopy capacity, the most recent paper authored by Wilschut and colleagues used micro-simulation modelling with the MISCAN-Colon model (one of 3 internationally validated models which evaluate screening programs). This
included 48 variations of the background screening program within which 2 surveillance strategies were simulated. Although this study presented results for variation in the primary screening strategy, the reported results do not permit comparison of the two surveillance strategies considered. The analysis considered whether it would be appropriate to offer colonoscopy surveillance under increasingly tight colonoscopy capacity constraints. They found that an affordable ICER was achievable for colonoscopy surveillance when capacity was greater than 20 colonoscopies per 1,000 individuals. However, if the capacity of colonoscopy was <5 per 1,000 individuals offering low-risk groups surveillance colonoscopy was no longer considered an effective allocation of a scarce health resource.

Wilschut et al.’s analysis adds a modelling feature, not commonly employed in the other papers, that permits the simulation of the impact of both primary screening and subsequent surveillance. Moreover, it has the ability to evaluate issues of service capacity, alternate types of testing and a mix of tests which more accurately reflects the complexity of choice facing decision makers. By comparison, the models used in other studies reviewed only characterise limited aspects of the decision problem.

Hassan et al estimated the benefit of early annual colonoscopy compared with not doing an early annual colonoscopy, since their descriptions are not clear we clarify that they compare providing a 1 year to a 3 year test. They report an ICER of $66,136 per life year gained (LYG) for a comparison of annual colonoscopy to no yearly test (where ‘no test’ is modelled as a 3 year test). However, the paper did not report total costs or total effects for the strategies considered; consequently, it was difficult to assess this ICER or its basis. The modelling conducted in this comparison is for persons aged 60 years on entry to the
surveillance program. This comparison may have somewhat limited clinical relevance in its chosen setting, as the recommended age to start screening in the US is 50 years\textsuperscript{7}. The finding that an annual colonoscopy may be cost-effective relative to a three year colonoscopy is in keeping with the results of an application of the U.K. guidelines in the U.S. which suggested a subset of high risk patients may warrant a one-year clearing colonoscopy\textsuperscript{69}.

**Chemoprevention**

Although none of the reviewed studies considered tests other than colonoscopy, a range of chemoprevention strategies were evaluated\textsuperscript{64–66}, one of which demonstrated that a strategy employing aspirin combined with colonoscopy is cost-effective\textsuperscript{66}. Focusing on the absolute differences in benefit, this study estimated that compared with no intervention, colonoscopy surveillance accrued +0.0124 life years saved (LYS) whilst aspirin combined with colonoscopy surveillance provided +0.0138 LYS\textsuperscript{66}.

DuPont et al reported ICERs for aspirin alone, colonoscopy surveillance alone and a combined intervention of aspirin with colonoscopy surveillance, which showed an ICER of $87,609/ LYS, $78,226/ LYS, $60,492/ LYS respectively\textsuperscript{66}. These ICERs however appear to have been calculated differently from the conventional interpretation\textsuperscript{58}. As such, the reported ICER in the paper effectively becomes an average cost-effect, that is, the ratio of the cost to benefit of an intervention without reference to a comparator\textsuperscript{70}. Accordingly, we recalculated the ICERs from the reported costs and effects and the replicated cost-effectiveness estimations plotted on a cost-effectiveness plane. These re-estimated ICERs are reported in Web Table 1 alongside the reported figures from the paper. This
reinterpretation of the results indicates that aspirin chemoprevention alone was subject to extended dominance, as was colonoscopy surveillance alone, meaning that they are not preferred from the cost-effectiveness perspective. The combination of 3/5yr colonoscopy combined with aspirin had an ICER of $73,927/ LYS and as such remains a cost-effective strategy for the US. This result shows that combination therapy is more cost-effective than either intervention alone, which is noteworthy and merits further investigation, given the role of aspirin in the prevention of premature mortality due to other causes.

Arguedas and colleagues compared colonoscopy surveillance with no surveillance and demonstrated an incremental benefit of 0.01995 LYS (8.48482 life years vs 8.45487 life years), whilst celecoxib was estimated to provide a greater absolute gain in LYS, generating a further 0.00579 LYS relative to colonoscopy surveillance. Although celecoxib chemoprevention was estimated to be more effective than colonoscopy, the ICER of $1,715,199/ LYS was significantly above US thresholds.

Notably the DuPont et al, aspirin paper and the Arguedas paper evaluating chemoprevention using celecoxib, shared co-authors and employed similar models. Whilst the DuPont et al title described addressing increased risk for CRC, the Arguedas et al paper described average-risk patients, however both models explained colonoscopy surveillance as colonoscopy ‘occurring 3 years after index colonoscopy’. Such description left it unclear if individuals eligible for a 5 year surveillance test were included. Whether these models in fact incorporated only those in the high-risk group, according to US guidelines, was not fully supported by the parameter estimates for the models. The cited reference for malignant transformation rate (0.1065) was taken from published data.
for untreated polyps, rather than reported high-risk transformation. The probability reported differed in the aspirin analyses, where malignant transformation was reported as 0.01, citing one shared reference, consequently in the absence of clear reporting, we cannot draw any firm conclusions about whether all those eligible for surveillance were modelled in either paper.

The effectiveness of supplemental calcium as a chemoprevention was evaluated by Shaukat et al. That analysis assessed a dose of 1.2g/day for age 50-80 years, not at the 3-4g/day dose mentioned in the article as providing a reduction in adenoma recurrence of 22% compared to placebo in meta-analyses. The article does not present a clear argument for using the lower dose selection.

Like DuPont et al, Shaukat et al report an ICER for calcium chemoprevention alone however this strategy is subject to extended dominance and so would not be a preferred strategy and should not have an ICER reported for it. The recalculated ratios of costs and effects were replicated to provide cost-effectiveness estimations which were plotted on a cost-effectiveness plane, reported in Web Table 1. Based on the recalculated ratios the resultant ICER for surveillance colonoscopy was $20,494/LYG when compared to natural history, with the combination of calcium chemoprevention and colonoscopy generating an ICER of $2,823,333/LYG, based on the 0.0003 incremental LYG reported, an ICER which is once again greater than US thresholds.

This reassessment of the reported results indicated that surveillance colonoscopy alone is cost-effective, whilst the ICERS indicated that the incremental cost, of additional health
benefits from chemoprevention by celecoxib alone or calcium combined with colonoscopy was likely to be very high, relative to the health gains. It can tentatively be claimed that aspirin chemoprevention combined with surveillance colonoscopy appears to be cost-effective, but given the ambiguity regarding risk-groups within the DuPont et al paper, the results merit further investigation to clarify if these are sub-group dependent or if they might apply to all adenoma patients.

From the review we believe the salient points from the conclusions of these cost-effectiveness evaluations of colonoscopy based surveillance programs to be:

(a) Colonoscopy capacity can, at lower levels, prohibit the ability of health systems to offer colonoscopy based surveillance to low-risk groups.

(b) Compared with a ten-year low-risk colonoscopy, offering a five-year colonoscopy to low-risk groups was above US thresholds at $296,266/ quality adjusted life year.

(c) Compared to a three-year high-risk colonoscopy, there is evidence to support offering a one-year high-risk* colonoscopy – *for persons aged 60 years entering surveillance.

(d) Aspirin combined with surveillance colonoscopy generated greater life years saved than aspirin or colonoscopy alone and in given its role in the prevention of premature mortality due to other causes, this combination merits further evaluation.

There were quality and reporting issues with a number of the papers evaluated. These shortcomings suggest that questions remain regarding the cost-effectiveness of post-polypectomy surveillance programs.
The trade-off between costs (resource use) and beneficial or adverse effects that need to be considered in a decision to adopt or reject personalising surveillance

Cost calculations of the strategies should account for all resources used. Whilst all models included the costs of colonoscopy and polypectomy, program and administration costs were only described in two papers\textsuperscript{59,62}. Only one reviewed study attempted to address the treatment costs, accounting for newer therapies such as oxaliplatin, now recommended in advanced stage cancer, and terminal care costs\textsuperscript{59,71,72}. Where these costs were estimated for the final year of life, there was some uncertainty as to how these were adjusted for according to heterogeneity by stage\textsuperscript{59}. There was no consistent approach to adjusting treatment costs according to the stage of disease\textsuperscript{62–64,72}. Adjustments for inflation were also unclear in some of the papers\textsuperscript{59,67}. Use of biologics, such as Cetuximab or Bevacizumab in treatment costs assumptions was not noted.

Study costs were commonly taken from Medicare fee schedules for colonoscopy, polypectomy, complications and pathology\textsuperscript{62–66}, or in some cases national reports\textsuperscript{62}. Only one study reported the type of distribution used for costs in probabilistic sensitivity analyses\textsuperscript{62}. Indirect costs, in the form of lost income to the patient and an escort, were included in only one study\textsuperscript{63}. Somewhat strangely, one study cited a long term arthritis trial for their $100,000 costs per CRC case, the provenance of which was uncertain given the source cited\textsuperscript{65}.

Resource costs for aspirin were given from a trial with wholesale prices used in sensitivity analyses\textsuperscript{66}. Calcium costs were described as constant over the period 2005-2008 prices\textsuperscript{64}. There were some inconsistencies in referenced costs for “incurable” CRC\textsuperscript{66}, citing a base
case scenario ($40,000) with a maximum in the range ($100,000) from a source that used this maximum as its base case\textsuperscript{65}.

Health effects were calculated based on the estimated effect of colonoscopy and polypectomy and weighted by the risk of adenoma transformation in all models. The use of a preference-weighted health state classification system such as the EuroQol-5D\textsuperscript{73} were not consistently reported. No citations were presented for the utility estimates used in some models (for CRC at diagnosis and subsequently)\textsuperscript{59} while in others no measures for utility were given\textsuperscript{63}.

We noted a large difference in the modelled life expectancy (between 8.45/ LYS\textsuperscript{65} and 12.28/ LYS\textsuperscript{66}) under no surveillance of celecoxib and aspirin from two studies that used related models in which the same discount rate was used and individuals were modelled from age 50 in both cases. While the difference in life expectancy may relate to differences in risk subgroups between the analyses the difference still seems large, and was not readily explained\textsuperscript{65,66}. Whilst there is a 6 year gap in publishing, it is unclear whether this difference can be directly attributed to the characteristics modelled, surveillance program or to differences in the quality or practice of colonoscopy techniques over time\textsuperscript{24,74}, or to treatment improvements\textsuperscript{72}.

It is inevitable that colonoscopy carries the risk of missed lesions, given as approximately 22\% by meta-analyses\textsuperscript{75,76}. Missed polyps clearly have the potential to become interval cancers. Only two of the studies reported a probability of a missed polyp; and there was a
noticeably large variation ranging from 0.08-0.21\textsuperscript{62,66} (where reported as a percentage, small adenoma=17.8\% and large adenoma=4.6\%\textsuperscript{64}). The remaining studies have not reported this within model parameters and it this implies it is not assessed within the analyses\textsuperscript{59,63,65}.

The risk of colonic perforation, as an adverse effect, was considered in all but one of the models\textsuperscript{63}. This was modelled with various probabilities; a base case probability of 0.0006\textsuperscript{62}, 0.003 for colonoscopy alone\textsuperscript{65,66}, or 0.02 with polypectomy\textsuperscript{65}. The origins of these rates are uncertain from the reported literature. Although relatively rare, perforation can cause significant morbidity and even death (30 day morbidity rates of 21\%-53\% and mortality rates of 0\%-26\%, with hospital stay of up to 3 weeks\textsuperscript{77}).

**Discussion**

The main policy-relevant issue emerging from this review was that no studies were found that evaluated the cost-effectiveness of colonoscopy against other tests, such as FIT or other non-invasive testing. Colonoscopy has been the primary approach to post-polypectomy surveillance since the early 1990s but it has not been compared with other tests in the surveillance of patients after polypectomy. This is in spite of the availability of alternatives, such as FIT, which have been compared with colonoscopy in index screening evaluations\textsuperscript{78-82}.

Critically we acknowledge the gaps in cost-effectiveness reporting by sub-group. Since it is possible to implement different treatment decisions for patients with different characteristics, models should consider the potential for their results to vary across different subgroups to facilitate different policy decisions\textsuperscript{83}. As demand for testing changes over time
in screening programs, through the introduction of newer technologies and with trends in adherence and variable adenoma detection rate\textsuperscript{84}, these issues require attention from policy makers and modellers to understand and explore the potential of modelling to provide a clear understanding of the risks and benefits in the choice of interventions adopted.

Prior work has shown that FIT threshold for positivity can be adjusted within a screening program to optimise detection according to available colonoscopy capacity\textsuperscript{85}, therefore post polypectomy surveillance could follow such an approach. The role of FIT is being considered in surveillance with a trial in the UK currently comparing FIT vs colonoscopy\textsuperscript{86}. FIT offers improved performance over older stool-based testing techniques and its ability to adjust cut-off levels may allow for greater optimisation of resources given colonoscopy capacity constraints.

The UK NHS Bowel Cancer Screening Programme recently recommended the primary test used be changed from guaiac-based faecal occult blood testing (gFOBT) to FIT\textsuperscript{87–89}. Such a change in the primary test used will likely affect the numbers of patients detected with advanced adenoma, and with it those eligible for surveillance\textsuperscript{90}. As part of this change there are planned adjustments to the FIT positivity cut-off value used, in order to continue to optimise the effectiveness of the planned technology in line with capacity changes and service transition. These recommendations have acknowledged the likely systemic effect on colonoscopy capacity; as such it would seem pragmatic to consider not only the adjustment of FIT cut-off for screening but also its role within the surveillance context. Whether
surveillance guidelines might be developed or modified to account for colonoscopy capacity
is one issue that might be explored in future modelling studies.

FIT has the potential to be an effective post-polypectomy surveillance test for suitable risk-
groups. Reported uses in screening other high-risk groups (e.g. first-degree relatives of
patients with CRC) has revealed that annual FIT screening (over 3 years) detected all CRCs
and proved equivalent to colonoscopy in detecting advanced neoplasia. FIT, when used
between scheduled surveillance colonoscopies, has been shown to have detected neoplasia
sooner than scheduled surveillances. Interval FIT analyses could be effectively used to
detect missed or rapidly developing lesions in surveillance programs. FIT has a useful
diagnostic role and it has also been suggested that FIT has a predictive capacity, with
interval cancers independently predicted by faecal haemoglobin concentration (FHbC),
which may be applied for tailored case management and modification based on FHbC.

The use of existing tests such as FIT, in innovative and adaptive ways, might help accrue
benefits in more risk appropriate, prescribed and personalised surveillance-based
approaches. Addressing and personalising other known features of risk of CRC, such as diet
and lifestyle, might offer increased precision and optimise the prevention of CRC. Offering
personalised surveillance with diet and lifestyle evaluation as a companion to non-invasive
testing alternatives might support adopting a primary care rather than secondary care
service design for prevention interventions to address the risk of colorectal cancer.
In future, there may be more scope for increased personalisation of surveillance programs. Novel blood based tests such as predictive micro-RNAs, or combined biomarkers (β-catenin nuclear localisation, Cox-2 expression and p53 nuclear expression, were significantly associated with adenoma recurrence after 3 years (β-catenin: p=0.002; Cox-2: p=0.001; p53: p=0.001). These tests put forward predictions of adenoma recurrence with high negative predictive value (88.5%) and sensitivity (94.6%), which if validated, would be equivalent to or better than current clinical risk stratification approaches based on adenoma size and frequency\textsuperscript{28,32}.

**Clinical Issues**

The most clinically-relevant issue raised by this review is that of the role of aspirin chemoprevention, recently endorsed by the updated US Preventative Task Force Recommendations\textsuperscript{96} and described as the first pharmacological agent to be endorsed for cancer chemoprevention\textsuperscript{97}. We have highlighted that aspirin combined with colonoscopy surveillance results in a reported ICER of $60,942 (recalculated to be $73,927/ LYG), in what we might reasonably infer to be high-risk groups and might be considered a strategy for personalised surveillance. Since some methodological issues were raised in the model reviewed within this paper, we believe it is highly relevant to consider an updated model which addresses the role of aspirin, taking cognisance of the known likelihood of a future precision medicine approach that is based on aspirin’s mechanism of action.

The other key clinical issue highlighted in the review, was how readily results may be affected by differences in capacity of colonoscopy services or may be influenced by other
quality assurance issues such as adenoma detection rates. As shown in other evaluations of screening, the adenoma detection rate was recognised as influencing the cost-effectiveness of screening programmes\(^{84}\). There are recognised differences in this rate between screening and surveillance, which was significantly higher in surveillance colonoscopies (37%), compared with screening colonoscopies (25%; P < .001)\(^{24}\). Future work acknowledging the impact of examination quality as characterised by adenoma detection rates, within decision models or colonoscopy capacity planning would allow robust evaluation of the benefits of surveillance. In so doing, we can more fully evaluate if infrequent high-quality colonoscopy exams are indeed more effective in preventing CRC than are frequent low-quality colonoscopy exams\(^{35}\).

**Limitations**

Potential limitations of the review are that as a result of our search strategy we do not characterise the grey literature related to the economic evaluation of surveillance in colorectal adenoma post-polypectomy surveillance.

**Conclusion**

We suggest a cautious interpretation of the findings of cost-effectiveness of colonoscopy-based post-polypectomy surveillance due to the small number of studies addressing the topic. Based on the reviewed literature we would suggest that future investigations update and confirm the benefits reported, in particular exploring comparisons of the cost-effectiveness of newer testing alternatives, such as FIT or newer tests like micro-RNA. In particular, we suggest examination of where FIT may provide clinically accessible adjustments to cut-off levels, and triage national or regional resources optimally based on
national or regional quality indicators and capacity. The insights on cost-effectiveness of combined aspirin and colonoscopy merit further exploration in light of the updated literature on the role of aspirin in chemoprevention and its likely role the in prevention of premature mortality due to other causes. Taken together, these results suggest that there are valuable alternatives to current guidelines which should be explored in updated cost-effectiveness models.
### Table 1 – Guidelines for surveillance following polypectomy

<table>
<thead>
<tr>
<th>Location</th>
<th>Year</th>
<th>Surveillance Recommendations</th>
<th>Interval</th>
<th>Reference No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK / New Zealand</td>
<td>2011</td>
<td>Low Risk - one or two adenomas smaller than 10 mm.</td>
<td>Consider colonoscopy at 5 years or return to screening (by gFOBT)</td>
<td>98,99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate Risk - three or four adenomas smaller than 10 mm or one or two adenomas if one is 10 mm or larger.</td>
<td>3 year colonoscopy</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High-risk - five or more adenomas smaller than 10 mm or three or more adenomas if one is 10 mm or larger.</td>
<td>1 year colonoscopy</td>
<td></td>
</tr>
<tr>
<td>US</td>
<td>2012</td>
<td>No polyps / distal small (&lt;10 mm) hyperplastic polyps.</td>
<td>10 year colonoscopy</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1–2 tubular adenomas &lt;10mm</td>
<td>5-10 year colonoscopy</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3–10 adenomas</td>
<td>3 year colonoscopy</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;10 adenomas</td>
<td>&lt;3 year colonoscopy (states - ‘no basis for less than 3 years,’ &lt; symbol as shown in paper).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>one or more tubular adenomas &gt;10 mm / one or more villous adenomas / adenoma with high grade dysplasia</td>
<td>3 year colonoscopy</td>
<td></td>
</tr>
<tr>
<td>European Society of Gastrointestinal Endoscopy (ESGE)</td>
<td>2013</td>
<td>Low risk group (patients with 1–2 tubular adenomas&lt;10mm with low grade dysplasia),</td>
<td>Participation in existing National screening</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High-risk group (patients with adenomas with villous histology or high grade dysplasia or ≥10mm in size, or ≥3 adenomas)</td>
<td>programmes 10 years after the index colonoscopy.</td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>2011</td>
<td>Low risk adenomas (patients with one or two small (&lt;10 mm) tubular adenomas).</td>
<td>Colonoscopy at 5 years</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High-risk adenomas (three or more adenomas, ≥10mm, or with tubulovillous, or villous histology, or high grade dysplasia)</td>
<td>Colonoscopy at 3 years</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple (Five or more) adenomas</td>
<td>Follow up at 12 months</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Possible incomplete excision adenoma</td>
<td>Colonoscopy 3-6 months</td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>2015</td>
<td>Uncategorized - Comments: Management of diminutive adenoma (&lt;5 mm) has not been established. In brief, there is no uniform Japanese approach (removal or follow-up) for diminutive adenomas, and controversy remains.</td>
<td>‘Follow-up colonoscopy should be performed within 3 years after polypectomy’</td>
<td>102</td>
</tr>
<tr>
<td>EU</td>
<td>2012</td>
<td>Low Risk (1–2 small adenomas)</td>
<td>Routine screening</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate Risk (3 or more adenomas or an adenoma ≥10 mm)</td>
<td>3-year interval to the first surveillance colonoscopy</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High-risk (5 or more adenomas or an adenoma of size 20 mm or larger).</td>
<td>An additional clearing colonoscopy at 12 months may be warranted</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cut-off age for stopping surveillance is usually 75 years – does not preclude further surveillance for clinical or other reasons.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>2013</td>
<td>Revised guideline 2002 onwards recommended, patients with three or more patients with fewer than three adenomas</td>
<td>3 years colonoscopy</td>
<td>45,103</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 years colonoscopy</td>
<td></td>
</tr>
</tbody>
</table>
Table 2 – PICOTS Criteria Applied

<table>
<thead>
<tr>
<th>PICOT Category</th>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Patients diagnosed with (resected) colorectal adenomatous polyp(s)</td>
<td>Patients with diagnosed colorectal cancer or sessile serrated adenomas*</td>
</tr>
<tr>
<td>Intervention</td>
<td>Interventions given for the management of colorectal cancer risk associated with the presence of a baseline adenoma, i.e. a follow up examination, surveillance test or reassessment by an appropriate means including colonoscopy and comparators listed below;</td>
<td>Interventions not currently in clinical use outside of trial for e.g. novel biomarkers</td>
</tr>
<tr>
<td>Comparison</td>
<td>Endoscopy, FOBT, FIT or CTC</td>
<td>Tests in development / biomarker based tests not currently in clinical use outside of trial for e.g. novel biomarkers</td>
</tr>
<tr>
<td>Outcome</td>
<td>Incidence of adenoma; recurrent /metachronous adenoma; colorectal cancer; ‘positive’ tests (in the case of qualitative FOBT, FIT, +/- other investigational tests) where a positive results indicates the need for further clinical investigation to treat/ resect potential lesions detected; Costs, LYG, Quality Adjusted Life Years, Disability Adjusted Life Years or other unit of health gain.</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>No time limits were imposed</td>
<td></td>
</tr>
<tr>
<td>Study designs</td>
<td>Economic evaluations where published as academic papers are eligible for inclusion</td>
<td>Case series, case reports, and reports from grey literature and conference proceedings; excluded from the review owing to the high potential for bias. RCTs and controlled trials reported effects and other formats than controlled trials, cohort studies, case–control whilst considered within the quality evaluation within models are not directly included.</td>
</tr>
</tbody>
</table>

FOBT = faecal occult blood testing, CTC = computed tomographic colonography, RCTs = Randomised controlled trials

* Sessile serrated lesions are often added to guidelines addressing the umbrella term polyp/adenoma, clear pathological and molecular distinctions are now recognised, thus we refer to comprehensive recent work on this pathology for further clinical - 25,26.
References


20. Stryker SJ, Wolff BG, Culp CE, Libbe SD, Ilstrup DM, MacCarty RL. Natural history of 

21. Rutter CM, Yu O, Miglioretti DL. A hierarchical non-homogenous Poisson model for 

2010;2(December):89. doi:10.3410/M2-89.


24. Anderson JC, Butterly LF, Goodrich M, Robinson CM, Weiss JE. Differences in 
detection rates of adenomas and serrated polyps in screening versus surveillance 
colonoscopies, based on the new hampshire Colonoscopy Registry. *Clin Gastroenterol 

25. Crockett SD, Snover DC, Ahnen DJ, Baron J a. Sessile Serrated Adenomas: An 


Adenoma and Incident Colorectal Cancer Based on Findings of the Baseline 

29. Leslie A, Carey FA, Pratt NR, Steele RJC. The colorectal adenoma-carcinoma sequence. 

for Colonoscopy Surveillance After Polypectomy: A Consensus Update by the US 
Multi-Society Task Force on Colorectal Cancer. *Gastroenterology*. 2012;143(3):844- 
857. doi:http://dx.doi.org/10.1053/j.gastro.2012.06.001.

doi:10.1016/j.gie.2014.01.014.

32. Brand L, Munding J, Pox CP, et al. β-catenin, Cox-2 and p53 immunostaining in 
colorectal adenomas to predict recurrence after endoscopic polypectomy. *Int J 

Patients with Low-Risk Polyps on Initial Colonoscopy. *J Clin Gastroenterol*. 

34. Lansdorp-Vogelaar I, Fedewa S, Lin CC, Virgo KS, Jemal A. Utilization of surveillance 

35. Karsa L von, Patnick J, Segnan N, European Colorectal Cancer Screening Guidelines 
Working Group. *Quality Assurance in Endoscopy in Colorectal Cancer Screening and


80. Quintero E, Castells A, Bujanda L, et al. Colonoscopy versus fecal immunochemical...
885 86. Atkin W. Faecal immunochemical testing (FIT) for adenomas surveillance / FIT for follow up study. NIHR Heal Technol Assess Program. 2011;(May).
894 94. Anderson AS, Craigie AM, Caswell S, et al. The impact of a bodyweight and physical activity intervention (BeWEL) initiated through a national colorectal cancer screening programme: randomised controlled trial. BMJ. 2014;348(March):g1823. doi:10.1136/bmj.g1823.
adenomas. 2011;(March).


Figure 1 - PRISMA Flow Diagram

- Records Identified Through Database Searching: n = 1,273
- Additional Records Identified Through Other Sources: n = 1
- Duplicates Removed: n = 264
- Records Screened (1st and 2nd Author): n = 1,010
- Records Excluded: n = 978
- Full-Text Articles Assessed for Eligibility (1st and 2nd Author): n = 32
- Full-Text Articles Excluded, with Reasons: n = 25
- Studies Included in Qualitative Synthesis: n = 7