**Economic Evaluations of Follow-up Strategies for Cancer Survivors: A Systematic Review and Quality Appraisal of the Literature.**

Michela Meregaglia a,b, John Cairns a

a Department of Health Services Research and Policy; Faculty of Public Health and Policy; London School of Hygiene and Tropical Medicine. Address: 15-17 Tavistock Place London WC1H 9SH (UK). Email: John.Cairns@lshtm.ac.uk. Tel: +44 (0) 207 927 2395. Fax: +44 (0) 207 927 2701.

b CeRGAS (Research Centre on Health and Social Care Management), Bocconi University. Address: Via Roentgen 1, 20136 Milan (Italy).Email: michela.meregaglia@unibocconi.it. Tel: +39 (0) 2 5836 5935.

**Abstract**

The aim of this study was to review and critically assess the health economics literature on post-treatment follow-up for adult cancer survivors. A systematic search was performed using PubMed, EMBASE and the Cochrane Library. The CHEERS checklist was adopted to assess the quality of the included studies. Thirty-nine articles met the eligibility criteria. Around two thirds of the studies addressed the most common cancers (i.e. breast, colorectal, cervical and lung). Twenty-one were based on a single clinical study, while the rest were modeling papers. All types of economic evaluations were represented other than cost-benefit analysis. The overall quality was generally high with an average proportion of 74% of checklist criteria fulfilled. The cost-effectiveness results supported the current trend towards less intensive, primary care-based and risk-adapted follow-up schemes.

***Keywords:*** *cancer, follow-up, economic evaluation, cost-effectiveness, CHEERS, quality, review*

**Introduction**

In recent years, oncological services worldwide have experienced an increasing number of cancer survivors due to advances in diagnostic tools, curative treatments and prevention campaigns. After being treated for their primary cancer, patients usually enter a program of post-treatment follow-up which may last for several years [1]. Follow-up programs usually involve hospital-based consultations with specialist cancer physicians, but the frequency of visits and the healthcare settings and professionals involved may vary according to geographical contexts and tumor sites [2]. Routine surveillance is primarily aimed at detecting loco-regional recurrences of cancer, metastases or second primaries at the earliest opportunity in order potentially to administer salvage treatments [3-4]. Secondary aims include addressing treatment-related side effects, managing the rehabilitation process and providing psychological and social support to patients and caregivers [2,5].

A variety of recommendations have been provided at national and international level to guide clinicians in the follow-up process of cancer care; however, the most efficient scheme for monitoring patients after the end of primary treatment is still under debate [6]. Firstly, whether or not repeated investigations can improve long-term clinical outcomes in cancer survivors remains controversial in oncology [3]. Secondly, the clinical benefit of early detection of a cancer relapse strictly depends on the availability of secondary treatments able to extend survival. Thirdly, cancer patients who are monitored intensively after the end of primary treatment may experience either positive (reassurance, relief) or negative (discomfort, anxiety) feelings [3,7]. Cancer surveillance schemes extended in time also raise economic considerations. The opportunity cost of delivering post-treatment services is significantly high [3] and the long-term sustainability of these programs must be carefully evaluated.

The assessment of the effectiveness and the cost-effectiveness of follow-up services in oncology is not straightforward. Any health benefits may become evident long after the intervention and post-treatment programs often involve different medical specialties and consume a variety of health care resources. Limited cost-effectiveness data are available to discriminate among different surveillance strategies especially in least common malignancies. Moreover, the optimal way to assess and combine health outcomes and cost parameters arising from these interventions is still unclear. The aim of this study was to review systematically and critically the clinical and economic evidence on follow-up strategies for adult cancer patients treated with curative intent. Published economic evaluations are summarized to identify the recent common and conflicting issues around cancer follow-up. The quality of the studies is appraised to help clinicians and decision makers understand the methodological limitations of the existing literature.

**Methods**

***Study identification and selection***

A systematic literature review was undertaken searching three major electronic databases (i.e. PubMed, EMBASE, and the Cochrane Library). The Preferred Reporting System for Systematic Review and Meta-Analysis (PRISMA) strategies were used to ensure systematic selection of studies [8]. Keywords were defined according to PICOS (population, intervention/comparator, outcome, study design) elements (Table 1). Economic evaluations comparing (two or more) follow-up interventions for adult patients (i.e. ≥ 18 years) after curative treatment for any cancer were included; both health and economic outcomes (i.e. costs) had to be reported. Childhood malignancies were excluded due to different outcomes and costs trajectories and longer time horizon. Other reasons for exclusion were: follow-up strategies for premalignant lesions not yet treated; screening programs for high-risk populations; clinical studies not reporting cost data; cost analyses focusing on only one alternative.

All searches were for studies published in the period 2000-2014. The reference list of the relevant articles was searched to avoid missing other pertinent studies. Only original full-text articles (i.e. not conference abstracts or editorial comments) were selected; no language restriction was applied to the search. The retrieved studies were reviewed by a single researcher (MM) in close consultation with a senior author (JC) and, in case of disagreement, issues were resolved by consensus. A data extraction template was designed to include all relevant information from the studies identified including country, setting, patient population, number of patients, intervention and comparator, type of economic analysis (e.g. modeling vs. clinical study-based), health and economic outcomes, time horizon, cost perspective, currency, conversion and discounting, uncertainty analysis, data sources, study results, and conclusions.

***Study quality assessment***

The CHEERS (Consolidated Health Economic Evaluation Reporting Standards) Statement was adopted for the critical appraisal of the studies. The CHEERS checklist was recently developed by the International Society of Pharmacoeconomics and Outcomes Research (ISPOR) and jointly endorsed by BMJ and nine other journals. Previous health economic evaluation guidelines were aggregated into a single standard to help authors report studies or reviewers assess them for publications. CHEERS consists of a 24-item checklist composed by five broad categories: Title and abstract (2 items); Introduction (1 item); Methods (14 items); Results (4 items); Discussion (3 items) [9]. The 24-item checklist was completed for each study included in the review, indicating “yes” when the criteria were met, “no” when they were unfulfilled and “not applicable” when they were not required for that type of study. Although the CHEERS checklist is not a scoring instrument, papers were divided into three quality categories according to the proportion of items achieved: high (≥75%), average (50%-75%) and poor (<50%) on the basis of other review studies adopting the same tool [10-13].

**Results**

Figure 1, a PRISMA diagram, displays the data for the number of titles initially identified (n=2490), 929 of them were duplicates. After title and/or abstract screening of the remaining 1561 records, 1507 publications were excluded for a variety of criteria (mainly studies on cancer treatment and cancer prevention/screening in high-risk populations). Fifty-four full-text articles were assessed for eligibility in the study, but only 39 finally met all inclusion criteria and were included in the review.

***Study characteristics***

A synthesis of the characteristics of the 39 included papers [14-52] is provided in Table 2. Of these, nine studies were conducted in the US, seven in the Netherlands, five in the UK, three each in Canada and Sweden, two each in Italy, France and Spain; the remaining studies, one per country, were carried out in Australia, Germany, Finland, Israel, Korea, and Norway.

The range of tumor sites was quite wide. Ten articles (26%) were related to breast cancer, followed by cervical (18%), colorectal (15%), lung (13%), bladder (8%), esophagus (5%), Hodgkin’s lymphoma (5%); finally, one each dealt with anal, pancreatic, prostate cancer and melanoma. The age range of patients was not systematically recorded, but all were adult subjects according to the review inclusion criteria.

Sixteen studies explicitly adopted a modeling framework (twelve Markov models; two discrete event simulations; one decision tree; one semi-Markov model) involving data extrapolation and/or evidence synthesis. Two papers [15,40] were based on unspecified modeling. Fourteen studies (36%) were analyses of empirical data from clinical trials. Among the remaining articles, three each were (non-randomized) prospective [14,27,33] and retrospective studies [19,28,44]; one paper [49] was classified as a retrospective study plus modeling. The number of participants recruited in non-modeling studies (n=20) ranged between 69 [42] and 472 [37], averaging at 231 patients.

In table 2, studies were clustered according to different types of follow-up programs compared. Around one third of the studies (12 out of 39) were classified as comparisons of follow-up strategies of different ‘intensity’; among them were included studies addressing follow-up schemes with different timing of controls (e.g. annual or 2-year mammography) or with novel diagnostic tests (e.g. PET-CT imaging) added to routine investigations. Six studies [16,26,33,42,44,51] examined outcomes and costs of a variety of diagnostic tools (e.g. positron-emission tomography vs. computed tomography) while four studies [14,23,38,41] related to cervical cancer compared Human Papilloma Virus (HPV) test versus cytology as potential instruments to detect new lesions. Four studies [25,30,34,49] compared costs and health gains arising from a surveillance program versus a ‘do nothing’ strategy; eleven studies examined traditional hospital-based follow-up programs in comparison with programs led by other healthcare professionals (i.e. in eight cases [20,31,35,36,43,46,48,52] the nurse and in three [17,19,28] the family physician). One study [15] compared a mobile-app follow-up versus traditional in-person consultations and another one [24] described two programs with (and without) an educational session.

Almost half of the papers (19 out of 39) adopted a limited healthcare perspective where only direct medical costs were considered; one study [43] from UK included the costs borne by social services as well. Seven US-based studies [16,29,33,34,41,49,50] carried out the analysis from the national social insurance program (i.e. Medicare) perspective. Two authors [14,44] calculated the costs borne by the health insurance companies in their countries (i.e. Israel and Korea). A societal perspective estimating broader costs to society (i.e. out-of-pocket costs, informal care, and productivity losses) irrespective of the payer was adopted by seven studies. The remaining three articles [20,24,42] presented study results according to both healthcare and societal perspectives.

Fifty-nine percent of the included papers (23/39) compared post-treatment surveillance outcomes and costs over a period between 1 and 5 years. Six studies [15,33,35,43,44,46] adopted a shorter timeframe (≤ 1 year), and three studies [21,22,38] a longer one (6-10 years). For six studies [16,25,29,30,41,50] the model was run over a lifetime horizon. In one case [48] the time horizon was not specified. The nine studies adopting a longer timeframe (i.e. >5 years) were all modelling studies; conversely, when primary data collection was performed (i.e. RCTs and cohort studies) the time horizon did not exceed 5 years.

Twenty-four articles did not report any discount rate for future costs and health outcomes. Five papers [18,20,23,38,40] applied a discount rate to costs only (i.e. either 3.0% or 3.5%). In nine articles costs and effects were discounted at the same rate (i.e. 3.0% in eight cases, 5.0% in one). The last paper [51] adopted a 1.5% discount rate for outcomes and 4.0% for costs. Among the studies (32 out of 39) adopting a time horizon ≥ 1 year, discounting was not applied in more than half of the cases (17/32), while five studies applied a discount factor to costs only; the remaining ten articles discounted both costs and health outcomes at rates ranging between 3% and 5%.

Eleven studies were cost-effectiveness analyses where outcomes were expressed in natural units, either intermediate (e.g. number of recurrences detected) or final ones (e.g. survival). Twelve articles assessed the effects of an intervention in terms of QALYs, either alone or in combination with physical outcomes (e.g. survival). Five studies were cost-minimization analyses assuming the equivalence of outcomes between interventions and comparators; eleven were cost-consequences analyses where costs and effects were not combined in a unique index and economic results were usually expressed in terms of cost per patient or total costs.

***Study findings***

Among the four studies [25,30,34,49] comparing a follow-up program for cancer survivors versus not performing one (i.e. ‘do nothing’ option), all showed the intervention strategy was a cost-effective option through the calculation of incremental cost per QALY (or life year) gained or recurrence detected.

Six studies compared several diagnostic tools to be adopted in post-treatment setting and the results in terms of costs and health outcomes varied considerably. One study [51] evaluating positron-emission tomography combined with computed tomography (PET-CT) showed that PET-CT was a cost-effective diagnostic instrument compared to CT alone or X-ray in lung cancer patients. Another study dealing with lung cancer [42] and comparing coincidence detection system imaging (CDET) with 18-fluorodeoxyglucose (FDG) with conventional technique imaging concluded that the two groups were similar in term of recurrences detected and survival, and FDG-18 was more costly. Three authors [26,33,44] reported that selected innovative tools (i.e. semi-automated microsatellite-analysis and NMP22 bladder check) were not cost-effective in detecting bladder cancer recurrences. One paper [16] comparing several diagnostic strategies for HIV-infected men with anal cancer found that a combination of high-resolution anoscopy (HRA) and cytology provided the greater benefit at an acceptable cost/QALY gained.

Four papers aimed at women with cervical intraepithelial neoplasia (NIP) compared follow-up strategies involving diagnostic tools such as HPV testing, cytology and colposcopy. Three of these studies [14,23,38] concluded that HPV testing was cost-effective compared to a conventional cytological approach; on the contrary, one study [41] concluded that HPV testing added limited improvement to survival at higher costs than cytology.

A broad group of papers (12/39) was categorized as economic evaluations of follow-up programs of different ‘intensity’. In five cases [18,22,27, 32, 45,], one or more diagnostic tests were added to routine surveillance (e.g. PET-CT imaging plus standard practice). All these papers concluded that a less-intensive follow-up program was clinically and economically justified for a variety of malignancies (i.e. breast, cervical and colorectal cancer, and melanoma). Three studies [21,37,40] compared the same types of diagnostic exams but administered with different timing, either in terms of number of tests per year or of follow-up length (or both). Among them, two studies stated that less frequent options were cost-effective [21] or even cost saving [37] compared to a more intensive program. On the contrary, the study by Macafee [40] showed that an intensive strategy detected more recurrences but at additional costs in colorectal cancer follow-up; as results were expressed in terms of cost per additional resectable recurrence, the cost-effectiveness of the intervention program compared to standard guidelines was not stated in this study. In two articles [29,50] the definition of ‘intensity’ combined both concepts (i.e. increased frequency and additional tests); in both these cases, less intensive options were preferred since they provided comparable clinical outcomes (i.e. overall survival, number of recurrences detected) at significantly lower costs. One study [47] showed that a risk-adapted follow-up, with the timing of clinical controls and radiological investigations modulated according to the risk of recurrence, was cost saving compared with a common strategy for all. Lastly, the study by Lu [39] performed a comparison of three follow-up strategies where the time in hospital was progressively shortened by a shift of care to the GP; once again, the simplified follow-up showed an acceptable cost-effectiveness profile.

The results from twelve studies dealing with different organizational aspects of post-treatment surveillance supported the current trend of moving towards less structured healthcare programs. Seven [31,35,36,43,46,48,52] out of eight papers comparing hospital-based versus nurse-led follow-up revealed the latter option was less costly without compromising patients health or acceptability. A UK-based cost-minimization analysis [20] concluded that, given the equivalence of health outcomes, a nurse-led telephone follow-up compared to a traditional hospital-based one might reduce patient’s travel and productivity costs but did not lead to cost or salary savings in the National Health Service perspective. In a similar way, three studies [17,19,28] showed that a general practitioner-led follow-up did not affect survival, quality of life or time to detection of recurrence. The last study [15] concluded that a mobile-app-based follow-up was cost-effective compared to a traditional in-person approach.

Finally, the study by Coyle [24] assessed a survivorship care plan (SCP) for women after breast cancer treatment, including an educational session for patients and full follow-up guidelines for general practitioners. SCP was not cost-effective since the control group had better outcomes and lower costs than the SCP group.

***Study quality assessment***

The quality evaluation of the included studies based on the CHEERS checklist is summarized in Table 3. Three items (i.e. preference measurement, model choice and model assumption) were applicable only to a limited number of studies; in detail, item 12 was related to cost-utility analyses using QALYs, while items 15-16 concerned modeling studies only. For this reason, the quality judgment was expressed in terms of percentage (instead of absolute number) of checklist criteria met.

Nineteen papers were categorized as high quality studies as more than 75% of criteria were fulfilled. For seventeen studies the quality estimated was of average level (between 50% and 75% of items met), while the remaining three studies met less than 50% of the checklist criteria and were categorized as poor. The average proportion of items achieved was 74% (range: 33% - 100%).

The most commonly missing quality criteria were: not accounting for patients’ heterogeneity in reporting results (Item 21; missing: 72%); a non-explicative title of the interventions compared and the economic evaluation performed (Item 1; missing: 59%); outcomes and costs not discounted and/or justification not given for the adopted (or not adopted) discount rate (Item 9; missing: 59%); incremental analysis between alternatives not performed and/or incremental cost-effectiveness ratio not reported (Item 19; missing: 44%).

On the other hand, the CHEERS items most often reported in the studies were the specification of the target population (Item 4; fulfilled: 97%) and the time horizon of the study (Item 8; fulfilled: 97%), followed by description of health outcomes and the methods and/or sources adopted to measure them (Items 10-11; fulfilled: 95%).

Despite the estimated quality being rather high (74%) across studies, a number of CHEERS points have been weakly considered by selected authors and need further discussion. First, discounting (Item 9) was disregarded by 59% of the studies; in sixteen cases, time horizon was longer than 2 years, thus a discount factor was required. Moreover, even in studies adopting a short timeframe (i.e. ≤1 year) CHEERS recommends reporting a 0% rate for clarity. Among the few studies which carried out some discount technique, most explicitly referred to published guidelines or health jurisdictions. All UK papers in this group adopted the most recent recommendations from NICE for discounting (i.e. 3.5%), but applied the rate to costs only.

A second issue is the specification of model choice and assumptions (Items 15-16). In fact, only three studies [18,21,22] gave reasons for the specific type of model used. Five modeling studies did not provide a graphical representation of the model structure and four did not properly describe the assumptions underlying the decision-analytic model. Moreover, a table reporting cost (and utility, if required) parameters with probability distributions (Item 18) was not provided by fifteen studies.

Thirdly, with reference to analytic methods (Item 17) a positive judgment (‘yes’) has been given if at least one of the following was reported in the article: methods for dealing with skewed, missing or censored data; approaches to validate or adjust a model (e.g. half-cycle correction); methods for handling heterogeneity or uncertainty. Uncertainty was considered by the great majority (87%) of the studies, either through sensitivity analyses in modeling studies and statistical tests (e.g. t-test, chi-square) or 95% confidence intervals in RCTs; conversely, the other two methodological requirements were fulfilled by a limited number of papers. Data skeweness was taken into account by six studies [17,20,35,43,46,52] which adopted non-parametric bootstrapping techniques (e.g. Mann-Whitney test) for skewed cost data. One study [24] used standard methods for handling censored data; no study specified how to deal with missing data. Among the modeling studies, only one [25] reported adjustments for lead-time bias.

Fourthly, the checklist section about costs (Items 13-14) was only partially fulfilled. In two papers adopting a societal perspective [23,26] the source of travel, production and other patient costs was not reported. Among the papers which met this requirement, four studies [17,20,24,35] empirically collected non-medical direct costs (i.e. transportation, co-payments, other patient/family expenses) through ad hoc surveys, cost diaries and local sources; in two studies [15,42] transportation costs were calculated as a function of the distance between home and hospital. Productivity losses were estimated on the basis of average national wages [15,20,30].

With reference to Item 14 (currency, price date and conversion), eight studies [19,26,27,39,44,45,47,48] did not specify the year of reported costs, while two authors [27,32] did not report the exchange rate between Euros and US dollars. In the nine US-based studies that met this requirement, medical costs (in US dollars) were derived from Medicare reimbursement data and adjusted using the medical component of the consumer price index (PCI) for the study year. In the only study from Norway [17] cost elements were converted from Norwegian kroner (NOK) into British pounds at the study year exchange rate. The remaining three studies from Australia, Israel and Korea reported healthcare costs in Australian dollars, US dollars and Korean Won, respectively.

Moreover, thirteen studies only reported results as incremental cost-effectiveness ratios (ICER), either as cost/QALY, cost/LYG or cost/additional treatable recurrence. Eight more studies reported mean differences in effects and costs between (or among) the alternative interventions, without combining them in an ICER. The remaining eighteen studies simply indicated mean values for the main categories of estimated costs and health outcomes for each follow-up strategy analyzed; this approach is typical of cost-consequences analyses, where costs and consequences are not aggregated into a single measure.

Finally, intervention results in cancer care may vary according to patients’ characteristics and disease severity; however, a limited group of papers (11 out of 39) handled population heterogeneity in the decision model. In these studies, results were stratified according to a number of factors (i.e. age, treatment, cancer stage, smoking status, presence/absence of metastasis, high/low recurrences risk, and symptomatic/asymptomatic disease).

**Discussion**

***Summary of evidence***

To our knowledge, this is the first time health and economic outcomes from post-treatment follow-up interventions have been compared across countries and for all types of cancer. We found around twenty-five reviews published in the last five years dealing with cancer follow-up care, but they were all limited to clinical studies only, specific cancer populations or types of interventions (e.g. primary versus secondary care). A dated paper by Edelman [53] reviewed surveillance strategies and assessed follow-up costs for the most common malignancies based on the studies retrieved. A study by Hex [54] aimed at reviewing the cost-effectiveness of follow-up care in pediatric tumors and highlighted a trend towards risk-based personalized approaches for long-term childhood cancer survivors. The target population addressed by Hex was exactly complementary to that of the current study, as our search was focused on adult cancer patients only. Moreover, none of these reviews addressed the quality of the included studies.

Our review provides insights into the clinical and economic value of a variety of post-treatment follow-up programs across many types of malignancies. Due to recent improvements in cancer therapies and survival rates, the number of patients requiring post-treatment services is rapidly increasing and posing a substantial burden on health care systems. The thirty-nine studies included in the review represent the best economic evidence available around cancer follow-up. From study findings, a general tendency emerged towards less intensive options in terms of frequency of visits and/or length of program, risk-adapted follow-up according to age or tumor stage, and service delivery in primary care or through mobile-app technologies replacing traditional hospital-based investigations. In most studies, these simplified follow-up schemes were to be preferred to the more intensive ones according to their favorable cost-effectiveness profile. However, some cost-effectiveness estimates for the same type of surveillance scheme were contradictory and estimates varied considerably by study setting and cancer type. Moreover, most studies reported the equivalence (or a non-significant difference) of health outcomes between traditional and novel options for post-treatment surveillance, with a cost saving when less intensive or non-hospital-based programs were implemented; indeed, none of them led to significant improvements in health outcomes such as the number of recurrences detected, overall survival or patient’s satisfaction.

The average quality score (74%) of the studies retrieved was good, with the majority of them (49%) performing very well in reporting economic evaluations. Among high-quality studies, we identified mainly cost-utility analyses, Markov model- and UK/US-based studies; cost-utility analyses, indeed, generally measure outcomes in terms of QALYs and cost/QALY is the ratio adopted by more recent studies to determine the cost-effectiveness of healthcare programs. Markov model is an appropriate instrument to conduct economic evaluations in chronic diseases where the occurrence of the events (e.g. cancer recurrence) is uncertain and these may happen more than once [55]. Compared to other types of modeling (e.g. discrete event simulation) equally valid for the purpose, the Markov model requires less clinical information but a validation of the underlying assumptions. Studies conducted in UK and US are more likely to adhere to recommendations from NICE or other HTA agencies. Finally, the average quality score (i.e. 80%) of the studies published in the last 5 years (n=18) is higher than the mean score (i.e. 61%) across the articles dated between 2000 and 2005. Thus, it is reasonable to assume that more recent economic evaluations tend to better adhere to published recommendations in this field.

***Critical issues***

A number of issues characterizing long-term cancer survivorship should be carefully evaluated by the health economics literature in this field. First, the setting (e.g. GP- versus hospital-based) where follow-up care is conducted can greatly affect patient’s quality of life and private costs (i.e. travel expenses and productivity losses), especially in rural areas with long distances to travel to hospitals. Secondly, little is known about potential damages (and related costs) of follow-up, in terms of patient’s dissatisfaction, long-term toxicity, and false positive results; thus, economic evaluations including also these cost categories are encouraged. Thirdly, the topic of heterogeneity in cancer patients affected by the same malignancy is still unexplored in the literature. Post-treatment surveillance programs, indeed, may yield different survival gains according to age, cancer stage and comorbidities; thus, economic evaluations should routinely report differences in cost-effectiveness results for relevant subgroups of patients. A further weakness observed in the reviewed articles is related to the choice of health outcomes; most studies, indeed, evaluate follow-up interventions in terms of number of recurrences detected. However, the economic value of an early diagnosis of cancer relapse is closely linked to the availability and effectiveness of secondary treatments able to extend survival. Future economic studies are encouraged to adopt longer timeframes in order to catch the full health effects and cost paths arising from different surveillance options and potential curative treatments administered in case of recurrence. A longer time horizon is also necessary in order to capture long-term side effects related to intensive radiological examinations. As medical advances have improved post-treatment prognosis, the health issues experienced by cancer survivors tend to be more episodic and to occur over a longer timeframe. Thus, model-based economic evaluations are increasingly required to extend clinical trial results over patient’s lifetime.

***Strengths and limitations***

This systematic review presents a number of limitations. First, the databases we searched were limited to the most common databases (i.e. PubMed, Medline, and EMBASE). This may result in selection bias we tried to avoid by searching the references of each selected article manually. Secondly, as in all searches of the published literature there may be publication bias and unpublished studies could affect the review findings; however, this is particularly common in studies funded by private companies of which there are few in our review as it deals with non-pharmaceutical interventions. Thirdly, all the included studies were conducted in the developed world thus likely representing the most healthy and affluent group of cancer survivors [2]. For these reasons, the extension of our results to other settings needs to be done cautiously.

Moreover, comparison of economic outcomes is complicated due to the high variability in time periods, currencies, and health systems involved. Although it is reasonable to assume that the relative price of some common diagnostic technologies (e.g. PET-CT scan) might be similar across many developed countries [10], other cost categories (e.g. nurse salaries, consultation fees) or reimbursement policies may differ a lot. Even in studies assessing standard health outcomes such as survival or QALYs, comparison is hard to perform due to eleven different types of cancer addressed in the review and, even across studies related to the same malignancy, heterogeneous follow-up interventions and patient populations by age or cancer stage. Moreover, the concentration on economic evaluations as inclusion criterion may have excluded other important clinical studies in cancer follow-up research.

A further limitation is the use of the CHEERS checklist as a measure of quality in economic evaluations. This checklist gives an indication of how much the published studies adhere to reporting criteria but does not state their relative importance. For example, reporting discounting may be more relevant than funding source and a simple addition of the criteria met may result in a misleading assessment of quality. However, as the CHEERS checklist does not provide any weights to be applied to quality criteria, summing the number of items achieved with a qualitative discussion in text was assumed an appropriate methodology to differentiate the quality of studies.

In general, the use of checklists to evaluate the quality of economic evaluation studies can be viewed as overly simplistic. For example, the standard checklists emphasize how well the study is reported rather than whether or not it can inform good policy decisions, since they evaluate the publication itself and not its implications for clinical practice. Moreover, not reporting an element in the article does not necessarily mean that the authors in the analysis have not addressed that aspect. In addition, most checklist criteria require the reasons behind the choice of a given item (e.g. time horizon) are specified, but the authors often disregard this aspect and just indicate the parameter value; thus, the simple addition of the items reported by each study does not inform about the appropriateness of a method. Especially in model-based economic evaluations, different assumptions around the study parameters may significantly alter the cost-effectiveness results with important consequences for evidence-based medical and policy decisions. A well-reported study will not necessarily be fit for purpose but at least it may be easier to determine whether or not it is fit for purpose compared with a less well-reported study.

In spite of these limitations and compared to previous guidelines (e.g. Drummond [56]), the CHEERS checklist appears more comprehensive and suitable for model-based studies which are becoming increasingly important, partly due to the financial and logistic constraints on performing primary data collections (e.g. RCTs) but also because of the well-known limitations of trial-based evaluations for informing decision making. The checklist, indeed, asks authors to specify more details about model assumptions and analytic methods (i.e. skewed, missing, censored, and pooled data) that were not provided in older guidelines. Moreover, a greater emphasis is given to the need of characterizing heterogeneity in reporting study results, in line with the recent trend towards a more personalized medicine according to patient’s characteristics.

**Conclusions**

Health economic analyses are increasingly used to inform policy makers about the efficient allocation of limited healthcare resources. Economic evaluations in cancer care have been mainly applied to drug therapies, while less evidence is available for other types of interventions. This review summarizes the current body of knowledge regarding economic evaluations of follow-up services for cancer survivors. The quality of the studies retrieved from the literature is generally high. Although judging the quality of scientific work is inevitably controversial, the CHEERS checklist appears suitably up-to-date and comprehensive in order to facilitate this task. Our review indicates that less intensive, primary-care based follow-up is often clinically equivalent and economically justified in oncology. HPV testing appears a cost-effective alternative to traditional cytological approach for cervical cancer patients. There is also evidence of increasing interest in delivering post-treatment services using technology (e.g. mobile apps), although not yet sufficient to state whether or not remote cancer follow-up may turn out to be cost-effective. Our study may help oncologists and policy-makers interpret health economic results in cancer follow-up setting according to study quality and update post-treatment surveillance schemes based on a sound scientific evidence.

**Expert commentary**

In recent years, a debate arose among oncologists around how to design post-treatment programs for cancer survivors. Due to improvement of clinical outcomes after treatment for primary cancer, the number of patients attending follow-up clinics continues to grow challenging the economic sustainability of the most intensive surveillance schemes. It was therefore important to summarize the cost-effectiveness results from the published studies around post-treatment surveillance in cancer care and assess their quality using a well-validated tool as the CHEERS checklist.

**Five-year view**

A consensus has not been reached yet in the oncology community regarding whether an intensive, prolonged follow-up program improves survival and quality of life in cancer patients. Moreover, personalized healthcare programs are increasingly suggested by clinicians to rise patient’s satisfaction and adherence. In an era of paucity of economic resources there is also growing interest in economic as well as clinical outcomes of healthcare interventions. Economic evaluations in oncology have traditionally focused on drug treatments and have often featured common malignancies, such as breast and colorectal cancer. Thus, well-conducted health economic analyses of post-treatment programs for all types of cancer are increasingly required. Patient preferences for different aspects of healthcare programs have been also addressed in recent health economic research. At the same time, evaluating the quality of the existing studies is useful to interpret the available evidence.

**Disclosure**

The authors declare they have no competing interests with the content of the manuscript. No writing assistance was utilized in the production of this manuscript.

**Key issues**

* Dissenting opinions exist among oncologists around the effectiveness of traditional intensive, hospital-based follow-up programs for cancer survivors.
* Surveillance schemes often lack a sound scientific base and impose a significant economic burden to healthcare systems and societies.
* A systematic literature review of published studies between 2000 and 2014 was performed searching PubMed, EMBASE and Cochran Library databases.
* A total of thirty-nine studies met all the inclusion criteria and were reviewed.
* Most of the included studies dealt with breast, colorectal, cervical and lung cancer and were conducted in the US.
* A quality appraisal of the literature was performed through a well-validated instrument (CHEERS checklist) and summing the number of items addressed by each study.
* The quality of the studies was generally high averaging at 74% of checklist criteria fulfilled.
* A common trend favoring less intensive, primary-care based and personalized follow-up programs emerged from the study findings.
* Well-conducted health economic analyses in cancer management phases other than treatment and for least common malignancies are increasingly required.

**References**

Papers of special note have been highlighted as:

\* of interest

\*\* of considerable interest

1. Lewis RA, Neal RD, Williams NH et al. Follow-up of cancer in primary care versus secondary care: systematic review. Br J Gen Pract, e234-e247 (2009).

2. Dickinson R, Hall S, Sinclair JE et al. Using technology to deliver cancer follow-up: a systematic review. BMC Cancer 311 (14), 1-16 (2014).

3. Jeffery M, Hickey BE, Hider PN. Follow-up strategies for patients treated for non-metastatic colorectal cancer (Review). The Cochrane Library 1, 1-37 (2007).

4. Pita-Fernandez S, Alhayek-Ai M, Gonzalez-Martin C, et al. Intensive follow-up strategies improve outcomes in nonmetastatic colorectal cancer patients after curative surgery: a systematic review and meta-analysis. Annals of Oncology 00, 1-12 (2014).

5. Digonnet A, Hamoir M, Andry G et al. Post-therapeutic surveillance strategies in head and neck squamous cell carcinoma. Eur Arch Otorhinolaryngol. 270, 1569-1580 (2013).

6. Nama V, Nordin A, Bryant A. Patient-reported outcome measures for follow-up after gynecological cancer treatment (Review). The Cochrane Library 11, 1-15 (2013).

7. Robertson C, Arcot Ragupathy SK, Boachie C et al. The clinical effectiveness and cost-effectiveness of different surveillance mammography regimens after the treatment for primary breast cancer: systematic reviews, registry database analyses and economic evaluation. Health Technology Assessment 34 (15), 1-6 (2011).

8 \*\*. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 4 (1), 1-9 (2015).

**\*\* The original study describing the development of PRISMA-P 2015 checklist**

9 \*\*. Husereau D, Drummond M, Petrou S, et al. Consolidated health economic evaluation reporting standards (CHEERS) – explanation and elaboration: a report of the ISPOR health economic evaluation publication guidelines good reporting practices task force. Value in Health 16, 231-250 (2013).

**\*\* The original study describing the development and explanation of CHEERS checklist**

10. Gordon LG, Rowell D. Health system costs of skin cancer and cost-effectiveness of skin cancer prevention and screening: a systematic review. European Journal of Cancer Prevention 24, 141-149 (2015).

11. Maru S, Byrnes J, Carrington MJ, et al. Systematic review of trial-based analyses reporting the economic impact of heart failure management programs compared with usual care. European Journal of Cardiovascular Nursing, 1-9 (2014).

12. Mihalopoulos C, Chatterton ML. Economic evaluation of interventions designed to prevent mental disorders: a systematic review. Early Intervention in Psychiatry, 1-8 (2014).

13\*.Palfreyman SJ, Stone PW. A systematic review of economic evaluations assessing interventions aimed at preventing or treating pressure ulcers. International Journal of Nursing Studies 52, 769-788 (2015).

**\* A study applying the CHEERS checklist to assess the quality of economic analyses of prevention and treatment interventions for pressure ulcers.**

14. Almog B, Gamzu R, Bornstein J, et al. Clinical and economic benefit of HPV-load testing in follow-up and management of women postcone biopsy for CIN2-3. British Journal of Cancer 89, 109-112 (2003).

15. Armstrong KA, Semple JL, Coyte PC. Replacing ambulatory surgical follow-up visits with mobile app home monitoring: modeling cost-effective scenarios. J Med Internet Res. 16(9), 1-34 (2014).

16. Assoumou SA, Mayer KH, Panther L, et al. Cost-effectiveness of surveillance strategies after treatment for high-grade anal dysplasia in high risk patients. Sex Transm Dis. 40(4), 298-303 (2013).

17. Augestad KM, Norum J, Dehof S, et al. Cost-effectiveness and quality of life in surgeon versus general practitioner-organized colon cancer surveillance: a randomized controlled trial. BMJ Open 3, 1-15 (2013).

18. Auguste P, Barton P, Meads C, et al. Evaluating PET-CT in routine surveillance and follow-up after treatment for cervical cancer: a cost-effectiveness analysis. BJOG 121, 464-476 (2014).

19. Baena-Canada JM, Ramirez-Daffos P, Cortés-Carmona C, et al. Follow-up of long-term survivors for breast cancer in primary care versus specialist attention. Family Practice 30, 525-532 (2013).

20. Beaver K, Hollingworth W, McDonald R, et al. Economic evaluation of a randomized clinical trial of hospital versus telephone follow-up after treatment for breast cancer. British Journal of Surgery 96, 1406-1415 (2009).

21. Bessen T, Karnon J. A patient-level calibration framework for evaluating surveillance strategies: a case study of mammographic follow-up after early breast cancer. Value in Health 17, 669-678 (2014).

22. Borie F, Combescure C, Daures JP, et al. Cost-effectiveness of two follow-up strategies for curative resection of colorectal cancer: comparative study using a Markov model. Word J. Surg 28, 563-569 (2004).

23. Coupé VMH, Berkhof J, Verheijen RHM, et al. Cost-effectiveness of human papillomavirus testing after treatment for cervical intraepithelial neoplasia. BJOG 114, 416-424 (2007).

24. Coyle BD, Grunfeld E, Coyle K, et al. Cost-effectiveness of a survivorship care plan for breast cancer survivors. Journal of oncology practice 10(2), e86-e92 (2013).

25. Das P, Ng AK, Earle CC, et al. Computed tomography screening for lung cancer in Hodgkin’s lymphoma survivors: decision analysis and cost-effectiveness analysis. Annals of oncology 17, 785-793 (2006).

26. de Bekker-Grob EW, van der Aa MNM, Zwarthoff EC, et al. Non-muscle-invasive bladder cancer surveillance for which cystoscopy is partly replaced by microsatellite analysis of urine: a cost-effective alternative? BJU International 104, 41-47 (2008).

27. Forni F, Ferrandina G, Deodato F, et al. Squamous cell carcinoma antigen in follow-up of cervical cancer treated with radiotherapy: evaluation of cost-effectiveness. Int. J. Radiation Oncology Biol. Phys. 69(4), 1145-1149 (2007).

28. Gilbert S, Reid KR, Lam MY, et al. Who should follow-up lung cancer patients after operation? Ann Thorac Surg. 69, 1696-700 (2000).

29. Guadagnolo BA, Punglia RS, Kuntz KM, et al. Cost-effectiveness analysis of computerized tomography in the routine follow-up of patients after primary treatment for Hodgkin’s disease. J Clin Oncol 24, 4116-4122 (2006).

30. Hassan C, Pickhardt PJ, Zullo A, et al. Cost-effectiveness of early colonoscopy surveillance after cancer resection. Digestive and liver disease 41, 881-885 (2009).

31. Helgesen F, Andersson S-O, Gustafsson O, et al. Follow-up of prostate cancer patients by on-demand contacts with specialist nurse. Scand J Urol Nephrol 34, 55-61 (2000).

32. Hengge UR, Wallerand A, Stutzki A, et al. Cost-effectiveness of reduced follow-up of in malignant melanoma. JDDG 5, 898-907 (2007).

33. Kamat AM, Karam JA, Grossman HB, et al. Prospective trial to identify optimal bladder cancer surveillance protocol: reducing costs while maximizing sensitivity. BJU International 108, 1119-1124 (2011).

34. Kent MS, Korn P, Port JL, et al. Cost-effectiveness of chest computed tomography after lung cancer resection: a decision analysis model. Ann Thorac Surg 80, 1215-23 (2005).

35. Kimman ML, Dirksen CD, Voogd AC, et al. Economic evaluation of four follow-up strategies after curative treatment for breast cancer: results of an RCT. European Journal of Cancer 47, 1175-1185 (2011).

36. Koinberg I, Engholm GB, Genell A, et al. A health economic evaluation of follow-up after breast cancer surgery: results of an RCT study. Acta Oncologica 48, 99-104 (2009).

37. Kokko R, Hakama M, Holli K. Follow-up cost of breast cancer patients with localized disease after primary treatment: a randomized trial. Breast Cancer Research and Treatment 93: 255-260 (2005).

38. Legood R, Smith M, Lew JB, et al. Cost-effectiveness of human papillomavirus test of cure after treatment for cervical intraepithelial neoplasia in England: economic analysis from NHS Sentinel Sites Study. BMJ 345, 1-10 (2012).

39. Lu W, Greuter MJW, Schaapveld M, et al. Safety and cost-effectiveness of shortening hospital follow-up after breast cancer treatment. British Journal of Cancer 99, 1227-1233 (2012).

40. Macafee DAL, Whynes DK, Scholefield JH. Risk-stratified intensive follow up for treated colorectal cancer - realistic and cost saving? Colorectal Disease 10, 222-230 (2007).

41\*. Melnikow J, Kulasingam S, Slee C, et al. Surveillance after treatment for cervical intraepithelial neoplasia. Obstetrics and Gynecology 116(5), 1158-1170 (2010).

**\* A high quality study evaluating costs and outcomes of surveillance strategies in women treated for cervical cancer using Markov model.**

42. Monteil J, Vergnenègre A, Bertin F, et al. Randomized follow-up study of resected NSCLC patients: conventional versus F-DG coincidence imaging. Anticancer Research 30, 3811-3816 (2010).

43. Moore S, Corner J, Haviland J, et al. Nurse led follow up and conventional medical follow-up in management of patients with lung cancer: randomized trial. BMJ 325, 1-7 (2002).

44. Ok BG, Ji YS, Ko YH, et al. Usefulness of urine cytology as a routine work-up in the detection of recurrence in patients with prior non-muscle-invasive bladder cancer: practicality and cost-effectiveness. Korean J Urol 55, 650-655 (2014).

45. Oltra A, Santaballa A, Munarriz B, et al. Cost-benefit analysis of a follow-up program in patients with breast cancer: a randomized prospective study. The Breast Journal 13(6), 571-574 (2007).

46. Polinder S, Verschuur EML, Siersema PD, et al. Cost comparison study of two different follow-up protocols after surgery for esophageal cancer. European Journal of Cancer 45, 2110-2115 (2009).

47. Secco GB, Fardelli R, Gianquinto D, et al. Efficacy and cost of risk-adapted follow-up in patients after colorectal cancer surgery: a prospective, randomized and controlled trial. EJSO 28, 418-423 (2002).

48. Strand E, Nygren I, Bergkvist L, et al. Nurse or surgeon follow-up after rectal cancer: a randomized trial. Colorectal Disease 13, 999-1003 (2011).

49. Tergas AI, Havrilesky LJ, Fader AN, et al. Cost analysis of colposcopy for abnormal cytology in post-treatment surveillance for cervical cancer. Gynecologic Oncology 130, 421-425 (2013).

50. Tzeng C-W D, Abbott DE, Cantor SB, et al. Frequency and intensity of postoperative surveillance after curative treatment of pancreatic cancer: a cost-effectiveness analysis. Ann Surg Oncol 20: 2197-2203 (2013).

51. van Loon J, Grutters JPC, Wanders R, et al. FDG-PET-CT in the follow-up of non-small cell lung cancer patients after radical radiotherapy with or without chemotherapy: an economic evaluation. European Journal of Cancer 46: 110-119 (2010).

52. Verschuur EML, Steyerberg EW, Tilanus HW, et al. Nurse-led follow-up of patients after esophageal or gastric cardia cancer surgery: a randomized trial. British Journal of Cancer 100, 70-76 (2009).

53. Edelman MJ, Meyers FJ, Siegel D. The utility of follow-up testing after curative cancer therapy. A critical review and economic analysis. J Gen Intern Med 12(5): 318-31 (1997).

54\*.Hex N, Bartlett C. Cost-effectiveness of follow-up services for childhood cancer survivors outside the cancer setting. Current Opin Support Palliat Care 7: 314-317 (2013).

**\* A systematic review of economic evaluation studies of follow-up strategies for pediatric cancer survivors.**

55\*.Yang M, Patel DS, Tufail W, et al. The quality of economic studies of cancer pharmacogenomics: a quantitative appraisal of the evidence. Expert Rev. Pharmacoecon. Outcomes Res. 13(5), 597-611 (2013).

**\* A well-conducted review and critical appraisal of economic studies in cancer pharmacogenomics.**

56. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. BMJ 313, 275-83 (1996).

**Figures and tables.**

**Table 1.** Search keywords according to PICOS elements.

|  |  |  |
| --- | --- | --- |
| **PICOS** | **Inclusion Criteria** | **Keywords** |
| Population | Adults patients after any cancer treatment | cancer OR carcinoma OR tumor OR neoplasm OR neoplasia |
| Intervention/Comparator | Post-treatment follow-up strategies | ‘follow up’ OR surveillance |
| Outcomes | Any health outcomes; costs | *Not specified* |
| Study design | Economic evaluations | ‘economic evaluation’ OR ‘cost effectiveness’ OR ‘cost utility’ OR ‘cost benefit’ OR ‘cost minimization’ OR ‘cost consequences’ |

**Figure 1.** Flow diagram.

**2490** citations identified by searches:

PubMed: N= 785

EMBASE: N= 1235

The Cochrane Library: N= 470

**1561** papers screened by title/abstract

**54** papers retrieved for detailed inspection

**39** papers included in the review

**929** duplicates removed

**1507** papers excluded:

**563** drug or surgical treatments, **389** population screening, **102** diagnosis/disease staging, **99** no cancer, **68** no economic data, **45** health promotion, **41** psycho-oncology, **39** methodological studies, **38** reviews, **19** conference abstracts, **19** HPV vaccination, **18** epidemiological studies, **17** cost-of-illness studies, **15** clinical guidelines, **8** genetic counselling, **7** no full-text, **6** quality of life/health utility assessment, **4** editorial comments, **4** pediatric cancer, **3** trial protocols, **3** fertility preservation

**15** papers excluded:

**6** no well-defined comparator, **3** no comparator group, **2** follow-up of low-grade precursor lesions, **1** pre-treatment, **1** not relevant comparator (surgery), **1** quasi-duplicate, **1** meta-analysis

**Table 2.** Key features of economic evaluation studies included in the review (n=39)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **First author** **(Year)** | **Country** | **Intervention**  | **Comparator** | **Study design** | **Cost perspective** | **Time horizon** | **Health outcomes** | **Economic** **outcomes** | **Quality score** |
| **Cancer** | **Economic evaluation** | **Discount factor** |
| ***Comparisons of follow-up strategies of different intensity*** |
| Auguste (2013) | UK | PET-CT imaging plus current practice | Current practice alone | Markov model | Healthcare system | 5 years | QALYs gained: 4.1096 | ∆Cost/∆QALY: £1 million | 96% |
| Cervical | CUA | Costs: 3.5% |
| Bessen (2014) | Australia | (1) Current annual mammography FUP (2) Mixed FUP | Less intensive FUP (2-year mammography) | DES model | Healthcare system | 10 years | QALYs gained: 0.002-0.006 (50-69 y old); 0.000-0.003 (70-79 y old) | ∆Cost/∆QALY: AU$21,481-AU$133,525 (50-69 y old); AU$40,706-AU$413,230 (70-79 y old) | 83% |
| Breast | CUA | No |
| Borie (2004) | France | CEA-based standard FUP | Simplified FUP | Markov model | Healthcare system | 7 years | QALYs gained:0.25 | ∆Cost/∆QALY: €3,114 | 58% |
| Colorectal | CUA | No |
| Forni (2007) | Italy | Simplified FUP (SCC antigen plus gynecologic examination) | Complete FUP | Prospective cohort | Healthcare system | 5 years | Rate of missed recurrences: 2.2% | Cost/patient: €298.5 vs. €3,653.4 | 52% |
| Cervical | CCA | No |
| Guadagnolo (2006) | US | Annual CT for 5 or 10 years | FUP with non-CT modalities only | Markov model | Societal (modified) | Lifetime | QALY gained: 0.0005 | ∆Cost/∆QALY: $9,042,300 | 87% |
| Hodgkin's lymphoma | CEA; CUA | 3.0% |
| Hengge (2007) | Germany | Established FUP practice | Less intensive FUP | Markov model | Healthcare system | 5 years | No difference in survival | Cost/QALY: €63,252 vs. €42,433  | 50% |
| Melanoma | CEA; CUA | No |
| Kokko (2005) | Finland | Four strategies combining different visit timing and diagnostic tools | RCT | Healthcare system | 5 years | Recurrences detected: 28-35 (range); no difference in SDF and OS | Cost/recurrence detected: €4,166 - €9,149 (range) | 57% |
| Breast | CEA | No |
| Lu (2012) | Netherlands | 5-year FUP with annual mammography  | Three less intensive strategies | DES model | Healthcare system | 5 years | No difference in recurrences detected | Cost (x1000)/1% increase in recurrences detected: range: €62.1 – €83.1 (current strategy) | 65% |
| Breast | CEA | No |
| MacAfee (2007) | UK | Intensive FUP | Standard FUP | Model (NS) | Healthcare system | 5 years | Additional recurrences detected: 853 | ∆Cost/∆recurrence detected: £18,077 | 78% |
| Colorectal | CEA | Costs: 3.5% |

**Table 2 (cont.).** Key features of economic evaluation studies included in the review (n=39)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **First author** **(Year)** | **Country** | **Intervention**  | **Comparator** | **Study design** | **Cost perspective** | **Time horizon** | **Health outcomes** | **Economic** **outcomes** | **Quality score** |
| **Cancer** | **Economic evaluation** | **Discount factor** |
| ***Comparisons of follow-up strategies of different intensity (cont.)*** |
| Oltra (2007) | Spain | FUP with annual mammography (n=63) | A more intensive FUP (n=58) | RCT | Healthcare system | 3 years | Recurrences detected: 11 (17.5%; 95% CI: 9.6%-25.3%) vs. 13 (22.4%; 95% CI: 13.4%-31.4%) | Cost/patient: €390 vs. €1,278. Total cost: €24,567 vs. €74,171 | 33% |
| Breast | CCA | No |
| Secco (2002) | Italy | Risk-adapted FUP (n=192) | Minimal surveillance (n=145) | RCT | Healthcare system | 5 years | Risk of recurrence: 52.6% vs. 57.2% (*p*<0.05) | No difference in cost | 48% |
| Colorectal | CCA | No |
| Tzeng (2013) | US | Four strategies of increasing intensity | No scheduled FUP | Markov model | Medicare | Lifetime | OS (months): 24.6 (no surveillance) vs. 32.8 (surveillance) | ∆Cost/LYG: US$5364 - US$294,696 (range) | 96% |
| Pancreatic | CEA | 3.0% |
| ***Comparisons among different diagnostic tools*** |  |
| Assoumou (2013) | US | Five strategies using high resolution anoscopy (HRA) and/or cytology | Markov model | Medicare | Lifetime | QALYs gained: 0.0723-0.1061 (range) | Cost/QALY: US$4,446-US$17,373 (range) | 96% |
| Anal | CUA | 3.0% |
| de Bekker-Grob (2008) | Netherlands | Semi-automated MA plus cystoscopy | Cystoscopy alone | Semi-Markov model | Societal | 2 years | Probability (no recurrence after 2 years): 86.3% vs. 86.6% | Cost/patient: €4,104 vs. €3,433 | 78% |
| Bladder | CCA | No |
| Kamat (2011) | US | Five strategies combining cystoscopy, cytology, NMP22 and FISH | Prospective cohort | Medicare | ≅4 months | Detection rate: 52%-72% (range) | Cost/recurrence detected: US$7,692 - US$26,462 (range) | 67% |
| Bladder | CEA | No |
| Monteil (2010) | France | CDET imaging with 18-FDG (n=36) | Conventional imaging (n=33) | RCT | Healthcare/Societal | 2 years | Recurrences detected: 16 (44.4%) vs. 9 (27.3%) (*p*=0.14). Time to recurrences detection (months): 12±9.9 vs. 18±11.8 | Cost/patient: €1,104.96 vs. €755.47 (*p*<0.001) | 67% |
| Lung | CCA | No |
| Ok (2014) | Korea | Seven strategies combining CT, cytology and urinalysis | Retrospective cohort | Health insurance | 6 months | Rate of recurrences detected: 24.5% - 77.6% (range) | Cost/recurrence detected: range: KRW11,049 - KRW100,647  | 67% |
| Bladder | CEA | No |
| Van Loon (2010) | Netherlands | (1) PET-CT scan; (2) chest-CT scan | Conventional chest X-ray scan | Markov model | Healthcare system | 5 years | QALYs: (1) 1.30 (2) 1.28; OS (months): (1) 25; (2) 24 | ∆Cost/∆QALY: (1) €69.086; (2) €264.033 | 92% |
| Lung | CUA | Costs: 4.0%;effects: 1.5% |

**Table 2 (cont.).** Key features of economic evaluation studies included in the review (n=39).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **First author** **(Year)** | **Country** | **Intervention**  | **Comparator** | **Study design** | **Cost perspective** | **Time horizon** | **Health outcomes** | **Economic** **outcomes** | **Quality score** |
| **Cancer** | **Economic evaluation** | **Discount factor** |
| ***HPV testing versus conventional cytology in cervical cancer follow-up*** |
| Almog (2003) | Israel | HPV testing (n=67) | Conventional cytology (n=63) | Prospective cohort  | Health insurance | ≅53 months | No difference in recurrences detected | Cost/recurrence detected: US$3,485 vs. US$3,573 | 57% |
| Cervical | CEA | No |
| Coupé (2007) | Netherlands | Six strategies with adjunct HPV testing | Current cytological FUP | Markov model | Societal | 5 years | Reduction in missed cases: 32%-77% (range) | Cost/patient: €178-€351 (range) | 70% |
| Cervical | CCA | Costs: 3.0% |
| Legood (2012) | UK | (1) Sentinel sites HPV test; (2) Extended HPV test  | Conventional cytology FUP | Markov model | Healthcare system | 10 years | Case averted: ≅8 | ∆Cost/∆Case averted: (1) -£1,120; (2) £6,474 | 87% |
| Cervical | CEA | Costs: 3.5% |
| Melnikow (2010) | US | Twelve strategies combining cytology, colposcopy and HPV testing | Markov model | Medicare | Lifetime | LYG: 0.001 vs. 0.108; QALY gained: 0.153 – 0.363 (range) | Cost/LYG: US$4,083 - US$1,160,000Cost/QALY: US$54 - US$5,246 (range) | 100% |
| Cervical | CEA; CUA | 3.0% |
| ***Comparisons between follow-up programs and ‘do nothing’ options*** |
| Das (2006) | US | Annual low-dose CT screening | No screening | Markov model | Societal (modified) | Lifetime | LYG: 0.64 and 0.16; QALYs gained: 0.58 and 0.14 (smokers and non-smokers) | ∆Cost/∆QALY: US$34,100 (smokers); US$125,400 (non-smokers) | 92% |
| Hodgkin's lymphoma | CEA; CUA | 3.0% |
| Hassan (2009) | US | 1-year endoscopy surveillance | No early endoscopy | Decision tree | Societal | Lifetime | LYG: 2,653 | ∆Cost/∆LYG: US$40,313 | 87% |
| Colorectal | CEA | No |
| Kent (2005) | US | Annual CT-based FUP | No annual CT-based FUP | Markov model | Medicare | 5 years | QALYs gained: 0.16 | ∆Cost/∆QALY: US$47,676 | 75% |
| Lung | CUA | 3.0% |
| Tergas (2013) | US | Colposcopy (n=27 low-grade Pap; n=60 high-grade Pap) | No colposcopy (n=23 low-grade Pap; n=18 high-grade Pap) | Retrospective cohort and model (NS) | Medicare | ≅34 months | Rate of recurrences detected: 8.3% vs. 0.0% | Cost/recurrence detected: US$7481 | 71% |
| Cervical | CEA | No |

**Table 2 (cont.).** Key features of economic evaluation studies included in the review (n=39).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **First author** **(Year)** | **Country** | **Intervention** | **Comparator** | **Study design** | **Cost perspective** | **Time horizon** | **Health outcomes** | **Economic** **outcomes** | **Quality score** |
| **Cancer** | **Economic evaluation** | **Discount factor** |
| ***Comparisons among different organizational aspects of post-treatment follow-up*** |
| Armstrong (2014) | Canada | Mobile-app FUP care  | Conventional, in-person FUP care | Model (NS)  | Societal | 30 days | None | Cost/patient: C$136 vs. C$381  | 61% |
| Breast | CMA | No |
| Augestad (2013) | Norway | GP-organized FUP (n=55) | Hospital surgeon-based FUP (n=55) | RCT | Societal | 2 years | No difference in QoL and time to recurrences detection | Cost/patient: £8,233 vs. £9,889 (*p*<0.001) | 95% |
| Colorectal | CMA | 3.0% |
| Baena-Canada (2013) | Spain | Primary care FUP (n=60) | Hospital specialist care FUP (n=38) | Retrospective cohort | Healthcare system | 5 years | No difference in recurrences detected or QoL | Cost/patient: €112.86 vs. €184.61 (*p*=0.0001) | 62% |
| Breast | CCA | No |
| Beaver (2009) | UK | Hospital-based FUP (n=183) | Nurse-led telephone FUP (n=186) | RCT | Healthcare/Societal | 2 years | No difference in psychological morbidity (STAI), recurrence rate or time to recurrence | NHS FUP cost/patient: £124 vs. £179. Recurrences treatment cost/patient: £143 vs. £182. Transport and productivity cost/patient: £67 vs. £19 | 95% |
| Breast | CMA | Costs: 3.5% |
| Gilbert (2000) | Canada | FP-led FUP | Surgeon-led FUP | Retrospective cohort | Healthcare system | 5 years | Recurrences detected: 78 (70.3%) vs. 26 (23.4%) | Cost/recurrence detected: C$1,105 vs. C$4,387 | 48% |
| Lung | CEA | No |
| Helgesen (2000) | Sweden | On-demand nurse-led FUP (n=200) | Urologist-led FUP (n=200) | RCT | Healthcare system | 3 years | No differences in medical safety and HAD scale | Cost/patient: SEK17,033 vs. SEK19,454  | 71% |
| Prostate | CCA | No |
| Kimman (2011) | Netherlands | Four strategies combining hospital-based or telephone nurse-led FUP with or without EGP | RCT | Societal | 12 months | QALYs: 0.776 vs. 0.772  | ∆Cost/∆QALY: €235.750 (hospital + EGP vs. telephone + EGP) | 91% |
| Breast | CUA | No |
| Koinberg (2009) | Sweden | Physician-led FUP (n=131) | On-demand nurse-led FUP (n=133) | RCT | Healthcare system | 5 years | No difference in HAS, patient satisfaction, recurrences and mortality | Cost/patient/year: €630 (95% CI: €557-€1,055) vs. €495(95% CI: €410-€797) | 81% |
| Breast | CMA | 3.0% |

**Table 2 (cont.).** Key features of economic evaluation studies included in the review (n=39).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **First author** **(Year)** | **Country** | **Intervention** | **Comparator** | **Study design** | **Cost perspective** | **Time horizon** | **Health outcomes** | **Economic** **outcomes** | **Quality score** |
| **Cancer** | **Economic evaluation** | **Discounting** |
| ***Comparisons among different organizational aspects of post-treatment follow-up (cont.)*** |
| Moore (2002) | UK | Nurse-led FUP (n=100) | Conventional medical FUP (n=103) | RCT | Health and social care system | 12 months | No difference in survival or disease progression | Cost/patient: £696.50 vs. £744.50 (*p*=0.66) | 71% |
| Lung | CCA | No |
| Polinder (2009) | Netherlands | Surgeon-led FUP (n=55) | Home-based nurse-led FUP (n=54) | RCT | Societal | 12 months | No difference in recurrences detected, patient satisfaction and QoL | Cost/patient: €3,798 vs. €2,592 (*p*=0.11) | 86% |
| Esophageal | CMA | No |
| Strand (2011) | Sweden | Nurse-led FUP (n=54) | Surgeon-led FUP (n=56) | RCT | Healthcare system | NS | Metastases detected: 8 vs. 7 (*p*=0.953) | Cost/patient: €51 vs. €55 (*p*=0.779) | 57% |
| Colorectal | CCA | No |
| Verschuur (2009) | Netherlands | Home-based, nurse-led FUP (n=54) | Standard outpatient clinic-based FUP (n=55) | RCT | Healthcare system | 13 months | No difference in QoL and patient satisfaction | Cost/patient: €2600 vs. €3800 (*p*=0.11) | 71% |
| Esophageal | CCA | No |
| ***Follow-up with education program versus follow-up without educational program*** |
| Coyle (2013) | Canada | Survivorship care plane (SCP) | Current practice (no SCP) | RCT | Healthcare/Societal | 2 years | No difference in QALYs | Cost/QALY: SCP is dominated | 86% |
| Breast | CUA | 5.0% |

CCA: cost-consequences analysis; CDET: Coincidence detection system; CEA: cost-effectiveness analysis; CMA: cost-minimization analysis; CT: computed tomography; CUA: cost-utility analysis; DES: discrete event simulation; DFS: disease-free survival; EGP: educational group program; FDG: fluorodeoxyglucose; FISH: fluorescence *in situ* hybridization; FP: family physician; FUP: follow-up; HAS: hospital anxiety and depression; HPV: human papilloma virus; KRW: Korean Won; MA: microsatellite-analysis; NMP22: BladderCheck®; NS: not specified; OS: overall survival; PET: positron emission tomography; QoL: quality of life; SEK: Swedish crowns; SCC: squamous cell carcinoma; SCP: survivorship care plane; STAI: State-Trait Anxiety Inventory.

**Table 3a.** Quality assessment of the included studies (n=39) based on CHEERS checklist (Items: 1-12).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Ref.** | **Title** | **Structured****abstract** | **Rationale/****objectives** | **Target****population** | **Setting** | **Study****perspective** | **Comparators** | **Time horizon** | **Discount rate** | **Health outcomes** | **Effectiveness measurement** | **Preferences****measurement** | **Items sum** | **Quality** |
| [14] | N | N | Y | Y | Y | Y | Y | Y | N | Y | Y | NA | 12/21 | Average |
| [15] | Y | Y | Y | N | N | Y | N | Y | N | Y | N | NA | 14/23 | Average |
| [16] | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | 23/24 | High |
| [17] | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | NA | 20/21 | High |
| [18] | N | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | 23/24 | High |
| [19] | N | Y | Y | Y | N | N | Y | Y | N | Y | Y | NA | 13/21 | Average |
| [20] | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | NA | 20/21 | High |
| [21] | N | Y | Y | Y | Y | N | Y | Y | N | Y | Y | Y | 20/24 | High |
| [22] | Y | N | N | Y | Y | Y | N | Y | N | Y | Y | Y | 14/24 | Average |
| [23] | N | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | NA | 16/23 | Average |
| [24] | N | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | N | 18/21 | High |
| [25] | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | 22/24 | High |
| [26] | Y | Y | Y | Y | Y | Y | Y | Y | N | N | Y | NA | 18/23 | High |
| [27] | N | Y | Y | Y | N | N | N | Y | N | Y | Y | NA | 11/21 | Average |
| [28] | N | Y | N | Y | Y | N | Y | Y | N | Y | Y | NA | 10/21 | Poor |
| [29] | N | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | 21/24 | High |
| [30] | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | NA | 20/23 | High |
| [31] | N | Y | Y | Y | Y | Y | Y | Y | N | N | Y | NA | 15/21 | Average |
| [32] | N | Y | Y | Y | Y | Y | N | Y | N | Y | Y | N | 12/24 | Average |
| [33] | N | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | NA | 14/21 | Average |
| [34] | Y | Y | N | Y | Y | Y | Y | Y | Y | Y | Y | N | 18/24 | High |
| [35] | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | Y | 20/22 | High |
| [36] | N | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | NA | 17/21 | High |
| [37] | N | N | Y | Y | Y | Y | Y | Y | N | Y | Y | NA | 12/21 | Average |
| [38] | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | N | NA | 20/23 | High |
| [39] | Y | Y | Y | Y | Y | N | Y | Y | N | Y | Y | NA | 15/23 | Average |
| [40] | N | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | NA | 18/23 | High |
| [41] | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | 24/24 | High |
| [42] | N | N | N | Y | N | Y | Y | Y | N | Y | Y | NA | 14/21 | Average |
| [43] | N | Y | N | Y | Y | Y | Y | Y | N | Y | Y | NA | 15/21 | Average |
| [44] | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | NA | 14/21 | Average |
| [45] | N | N | Y | Y | Y | N | Y | Y | N | Y | Y | NA | 7/21 | Poor |
| [46] | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | NA | 18/21 | High |
| [47] | N | Y | Y | Y | N | N | N | Y | N | Y | Y | NA | 10/21 | Poor |
| [48] | N | Y | Y | Y | Y | Y | Y | N | N | Y | Y | NA | 12/21 | Average |
| [49] | N | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | NA | 15/21 | Average |
| [50] | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | NA | 22/23 | High |
| [51] | N | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | 22/24 | High |
| [52] | N | N | Y | Y | Y | Y | Y | Y | N | Y | Y | NA | 15/21 | Average |
| **Y** | 16 | 41% | 33 | 85% | 34 | 87% | 38 | 97% | 34 | 87% | 32 | 82% | 34 | 87% | 38 | 97% | 16 | 41% | 37 | 95% | 37 | 95% | 9 | 23% |  |  |
| **N** | 23 | 59% | 6 | 15% | 5 | 13% | 1 | 3% | 5 | 13% | 7 | 18% | 5 | 13% | 1 | 3% | 23 | 59% | 2 | 5% | 2 | 5% | 3 | 8% |  |  |
| **NA** | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 27 | 69% |  |  |

**Table 3b.** Quality assessment of the included studies (n=39) based on CHEERS checklist (Items: 13-24).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Ref.** | **Resources/****costs** | **Currency/****price date** | **Model choice** | **Model****assumptions** | **Analytic methods** | **Study****parameters** | **∆costs/****∆outcomes** | **Uncertainty** | **Heterogeneity****(subgroups)** | **Discussion** | **Funding****source** | **Conflict interest** | **Items sum** | **Quality** |
| [14] | Y | Y | NA | NA | Y | N | N | Y | N | N | N | N | 12/21 | Average |
| [15] | Y | Y | N | N | Y | Y | N | Y | N | Y | Y | Y | 14/23 | Average |
| [16] | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | Y | 23/24 | High |
| [17] | Y | Y | NA | NA | Y | Y | Y | Y | N | Y | Y | Y | 20/21 | High |
| [18] | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | 23/24 | High |
| [19] | Y | N | NA | NA | Y | Y | N | Y | N | N | Y | Y | 13/21 | Average |
| [20] | Y | Y | NA | NA | Y | Y | Y | Y | N | Y | Y | Y | 20/21 | High |
| [21] | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | N | 20/24 | High |
| [22] | N | Y | Y | N | Y | N | Y | Y | Y | N | N | N | 14/24 | Average |
| [23] | N | Y | Y | Y | Y | Y | N | Y | N | N | N | N | 16/23 | Average |
| [24] | Y | Y | NA | NA | Y | Y | Y | Y | N | N | Y | Y | 18/21 | High |
| [25] | Y | Y | N | Y | Y | Y | Y | Y | Y | Y | Y | N | 22/24 | High |
| [26] | N | N | Y | Y | Y | Y | Y | Y | N | Y | Y | Y | 18/23 | High |
| [27] | Y | N | NA | NA | Y | Y | N | Y | N | N | N | Y | 11/21 | Average |
| [28] | N | N | NA | NA | Y | N | N | Y | N | N | Y | N | 10/21 | Poor |
| [29] | Y | Y | N | Y | Y | Y | Y | Y | Y | Y | N | Y | 21/24 | High |
| [30] | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | N | Y | 20/23 | High |
| [31] | Y | Y | NA | NA | Y | N | Y | Y | Y | N | Y | N | 15/21 | Average |
| [32] | Y | N | Y | N | N | N | N | N | N | N | Y | Y | 12/24 | Average |
| [33] | Y | Y | NA | NA | N | N | N | N | N | Y | Y | Y | 14/21 | Average |
| [34] | Y | Y | Y | Y | Y | N | Y | Y | N | Y | N | N | 18/24 | High |
| [35] | Y | Y | NA | NA | Y | Y | Y | Y | N | Y | Y | Y | 20/22 | High |
| [36] | Y | Y | NA | NA | Y | Y | Y | Y | N | Y | N | N | 17/21 | High |
| [37] | Y | Y | NA | NA | N | Y | N | N | N | N | Y | N | 12/21 | Average |
| [38] | Y | Y | N | Y | Y | Y | Y | Y | N | Y | Y | Y | 20/23 | High |
| [39] | Y | N | Y | Y | Y | N | N | Y | N | N | N | Y | 15/23 | Average |
| [40] | Y | Y | N | N | Y | Y | Y | Y | Y | N | N | Y | 18/23 | High |
| [41] | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | 24/24 | High |
| [42] | Y | Y | NA | NA | Y | Y | Y | Y | N | Y | N | Y | 14/21 | Average |
| [43] | Y | Y | NA | NA | Y | N | N | Y | N | Y | Y | Y | 15/21 | Average |
| [44] | Y | N | NA | NA | N | N | N | Y | N | Y | N | Y | 14/21 | Average |
| [45] | N | N | NA | NA | N | N | N | N | N | N | N | N | 7/21 | Poor |
| [46] | Y | Y | NA | NA | Y | Y | N | Y | N | Y | N | Y | 18/21 | High |
| [47] | N | N | NA | NA | Y | N | N | Y | Y | Y | N | N | 10/21 | Poor |
| [48] | Y | N | NA | NA | Y | N | N | Y | N | Y | N | N | 12/21 | Average |
| [49] | Y | Y | NA | NA | N | N | Y | N | Y | Y | N | Y | 15/21 | Average |
| [50] | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | Y | Y | 22/23 | High |
| [51] | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | N | Y | 22/24 | High |
| [52] | Y | Y | NA | NA | Y | N | N | Y | N | Y | Y | Y | 15/21 | Average |
| **Y** | 33 | 85% | 29 | 74% | 13 | 33% | 14 | 36% | 33 | 85% | 24 | 62% | 22 | 56% | 34 | 87% | 11 | 28% | 26 | 67% | 21 | 54% | 26 | 67% |  |  |
| **N** | 6 | 15% | 10 | 26% | 5 | 13% | 4 | 10% | 6 | 15% | 15 | 38% | 17 | 44% | 5 | 13% | 28 | 72% | 13 | 33% | 18 | 46% | 13 | 33% |  |  |
| **NA** | 0 | 0% | 0 | 0% | 21 | 54% | 21 | 54% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% |  |  |