The cost-effectiveness of biomarkers for predicting the development of oesophageal adenocarcinoma
Rubenstein J H, Vakil N, Inadomi J M

Record Status
This is a critical abstract of an economic evaluation that meets the criteria for inclusion on NHS EED. Each abstract contains a brief summary of the methods, the results and conclusions followed by a detailed critical assessment on the reliability of the study and the conclusions drawn.

Health technology
The study examined two strategies, biomarker-guided surveillance (MS) and biomarker-guided oesophagectomy (MOE), for the detection of oesophageal adenocarcinoma (OAC).

With MS, the patients underwent oesophagogastroduodenoscopy (EGD) with testing for a biomarker, or panel of biomarkers, at age 50. Those not found to have the biomarker did not undergo surveillance. Patients with a positive result underwent repeat EGD with biopsy for histology every 3 months, to survey for cancer. Biomarker testing was repeated at the time of each EGD in order to detect patients with initial false-positive biomarker results.

The MOE strategy was similar to MS and all patients underwent EGD and biomarker testing at age 50. Patients with a negative biomarker test did not undergo surveillance. Those with a positive biomarker test immediately underwent prophylactic oesophagectomy without any additional surveillance. In addition, those diagnosed with a prevalent cancer by the initial EGD underwent oesophagectomy if resectable.

Candidate biomarkers from oesophageal tissue included aneuploidy, increased tetraploidy, 17p (p53) loss of heterozygosity, 9p (p16) loss of heterozygosity, and cyclin-D1 overexpression. An idealised biomarker (or panel of biomarkers) was chosen for this analysis.

Type of intervention
Screening.

Economic study type
Cost-utility analysis.

Study population
The study population comprised a hypothetical cohort of 50-year-old Caucasian men with a history of GORD.

Setting
The setting was secondary care. The economic study was carried out in the USA.

Dates to which data relate
The effectiveness data were derived from studies published between 1976 and 2002. The dates when the resource use data were collected were unclear. The costs came from sources published between 1994 and 2001. Some costs referred to 2001 values, but the price year was not explicitly stated.

Source of effectiveness data
The effectiveness evidence was derived from a synthesis of completed studies and authors' assumptions.
Modelling
A Markov model was constructed to assess the clinical and economic impact of biomarkers. A simplified schematic of the model was reported. Two different pathways were assumed in the case of OAC predicted by dysplasia or by the biomarker. The natural history of progressing from Barrett’s with and without dysplasia to cancer was taken from a published model. In the model where the biomarker predicted OAC, the biomarker state was assumed to already be prevalent in a sub-group of 50-year-old men with symptoms of GORD. Patients lacking the biomarker at that age had no chance of developing it later. Patients with the biomarker could not regress from it. Further, there was no chance of developing cancer without the biomarker, and all people with the biomarker were destined to develop cancer if they did not die from other causes first. The health states considered were no Barrett’s, Barrett’s without dysplasia, Barrett’s with dysplasia, cancer, oesophagectomy and dead. The patients were followed until age 80 or death. The cycle length was one year.

Outcomes assessed in the review
The outcomes estimated from the literature were:

prevalence data,
transition probabilities (progression and regression rates),
cancer treatment probabilities,
test inaccuracy rates, and
utility values.

Study designs and other criteria for inclusion in the review
The authors stated that the transition rates between model health states were derived from a review of the literature. However, the methods of the review were not reported. Further, the sources of other model inputs were not stated.

Sources searched to identify primary studies
Not stated.

Criteria used to ensure the validity of primary studies
Not reported.

Methods used to judge relevance and validity, and for extracting data
Not stated.

Number of primary studies included
Forty-six studies provided the clinical evidence used in the model.

Methods of combining primary studies
Not stated.

Investigation of differences between primary studies
Not stated.
Results of the review
The prevalence of Barrett’s oesophagus in GORD was 0.1. The prevalence of low-grade dysplasia (LGD) in GORD was 0.01. The prevalence of high-grade dysplasia (HGD) in GORD was 0.007. The prevalence of cancer in Barrett’s oesophagus was 0.067.

With respect to annual progression rates, the estimates used were:

from no dysplasia to LGD, 0.05;
from no dysplasia to HGD, 0.01;
from no dysplasia to cancer, 0.005;
from LGD to HGD, 0.05;
from LGD to cancer, 0.025; and
from HGD to cancer, 0.055.

The cancer incidence in Barrett’s oesophagus was 0.005.

With respect to annual regression rates, the estimates used were:

from Barrett’s oesophagus to normal state, 0.0175;
from LGD to no dysplasia, 0.63;
from HGD to no dysplasia, 0.1; and
from HGD to LGD, 0.07.

Cancer treatment probabilities were:

0.5 for resectability without surveillance;
0.95 for resectability with surveillance;
0.05 for surgical mortality without surveillance;
0.027 for surgical mortality with surveillance;
0.2 for cancer cure without surveillance;
0.8 for cancer cure with surveillance; and
0.000021 for mortality due to endoscopy.

The utility value for post-oesophagectomy was 0.97.

Extensive information on the diagnostic inaccuracy rates was provided in the paper.

Methods used to derive estimates of effectiveness
The authors made some assumptions to derive clinical estimates used in the decision model.
Estimates of effectiveness and key assumptions
The annual probability of developing Barrett's oesophagus was 0.005.

The annual probability of progress from biomarker positive to cancer was 0.055.

The probability of mortality due to unresectable cancer was 0.9.

The utility value for cancer was 0.5.

The probability of a patient with a negative biomarker test being falsely diagnosed as having cancer was the same as the probability of a patient with LGD being falsely diagnosed with cancer, which was set at 5%.

The probability of a patient with no biomarker being falsely diagnosed as having cancer was the same as the probability of a patient without Barrett's oesophagus being falsely diagnosed with cancer, which was set at 0.5%.

In the best-case scenario, the biomarker was assumed 100% sensitive and 100% specific.

Measure of benefits used in the economic analysis
The summary benefit measure used was the quality-adjusted life-years (QALYs). These were calculated by combining survival and quality of life data derived from the literature or based on authors' assumptions. Expected survival was discounted at an annual rate of 3%. The numbers of cancer cases, cancer deaths and deaths from oesophagectomy were also reported as model outputs.

Direct costs
The analysis of the costs was carried out from the perspective of the third-party payer. The categories of costs included were endoscopy with biopsies, oesophagectomy, endoscopic palliation, annual postsurgical care, incurable cancer care and clinic visit. The unit costs were presented for most items, but information on the quantities of resources used was less clear. A breakdown of the cost items was not given. The source of the resource consumption data was unclear, although the data appear to have been derived from both published studies and authors' assumptions. The costs were derived from the literature and from Medicare and Medicaid Services. The authors set the cost of the biomarker. Discounting was relevant since long-term costs were included in the analysis, and a 3% discount rate was used. The price year was not stated, although some costs were based on the fiscal year 2001.

Statistical analysis of costs
The costs were treated deterministically.

Indirect Costs
The indirect costs were not considered.

Currency
US dollars ($).

Sensitivity analysis
Three-way sensitivity analyses were carried out for both MS and MOE on the threshold combination of sensitivity, specificity, and cost of the biomarker required for it to be cost-effective given an assumed willingness to pay. Further, univariate sensitivity analyses were performed on the sensitivity and specificity of the biomarker needed for MS or MOE to prevent more cancers and more cancer deaths than dysplasia-guided surveillance. In a hypothetical cohort of 10,000 GORD patients, observation alone (natural history) led to 366 cancers, 356 cancer deaths, and less than 1 death from oesophagectomy. The total QALYs per patient were 16.466. Dysplasia-guided surveillance prevented 14% of
cancers and 56% of cancer deaths compared with observation alone. The patients experienced an average of 16.637 QALYs.

Assuming 100% sensitivity and specificity of the biomarker, MS prevented 53% of cancers and 87% of cancer deaths compared with observation alone. The patients experienced an average of 16.691 QALYs.

Assuming 100% sensitivity and specificity of the biomarker, MOE prevented 76% of cancers and 93% of cancer deaths compared with observation alone. The patients experienced an average of 16.707 QALYs.

In all surveillance strategies, out of 10,000 GORD patients, 11 died as a complication of oesophagectomy.

Cost results
The average lifetime cost per patient was $104 with observation alone, $2,444 with dysplasia-guided surveillance, $2,356 with MS and $2,291 with MOE (assuming a cost of $100 for the biomarker).

Synthesis of costs and benefits
Incremental cost-utility ratios were calculated to combine the costs and benefits of the alternative surveillance strategies.

The incremental cost per QALY gained with dysplasia-guided surveillance over observation alone was $14,211. The incremental cost per QALY with MS over dysplasia-guided surveillance was $20. The incremental analysis showed that MOE dominated both MS and dysplasia-guided surveillance (by being both more effective and less costly). The incremental cost per QALY gained with MOE over observation alone was $9,055.

The univariate sensitivity analysis showed that for MS to prevent more cancer deaths than dysplasia-guided surveillance, the sensitivity of the biomarker had to be at least 70%, while for MOE the sensitivity had to be at least 53%. In general, regardless of the specificity of the biomarker, either MS or MOE always prevented more cancer deaths than dysplasia-guided surveillance.

In the comparison between MS and dysplasia-guided surveillance, assuming a willingness to pay of $50,000 per additional QALY gained, the three-way sensitivity analysis led to the following results. At a cost of $100 for the biomarker, MS would be cost effective if both the sensitivity and specificity of the biomarker were at least 80%. MS would also be cost effective if the sensitivity was 64% simultaneously with a specificity of 100%, or with a sensitivity of 100% simultaneously with a specificity of 62%. The threshold sensitivities and specificities remained relatively similar despite changing the cost of the biomarker to $10; however, if the cost of the biomarker was $1,000, the sensitivity and specificity would need to simultaneously be 90% for MS to be cost effective.

For MOE versus dysplasia-guided surveillance, at a biomarker cost of $100, MOE was cost effective if the sensitivity was as low as 53% with a specificity of 100%, but the specificity needed to be at least 95% even if the sensitivity was 100%. At a biomarker cost of $1,000, the sensitivity could be as low as 64% if the specificity was 100%, but the specificity needed to be at least 96% even with a sensitivity of 100%.

Authors' conclusions
Strategies for managing Barrett's oesophagus guided by a perfect biomarker could substantially reduce mortality from oesophageal adenocarcinoma (OAC). In addition, they would be more cost-effective than the current recommended dysplasia-guided strategy.

CRD COMMENTARY - Selection of comparators
A justification for the choice of the comparators was provided. They are appropriate surveillance strategies in the authors' setting. The current strategy was compared with no surveillance and with strategies based on idealised biomarkers. You should decide whether they are valid comparators in your own setting.
Validity of estimate of measure of effectiveness
The clinical evidence came from a review of the literature, although the methods and conduct of the review were not reported. For example, inclusion criteria were not stated and the methods used to pool the primary estimates were not provided. The issue of homogeneity across the primary studies was not addressed. Several assumptions were also made, mainly because of the lack of data on the accuracy of the biomarkers. The issue of uncertainty surrounding some clinical data was tested in the sensitivity analysis.

Validity of estimate of measure of benefit
The use of QALYs as the summary benefit measure was appropriate as they captured the effect of the interventions on quality of life and survival, which represent two relevant dimensions of care. Discounting was applied to life expectancy, as guidelines for economic evaluations suggest. Limited information on the source of the quality of life data was provided. The impact of alternative assumptions on quality of life was not investigated.

Validity of estimate of costs
The cost analysis was carried out in accordance with the perspective stated. Most costs were presented as macro categories and a detailed breakdown of the items was not reported. This might limit the possibility of replicating the analysis in other settings. The source of the data was reported, but there were few details of the methods used to derive the costs. Overall, the information given on the analysis of the costs was limited. Discounting was applied, but the impact of using alternative discount rates was not investigated. The costs were treated deterministically but some cost estimates were varied in the sensitivity analysis. The price year was not stated, which makes reflation exercises in other time periods difficult.

Other issues
The authors did not compare their findings with those from other studies, probably due to the lack of published cost effectiveness analyses of biomarkers for detecting OAC. The issue of the generalisability of the study results to other settings was addressed. The authors noted that caution is required when extrapolating their findings to other populations, owing to variability in the prevalence of Barrett’s oesophagus and OAC incidence. Sensitivity analyses were carried out on key clinical and economic estimates, which enhance the robustness of the model results. The authors stated that the model was biased in favour of biomarkers, in order to define the best scenario for such a new surveillance strategy.

Implications of the study
The study results support the development of accurate biomarkers for the surveillance of Barrett’s oesophagus in order to detect OAC. However, the authors pointed out that the results of their model are not immediately applicable to current decision making since a single biomarker with the characteristics delineated in the study is not available. Rather, the current analysis should guide future research (i.e. pilot studies) on these biomarkers. A further implication of the study was that the currently recommended surveillance strategy misses a substantial number of cancer cases.

Source of funding
Supported by the National Institutes of Health and the Department of Veterans Affairs, Veterans Health Administration, Health Service Research and Development Service.

Bibliographic details

PubMedID
16011672
DOI
10.1111/j.1365-2036.2005.02536.x

Other publications of related interest


Indexing Status
Subject indexing assigned by NLM

MeSH
Adenocarcinoma /diagnosis /economics; Biomarkers /blood; Cost-Benefit Analysis; Esophageal Neoplasms /diagnosis /economics; Humans; Male; Markov Chains; Middle Aged; Quality-Adjusted Life Years; Sensitivity and Specificity

AccessionNumber
22005001250

Date bibliographic record published
30/04/2006

Date abstract record published
30/04/2006