Costs and effects of ultrasonography in the evaluation of palpable breast masses
Flohe K Kessels A G Severens J L Beets G L de Koning H J von Meyenfeldt M F van Engelshoven 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
Several diagnostic strategies for the evaluation of palpable breast masses (PBM) were examined. The analysis focused on the possible role played by ultrasonography (US) as an adjunct to clinical examination (CE) and mammography (MAM). The diagnosis from US was scored on a 5-point grading scale of increasing suspicion for malignancy, ranging from 1 (normal) to 5 (malignant). For the definition of positive and negative cases, a cut-off point was used between benign (score of 2) and probably benign (score of 3) results. A combined diagnosis from MAM and CE together and results from core needle biopsy were scored on a similar grading scale. Fine-needle aspiration cytology (FNAC) was recorded on a 4-point grading scale, ranging from 1 (normal) to 4 (malignant) with a score of 0 for indeterminate. Four strategies were considered.

The conventional strategy consisted of routine FNAC after MAM and CE in all patients.

In experimental strategy 1, US was performed on all patients, and a cut-off point between benign and probably benign imaging diagnosis was used for further referral for additional FNAC. Patients were discharged when no abnormality was found. Palpable cysts were aspirated under US guidance. Solid benign structures, such as fibroadenoma, were removed by simple surgical excision. All suspicious findings on US were excised surgically.

In experimental strategy 2, the effect of shifting the cut-off point for referral for FNAC to normal and benign imaging diagnosis was studied. Assumptions were similar to those in experimental strategy 1.

In experimental strategy 3, US was only performed on patients with normal or benign results on MAM and CE, whereas immediate FNAC was performed in patients with suspicious lesions.

Type of intervention
Diagnosis.

Economic study type
Cost-effectiveness analysis.

Study population
The study population comprised a hypothetical cohort of women with PBM who were referred for diagnostic breast imaging.

Setting
The setting was a hospital. The economic study was carried out in the Netherlands.

Dates to which data relate
The effectiveness data and some resource use data were derived from studies published between 1990 and 2003. The price year was 2000.
Source of effectiveness data
The effectiveness data was derived from a synthesis of completed studies and experts' opinions.

Modelling
A decision model was used to compare the costs and benefits of the four alternative strategies. A graphical representation of the decision model was presented. The model included:

diagnostic procedures such as US, FNAC, and core needle biopsy;
primary therapeutic events such as surgical excision, radiotherapy, chemotherapy and hormonal therapy; and
the treatment of local recurrences, distant recurrences, and palliative care at the terminal stage of the disease.

True-positive and false-negative cases were considered. All the possible pathways assumed in the model were described graphically.

Outcomes assessed in the review
The outcomes estimated from the literature were the probability values associated with the performance of the diagnostic tests and surgery or palliative care, and life expectancy (LE).

Study designs and other criteria for inclusion in the review
A systematic review of the literature was not undertaken to identify the primary estimates. Most of the clinical data were derived from a large prospective study carried out in The Netherlands, which enrolled 3,835 breasts in 2,020 consecutive patients. Of these, 522 PBMs in 492 patients (mean age 49 years; breast cancer prevalence 19%) were found. The data from this sample were used to estimate the majority of the parameters. Other data were retrieved from life tables and other published studies.

Sources searched to identify primary studies
Not relevant.

Criteria used to ensure the validity of primary studies
The validity of the primary studies was not discussed, but the use of a large prospective study enhances the robustness of the primary estimates. Details of the other sources of evidence were not provided.

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

Number of primary studies included
Seven primary studies provided evidence.

Methods of combining primary studies
A narrative approach appears to have been used to combine the primary estimates.

Investigation of differences between primary studies
Not stated.
Results of the review
The following average probability values for experimental strategy 1 were estimated from the literature.

Suspicious (grade 3-4-5) US result, 0.349 (range: 0.314 - 0.385).

Malignant (grade 4) FNAC result after US, 0.324 (range: 0.267 - 0.386).

True-positive result surgery after malignant FNAC and US, 0.983 (range: 0.922 - 0.999).

Benign (grade 2) FNAC result after US, 0.439 (range: 0.363 - 0.517).

False-negative result surgery after benign FNAC and US, 0.074 (range: 0.026 - 0.162).

Malignant (grade 3-4-5) biopsy result after FNAC and US (true-positive), 0.435 (range: 0.333 - 0.541).

Surgery after benign (grade 2) US result, 0.126 (range: 0.098 - 0.160).

False-negative result surgery after benign US result, 0.023 (range: 0.001 - 0.106).

False-negative result of discharge after normal (grade 1) US result, 0.000 (range: 0.000 - 0.050).

The following average probability values for experimental strategy 2 were estimated.

Suspicious (grade 2-3-4-5) US result, 0.761 (range: 0.728 - 0.791).

Malignant (grade 4) FNAC result after US, 0.149 (range: 0.120 - 0.181).

True-positive result surgery after malignant FNAC and US, 0.983 (range: 0.922 - 0.999).

Benign (grade 2) FNAC result after US, 0.580 (range: 0.534 - 0.625).

False-negative result surgery after benign FNAC and US, 0.107 (range: 0.073 - 0.151).

Malignant (grade 3-4-5) biopsy result after FNAC and US (true-positive), 0.099 (range: 0.061 - 0.150).

Surgery after benign (grade 2) US result, 0.000 (range: 0.000 - 0.050).

False-negative result surgery after benign US result, 0.000 (range: 0.000 - 0.050).

False-negative result of discharge after normal (grade 1) US result, 0.000 (range: 0.000 - 0.050).

The following average probability values for experimental strategy 3 were estimated.

Suspicious (grade 3-4-5) CE plus MAM result, 0.753 (range: 0.712 - 0.784).

Suspicious (grade 3-4-5) US result, 0.078 (range: 0.043 - 0.128).

Malignant (grade 4) FNAC result after US, 0.200 (range: 0.037 - 0.507).

True-positive result surgery after malignant FNAC and US, 1.000.

Benign (grade 2) FNAC result after US, 0.500 (range: 0.193 - 0.807).

False-negative result surgery after benign FNAC and US, 0.000 (range: 0.000 - 0.050).

Malignant (grade 3-4-5) biopsy result after FNAC and US (true-positive), 0.250 (range: 0.001 - 0.106).

Surgery after benign (grade 2) US result, 0.109 (range: 0.066 - 0.168).
False-negative result surgery after benign US result, 0.000 (range: 0.000 - 0.050).
False-negative result of discharge after normal (grade 1) US result, 0.000 (range: 0.000 - 0.050).
The following average probability values common to all strategies were estimated.
Local recurrence after primary surgery, 0.100 (range: 0.05 - 0.150).
Palliative stage after diagnosis, 0.010 (range: 0.005 - 0.015).
Palliative stage after adjuvant therapy, 0.440 (range: 0.220 - 0.660).
Palliative stage after non-adjuvant therapy, 0.220 (range: 0.110 - 0.330).
Palliative stage after local recurrence, 0.670 (range: 0.335 - 1.000).

LE was as follows:
for palliative stage after diagnostic test, 1.5 years (range: 0.8 - 2.3);
for palliative stage after local recurrence, 5 years (range: 2.5 - 7.5);
for palliative stage after local and distant recurrence, 8 years (range: 4 - 12);
for disease-free after diagnostic test, 32.8 years (range: 16 - 49);
for disease-free after treatment, 32.8 years (range: 16 - 49).

Methods used to derive estimates of effectiveness
The opinions of an expert panel were used to derive some unpublished clinical estimates used in the decision model.

Estimates of effectiveness and key assumptions
For example, the following estimates for experimental strategy 1 were made.
Malignant (grade 4) FNAC result, 0.115 (range: 0.058 - 0.173).
True-positive result surgery after malignant (grade 4) FNAC, 0.983 (range: 0.492 - 1.000).
Benign (grade 2) FNAC result, 0.208 (range: 0.104 - 0.312).
False-negative result surgery after benign (grade 2) FNAC, 0.042 (range: 0.021 - 0.063).
Biopsy after probably benign (grade 0-3) FNAC result, 0.189 (range: 0.095 - 0.280).
Malignant (grade 3-4-5) biopsy result after FNAC (true-positive), 0.435 (range: 0.218 - 0.653).
Biopsy after normal (grade 1) FNAC result, 0.300 (range: 0.150 - 0.450).
Surgery after normal (grade 1) FNAC, 0.140 (range: 0.070 - 0.210).
Other assumptions were made for strategies 2 and 3 and were reported in the article.

Measure of benefits used in the economic analysis
The summary benefit measure used was the estimated LE associated with each diagnostic strategy. This was derived
using the decision model. Data on LE were obtained from the literature. Discounting was not applied.

**Direct costs**
The analysis of the costs took the perspective of the health care system. Thus, only the direct medical costs of FNAC, US, core needle biopsy, excision biopsy, curative surgery, adjuvant therapy, treatment for local recurrence, palliative care and treatment of disease-free patients were considered. For each category of costs, the main cost components were reported. These included hospital or nursing care, diagnostic radiology, pathology, surgery, radiotherapy, hormone therapy, chemotherapy and specialist control. The unit costs were not explicitly presented separately from the quantities of resources but, since resource use depended on different pathways of the model, were reflected by the probabilities of events. Also, unit costs were reported for most items. Resource use was estimated from published studies, while the costs were estimated from the Dutch Health Care Insurance Board, the Hospital Information System of the Maastricht University Hospital, the Pharmacotherapeutic Guidelines 2000-2001, and the Dutch Radiotherapy Guidelines for Breast Cancer. Discounting was not relevant as the costs per patient were incurred during a short timeframe. The price year was 2000.

**Statistical analysis of costs**
The costs were treated deterministically.

**Indirect Costs**
The indirect costs were not considered in the analysis.

**Currency**
Euros (EUR).

**Sensitivity analysis**
A univariate sensitivity analysis was carried out on almost all model inputs to assess the robustness of the costs and benefits. Ranges of values were based on 90% confidence intervals for clinical data. For deterministic inputs, ranges were derived using +/- 50% of the base-case value. A threshold analysis was also performed to identify the values at which the preferred strategy would alter.

**Estimated benefits used in the economic analysis**
The expected LE was 31.0 years with all diagnostic strategies.

**Cost results**
The total costs were:

- EUR 3,087 with the conventional strategy (EUR 182 for diagnosis, EUR 1,472 for treatment and EUR 1,433 for follow-up care),
- EUR 3,047 (EUR 165 for diagnosis, EUR 1,449 for treatment and EUR 1,433 for follow-up care) with experimental strategy 1,
- EUR 3,512 (EUR 294 for diagnosis, EUR 1,785 for treatment and EUR 1,433 for follow-up care) with experimental strategy 2, and
- EUR 3,013 (EUR 122 for diagnosis, EUR 1,458 for treatment and EUR 1,433 for follow-up care) with experimental strategy 3.
Strategies 1 and 3 were less costly than the conventional strategy, while strategy 2 was more expensive.

**Synthesis of costs and benefits**
A synthesis of the costs and benefits was not relevant as a cost-minimisation analysis was carried out. In effect, the same LE was associated with the four diagnostic strategies. The univariate sensitivity analysis showed that the most influential variables were the proportion of suspicious and benign FNAC results and the proportion of false-negative US results, which mainly affected the cost side of the analysis. LE was quite robust to any variation in the model inputs. However, the results of the sensitivity analysis were not reported. The threshold analysis suggested that the conventional strategy was the preferred diagnostic approach only when the cost of US increased to EUR 163 (it was EUR 117 in the base-case).

**Authors' conclusions**
Incorporating ultrasonography (US) in the triple assessment for the evaluation of palpable breast masses (PBM) could result in a reduction of the total costs for the diagnosis and treatment of breast cancer. The most efficient strategy consisted of US of lesions without suspicion of breast cancer based on mammography (MAM) and clinical examination (CE), and immediate fine-needle aspiration cytology (FNAC) of suspicious lesions. Despite small per-patient cost-savings, the authors suggested that patients might prefer less invasive diagnostic approaches such as US.

**CRD COMMENTARY - Selection of comparators**
The rationale for the choice of the comparators was clear, and the authors justified the alternative diagnostic strategies examined in the study. These options were clearly described. You should decide whether they are valid comparators in your own setting.

**Validity of estimate of measure of effectiveness**
The effectiveness analysis relied mainly on data derived from a large, prospective study of consecutive patients that was carried out in the Netherlands. Other data were derived both from selectively identified published studies and from a panel of experts. For example, LE came from Dutch statistics. Details of the approach used to elicit the experts' opinions were not reported. The issue of uncertainty surrounding clinical estimates was addressed in the sensitivity analysis, but the results of the sensitivity analysis were not presented.

**Validity of estimate of measure of benefit**
No summary benefit measure was used because a cost-minimisation analysis was conducted. However, LE was the main model output, which was appropriate since survival captures the final impact of the intervention on the patients' health. The analysis did not consider aspects related to quality of life, which would have been interesting. The use of LE makes the benefits considered in the current study comparable with those estimated in other economic evaluations. The reasons for not applying discounting were not stated.

**Validity of estimate of costs**
The analysis of the costs was performed from the perspective of a third-party payer, which guided the choice of the model inputs and the sources used to derive the costs. In effect, non-medical direct costs and indirect costs were not relevant and were not included. The costs were broken down into sub-categories and the unit costs were presented for most items. This enhances the possibility of replicating the analysis in other settings. The cost estimates were specific to the study setting but the effect of changes in cost estimates was investigated in the sensitivity analysis. Again, the results of this analysis were not reported. Statistical analyses were not performed. The price year was reported, which will facilitate reflation exercises in other time periods.

**Other issues**
The authors made limited comparisons of their findings with those from other studies. It was noted that caution is
required when transferring the current cost results to other countries, owing to the differences in treatment patterns. The issue of the generalisability of the study results to other settings was implicitly addressed in the sensitivity analysis. However, the results of this analysis were presented selectively. The authors highlighted that it was assumed that all malignancies detected by US in the experimental strategies were also detected by FNAC in the conventional strategy. This assumption might have overestimated the diagnostic performance of FNAC and underestimated that of US. The study referred to women with PBMs and this was reflected in the authors' conclusions.

**Implications of the study**
The study results supported the use of US in the diagnostic work-up of all patients with PBMs.

**Source of funding**
Supported by the Dutch Health Care Insurance Board, Amstelveen, the Netherlands.

**Bibliographic details**

**PubMedID**
15609793

**Other publications of related interest**


**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Adolescent; Adult; Aged; Aged, 80 and over; Biopsy, Fine-Needle /economics; Breast /pathology; Breast Neoplasms /diagnosis /economics; Costs and Cost Analysis; Decision Support Techniques; Female; Humans; Middle Aged; Palpation /economics; Prospective Studies; Ultrasonography, Mammary /economics

**AccessionNumber**
22004008423

**Date bibliographic record published**
31/05/2006

**Date abstract record published**
31/05/2006