Making cost-effectiveness analyses clinically relevant: the effect of provider expertise and biliary disease prevalence on the economic comparison of alternative diagnostic strategies


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
Three strategies for the diagnosis and treatment of biliary disease (BD) were examined. The strategies were endoscopic retrograde cholangiopancreatography (ERCP), magnetic resonance cholangiopancreatography (MRCP) followed by ERCP, and endoscopic ultrasonography (EUS) followed by ERCP.

Type of intervention
Diagnosis and treatment.

Economic study type
Cost-effectiveness analysis.

Study population
The patients included and the target population for the model were those with suspected BD who were referred for ERCP.

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

Dates to which data relate
The effectiveness evidence was obtained from studies dating from 1997 to 2002. The cost data and the price year related to year 2000.

Source of effectiveness data
The evidence was derived from a single study (a parent study), a review or synthesis of completed studies, and estimates based on authors' opinions.

Link between effectiveness and cost data
The costing was undertaken retrospectively on a different patient sample from that used in the modelling.

Study sample
The study sample was composed of 30 consecutive adult patients (more than 18 years of age) who were referred to a teaching medical centre for ERCP for suspected BD between July 1997 and December 1998 because of clinical signs and symptoms (biliary pain), and/or the results of traditional diagnostic tests (abnormal transcutaneous sonogram or liver enzyme levels). The patients underwent MRCP, EUS, or ERCP within 24 hours of referral. No power or sample size calculations were reported. Additional information may be available from the parent study (Scheiman et al. 2001,
see 'Other Publications of Related Interest' for bibliographic details).

**Study design**
The study was a diagnostic accuracy study. This consisted of an analysis of test operating characteristics, as well as an evaluation of inter-observer variability in the case of MRCP. This last component was used to assess the impact of MRCP provider expertise on cost-effectiveness. The authors employed reader-specific MRCP test characteristics derived from the analysis. The readers were blinded to prior interpretations of the ERCP and MRCP studies.

**Analysis of effectiveness**
The primary outcomes were the rate of BD prevalence, the sensitivity and specificity of each strategy, and the inter-observer variability to MRCP.

The sensitivity and specificity of each strategy were assessed using ERCP as the reference standard. Patients were also categorised as true-positive diagnosis with appropriate treatment, true-negative diagnosis without additional treatment, false-positive diagnosis with inappropriate follow-up ERCP, and false-negative diagnosis with subsequent symptom recurrence or complication.

For the determination of inter-observer variability, two experienced radiologists performed a post hoc analysis of MRCP data that had been prospectively acquired. The authors limited post hoc analysis of observer variability to MRCP since it was not as time-dependent as EUS and because ERCP, as the reference standard, was assumed not to be subject to observer variability.

**Effectiveness results**
The prevalence of BD was 0.29.

The diagnostic accuracy of the tests was as follows:

for MRCP, sensitivity 0.50 and specificity 0.65; and

for EUS, sensitivity 0.75 and specificity 0.95.

As the reference standard, both the sensitivity and specificity of ERCP were 1.00.

No statistically significant difference in specificity was found in the post hoc analysis comparing MRCP with EUS, (p>0.05 for both readers).

Over all readers, the average sensitivity of MRCP for detecting abnormality was 0.50 and the average specificity was 0.77.

The inter-observer agreement was moderate (0.57) for choledocholithiasis detection and slight (0.17) for stricture detection.

The inter-modality agreement was kappa 0.36 for choledocholithiasis and kappa 0.25 for stricture detection.

**Clinical conclusions**
To the authors’ knowledge, no previously reported studies have estimated the inter-observer and cost-effectiveness variances due to differences in MRCP provider expertise. They have empirically determined the performance characteristics of MRCP and the variability in MRCP interpretation (as a measure of provider expertise) in a population with low disease prevalence.

**Modelling**
A decision analytic model simulating the diagnosis and treatment of BD was built. The algorithms were designed to resemble clinical practice, with clinical decision-making based on the reported results of commonly available diagnostic tests.

**Outcomes assessed in the review**
The parameters used in the model were the numbers of true-positive and false-positive diagnoses of biliary abnormalities, and the rates of complication.

**Study designs and other criteria for inclusion in the review**
No inclusion criteria for a review of any of the parameters were reported.

**Sources searched to identify primary studies**
Not reported.

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

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

**Number of primary studies included**
The authors reported that three primary studies provided the effectiveness evidence.

**Methods of combining primary studies**
A narrative method was used to combine the studies.

**Investigation of differences between primary studies**
Not reported.

**Results of the review**
The parameters used in the model were as follows:

- true-positives were 0.29 for ERCP, 0.21 for EUS-ERCP, and 0.14 for MRCP-ERCP;
- true-negatives were 0.71 for ERCP, 0.68 for EUS-ERCP, and 0.47 for MRCP-ERCP;
- false-positives were 0 for ERCP, 0.04 for EUS-ERCP, and 0.25 for MRCP-ERCP; and
- false-negatives were 0 for ERCP, 0.07 for EUS-ERCP, and 0.14 for MRCP-ERCP.

The rates of pancreatitis were 0.05 after diagnostic ERCP and 0.07 after therapeutic ERCP.

The rate of ERCP-related complications was 0.05 for ERCP, 0.02 for EUS-ERCP, and 0.03 for MRCP-ERCP.

**Methods used to derive estimates of effectiveness**
The authors made assumptions to derive some estimates of effectiveness.
Estimates of effectiveness and key assumptions
For all strategies, the rate of recurrent symptoms in false-negative cases was assumed to be 80%, with a 10% risk of biliary colic or cholangitis requiring hospital admission.

To the authors' knowledge, there were no data on the prevalence of various renal lesions in the population suspected of having BD. Therefore, the authors focused solely on the ability of EUS and MRCP to depict BD, to prevent complicating the analysis unnecessarily and thus introducing an unacceptable level of uncertainty.

Measure of benefits used in the economic analysis
The outcome measures used in the economic analysis were the numbers of additional biliary abnormalities detected and false-positive diagnoses averted.

Direct costs
The direct costs were for outpatient procedures and hospitalisation for uncomplicated pancreatitis and cholangitis. Outpatient procedures covered MRCP, EUS, diagnostic ERCP, therapeutic ERCP, emergency room visits and transcutaneous ultrasound. The resource quantities and the unit costs were not reported separately. Procedural costs were derived from Medicare reimbursement data, while hospitalisation costs resulting from procedural complications were derived from diagnosis-related group reimbursements. Discounting was appropriately not carried out because of the short-term horizon of the study (i.e. less than 2 years). The quantities and the total costs were estimated through modelling. The price year was 2000.

Statistical analysis of costs
No statistical analysis of the costs was reported.

Indirect Costs
The indirect costs were not included in the study.

Currency
US dollars ($).

Sensitivity analysis
Two sensitivity analyses were performed to investigate areas of uncertainty related to variability in the data. First, to assess the independent effects of EUS and MRCP test characteristics, the authors performed analyses using the 95% confidence intervals for overall sensitivity and specificity for each modality. Second, to assess the robustness of the model to specific cost and clinical variables, the authors analysed sensitivity by altering the cost of each modality, the probability of recurrent symptoms, and the probability of biliary colic or cholangitis after a missed biliary stone or stricture. The results of the sensitivity analyses were used as inputs for a tornado analysis, which was performed to determine the relative influence of each variable examined on the relative cost-effectiveness of the three diagnostic strategies.

Estimated benefits used in the economic analysis
With a 6% rate of false-positive diagnoses (4 out of 71), 72% of all biliary abnormalities (21 out of 29) were detected with EUS-ERCP.

With a 35% false-positive rate (25 out of 71), 48% of all biliary abnormalities (14 out of 29) were detected with MRCP-ERCP.
Cost results
The cost per patient was $1,535 for ERCP, $1,145 for EUS-ERCP, and $1,148 for MRCP-ERCP.

Synthesis of costs and benefits
The costs and benefits were combined in the form of incremental cost-effectiveness ratios (ICERs).

Each true-positive diagnosis made with ERCP alone, which would not have been made with EUS-ERCP or MRCP-ERCP, added incremental costs of $4,875 to EUS-ERCP and of $2,580 to MRCP-ERCP.

Each false-positive diagnosis averted with initial ERCP, which would have been made with EUS-ERCP or MRCP-ERCP, added incremental costs of $9,750 to EUS-ERCP and of $1,548 to MRCP-ERCP.

Per patient, MRCP-ERCP cost $3 more to perform than EUS-ERCP. With a BD prevalence of 0.29 (the base-case), EUS-ERCP demonstrated more lesions than MRCP-ERCP and was less costly. Therefore, it dominated MRCP-ERCP.

Sensitivity analysis showed that MRCP-ERCP was less costly when EUS test performance diminished, or when improvements in MRCP performance decreased the cost of MRCP-ERCP. When the sensitivity of MRCP exceeded 0.75, MRCP-ERCP dominated EUS-ERCP. The tornado analysis demonstrated that the most influential variables were disease prevalence and the cost of the procedure.

In general, as the sensitivity and specificity of MRCP increased, it became less cost-effective to perform EUS first.

In relation to provider expertise, at a base-case disease prevalence of 0.29, the ICER for MRCP-ERCP compared with ERCP alone was $2,580 for reader A (prior reader), $3,300 for reader B (first reviewer), and $5,563 for reader C (second reviewer).

Sensitivity analyses for provider expertise showed that, as prevalence increased, the ICER decreased per additional correct diagnosis detected and it became more cost-effective to perform ERCP rather than MRCP first, regardless of the MRCP reader. In addition, initial ERCP was more cost-effective than initial EUS. At a prevalence greater than 0.32, the use of ERCP would cost less than $5,000 per additional correct diagnosis, regardless of the alternative strategy.

A comparison of MRCP-ERCP with EUS-ERCP showed that as MRCP performance improved across readers, the cost of performing EUS-ERCP increased compared with that of performing MRCP-ERCP. In fact, increasing provider expertise shifts preference from EUS-ERCP to MRCP-ERCP at a baseline prevalence of 0.29: for reader A, EUS-ERCP dominated MRCP-ERCP, while for readers B and C, MRCP-ERCP was the less costly diagnostic strategy.

Authors' conclusions
Initial magnetic resonance cholangiopancreatography (MRCP) and initial endoscopic ultrasonography (EUS) appear to have been cost-effective alternatives to initial endoscopic retrograde cholangiopancreatography (ERCP) for the evaluation of biliary abnormalities in a low-prevalence population. At very low prevalence levels, the incremental cost-effectiveness of initial MRCP compared with initial EUS is highly dependent on provider expertise. At extremely high levels, ERCP may be an economically viable initial test despite a moderately increased cost and risk of complications. At intermediate prevalence levels, procedural costs may be the determining factor. An understanding of the influence of these variables on cost-effectiveness may aid health care providers in deciding which diagnostic strategies will work best in their clinical practice.

CRD COMMENTARY - Selection of comparators
The authors gave a justification for the comparators. ERCP, the standard method for the evaluation of pancreaticobiliary disease, is moderately invasive and carries a risk of complications, including pancreatitis and cholangitis. Less invasive alternatives, such as EUS and MRCP, provide similar diagnostic capability and research studies have identified economic reasons for incorporating EUS and MRCP into routine clinical practice. You should judge whether these strategies are relevant in your own setting, or whether other procedures could also be relevant.
Validity of estimate of measure of effectiveness
The authors did not state that a systematic review of the literature had been undertaken. Although this is a common practice with models, it does not always ensure that the best data available are used. The authors used data from the available studies selectively, and used a convenience sample with a small number of patients for the single study. One cannot be sure that all of the relevant literature was identified, although the estimates of effectiveness were derived credibly from the studies identified. The effectiveness evidence for the review and modelling were derived from prospective clinical trials, which are an adequate source to estimate effectiveness. In addition, reader-specific MRCP characteristics derived from an analysis of inter-observer variability were used to test the impact of provider expertise. The authors justified their assumptions with reference to the medical literature. The estimates were investigated in sensitivity analyses using ranges derived from confidence intervals and medical literature.

Validity of estimate of measure of benefit
The authors performed the analysis in terms of the false-positive diagnoses averted. The principal input parameters for the model were derived from published medical literature and the single study they performed.

Validity of estimate of costs
The authors reported that the costs were estimated from the perspective of a third-party payer. Therefore, they used medical costs for the study based on a reimbursement schedule, and indirect costs were appropriately not included. The costs of hospitalisation resulting from procedural complications were included. However, the costs of complicated pancreatitis and cholangitis were not included, and this could have affected the authors' conclusions. The costs and the quantities were not reported separately, which hinders the generalisability of the results. Although no statistical analysis of the costs was reported, sensitivity analyses were conducted to assess the robustness of the estimates used. Discounting was, appropriately, not carried out since the time horizon of the model did not exceed 2 years. The price year was reported, which will aid any future reflation exercise.

Other issues
The authors made appropriate comparisons of their findings with those from other studies. However, they did not explicitly address the generalisability of the results to other settings. The authors’ conclusions reflected the scope of the analysis.

The authors acknowledged two main limitations to the study. First, the inherent level of uncertainty in all decision models, which was addressed through the sensitivity analyses. Second, the analysis assumed a payer's perspective and only included procedural direct costs. This is likely to have provided a conservative estimate of the incremental cost-effectiveness of alternative diagnostic strategies, and instituted a bias toward ERCP. In addition, the incorporation of indirect costs might also have biased the model toward MRCP as an alternative strategy. For example, accounting for lost revenue from prolonged recovery times with EUS, not present with MRCP, might shift preference toward MRCP.

The authors attempted to overcome acknowledged barriers to implementing the results of cost-effectiveness analyses in clinical practice. The societal perspective in which the outcome is the incremental cost per quality-adjusted life-year gained, although providing ease of comparison across a variety of interventions might not accurately reflect the perspective of a clinical practice. A clinical practice might have a shorter time horizon, or might view quality-adjusted life-years as an abstract and controversial end point, preferring instead a cost-effectiveness analysis conducted with intermediate end points such as true-positive diagnoses made and false-positive diagnoses averted.

Implications of the study
An understanding of the influence of disease prevalence, provider expertise, complications, and variable costs on cost-effectiveness might aid health care providers in deciding which diagnostic strategies will work best in their clinical practice. Cost-savings accruing from averted ERCP might be enhanced by patient pre-selection for initial MRCP, with ERCP being reserved for patients who have an increased probability of choledocholithiasis.
Source of funding
Supported by the General Clinical Research Center of the University of Michigan, the Robert Wood Johnson Clinical Scholars Program, and a GE-AUR Research Fellowship.

Bibliographic details

PubMedID
12809415

Other publications of related interest


Indexing Status
Subject indexing assigned by NLM

MeSH
Biliary Tract Diseases /diagnosis /economics /epidemiology; Cholangiopancreatography, Endoscopic Retrograde /economics; Cost-Benefit Analysis /economics; Decision Support Techniques; Endosonography /economics; False Negative Reactions; Follow-Up Studies; Health Personnel /economics; Humans; Magnetic Resonance Imaging /economics; Observer Variation; Prevalence; Professional Competence /economics; Prospective Studies; Sensitivity and Specificity

AccessionNumber
22003009472

Date bibliographic record published
31/08/2005

Date abstract record published
31/08/2005