ABSTRACT

Objective To compare the cost effectiveness of nurses and doctors in performing upper gastrointestinal endoscopy and flexible sigmoidoscopy.

Design As part of a pragmatic randomised trial, the economic analysis calculated incremental cost effectiveness ratios, and generated cost effectiveness acceptability curves to address uncertainty.

Setting 23 hospitals in the United Kingdom.

Participants 67 doctors and 30 nurses, with a total of 1888 patients, from July 2002 to June 2003.

Intervention Diagnostic upper gastrointestinal endoscopy and flexible sigmoidoscopy carried out by doctors or nurses.

Main outcome measure Estimated health gains in QALYs measured with EQ-5D. Probability of cost effectiveness over a range of decision makers’ willingness to pay for an additional quality adjusted life year (QALY).

Results Although differences did not reach traditional levels of significance, patients in the doctor group gained 0.015 QALYs more than those in the nurse group, at an increased cost of about £56 (£59, $78) per patient. This yields an incremental cost effectiveness ratio of £3660 (£3876, $5097) per QALY. Though there is uncertainty around these results, doctors are probably more cost effective than nurses for plausible values of a QALY.

Conclusions Though upper gastrointestinal endoscopies and flexible sigmoidoscopies carried out by doctors cost slightly more than those by nurses and improved health outcomes only slightly, our analysis favours endoscopies by doctors. For plausible values of decision makers’ willingness to pay for an extra QALY, endoscopy delivered by nurses is unlikely to be cost effective compared with endoscopy delivered by doctors.

Trial registration International standard RCT 82765705

INTRODUCTION

Gastrointestinal endoscopy is a common clinical procedure, and its use is increasing over time. To meet the increasing demand, endoscopy is becoming widely practised by nurses in the United Kingdom. There has been little evaluation of the cost effectiveness of procedures undertaken by nurses rather than by doctors.

Consideration of the economics of diagnostic procedures can be complex as the cost effectiveness of the consequent treatment of any discovered condition needs to be considered. Economic evaluations of screening tests often estimate a “cost per condition detected,” which is determined partly by the sensitivity and specificity of the test. We focused not on the cost effectiveness of endoscopy itself but on whether or not there is a difference in endoscopy delivered by doctors or nurses. We took a pragmatic approach to the evaluation of this complex intervention, in which we assumed that it is the method of delivery (nurse or doctor) that is under consideration, not the intervention itself. We assessed relative cost effectiveness as part of a pragmatic randomised controlled trial undertaken in the UK.

METHODS

Study design and interventions

In the cost effectiveness analysis we addressed two issues: the decision about who should undertake the intervention and the level of uncertainty around this decision. The incremental cost effectiveness ratio represents the relation between costs and outcomes and thus facilitates decision making. We characterised the uncertainty associated with this decision by estimating the probability that nurse delivered endoscopy is cost effective over a range of values of decision makers’ willingness to pay for an additional quality adjusted life year (QALY).

The clinical study, of which this economic evaluation was part, was a pragmatic randomised trial in 23 hospitals in England, Scotland, and Wales. A total of 1888 patients were allocated at random to either a doctor or a nurse for upper gastrointestinal endoscopy or flexible sigmoidoscopy. We collected health outcome measures at baseline, one day, one month, and one year after the intervention. Further details of the
trial conduct, and patient sample and characteristics, are described elsewhere.\textsuperscript{3} We take a UK National Health Service (NHS) perspective with effects assessed in terms of health gains measured in QALYs. As the time horizon of the study was one year, discounting was not appropriate. We did not extrapolate beyond one year as the study was not powered to detect differences between groups in factors influencing long term health outcomes. We used Bayesian analysis in which the parameters have probability distributions. Thus it was possible, and appropriate, to compute the probability of an event being effective or cost effective.

Data collection and outcome measures
We extracted information on resources used during endoscopy of trial patients from resource time sheets. Information collected included duration of endoscopy, number of patients undergoing endoscopy, staffing, and consumables used (except for therapeutic procedures). The duration of procedures was timed from the extubation of one patient to the extubation of the next.

We obtained data on resource use after the endoscopy from examination of patients’ medical records and patients’ questionnaires administered at baseline and 12 months (table 1). We estimated the cost of the intervention from data on the duration of intervention from the clinical trial multiplied by figures for the cost per minute (for doctor or nurse) estimated from Netten and Curtis.\textsuperscript{7} Table 2 shows unit cost estimates in 2002-3 prices.

We used the EQ-5D instrument\textsuperscript{8} to measure patients’ health states and to ascribe values to those states. This instrument measures health across five dimensions (mobility, self care, usual activities, pain and discomfort, and anxiety and depression). Patients have three possible responses (no problems, moderate problems, or severe problems) for each of these dimensions to reflect their perceived their health states. This puts each into one of 245 mutually exclusive health states (the 243 states arising from the EQ-5D, plus unconscious, and dead), each valued on a scale from 0 (representing death) to 1 (representing perfect health).

Table 1 | Details of resource use data and unit costs

<table>
<thead>
<tr>
<th>Item of resource use</th>
<th>Source of resource use</th>
<th>Source of unit cost data</th>
</tr>
</thead>
<tbody>
<tr>
<td>NHS staff and time</td>
<td>Resource time sheet recorded during endoscopy lists</td>
<td>UK NHS salary scales</td>
</tr>
<tr>
<td>inpatient stay</td>
<td>Hospital medical records examined at one year</td>
<td>UK NHS reference costs</td>
</tr>
<tr>
<td>Outpatient appointments</td>
<td>Patient questionnaires plus hospital medical records and primary care records at one year</td>
<td>UK NHS reference costs</td>
</tr>
<tr>
<td>GP visits</td>
<td>Patient questionnaires plus hospital medical records and primary care records at one year</td>
<td>Unit costs of health and social care\textsuperscript{4}</td>
</tr>
<tr>
<td>Medical management, such as drugs</td>
<td>Patient questionnaires and GP questionnaire</td>
<td>British National Formulary</td>
</tr>
<tr>
<td>Travel to and from appointments</td>
<td>Patient questionnaires at baseline, one month, and one year</td>
<td>AA</td>
</tr>
<tr>
<td>Private medical care</td>
<td>Patient questionnaires at baseline, one month, and one year</td>
<td>Unit costs of health and social care\textsuperscript{5}</td>
</tr>
</tbody>
</table>

AA=Automobile Association.

We converted all EQ-5D scores to “utilities” through a tariff derived from a representative UK population sample.\textsuperscript{4} We compared mean QALYs generated in the two groups over the 12 month period. We plotted utility at baseline and subsequent points and calculated the area under the curve to estimate QALYs gained (or lost) by each patient. This reflects the fact that the QALY is the product of time and utility\textsuperscript{10} and assumes that changes in utility over time follow a linear path.\textsuperscript{11} We adjusted these estimates for baseline EQ-5D as recommended by Manca et al\textsuperscript{12} and included sex and age as covariates. In a subgroup analysis we separately considered the cost effectiveness ofsigmoidoscopy and oesophagastroduodenoscopy.

Analysis
Resource use and EQ-5D data were missing in several patients, with some missing both. For analysis we assumed that data were missing at random. We used two methods to impute missing data. For EQ-5D, we used the last value carried forward, as patients’ last score is likely to be the best predictor of the missing value. For resource use we used regression to impute missing values from age, sex, and EQ-5D scores.

Traditionally, cost effectiveness analyses estimate incremental cost effectiveness ratios from mean differences in costs and effects between treatment and control groups, and 95\% confidence intervals from the independent samples t test. As interpretation of incremental cost effectiveness ratios derived from more than one quadrant of the cost effectiveness plane is problematic,\textsuperscript{13} we calculated net monetary benefit\textsuperscript{14} for each group from trial and imputed data.\textsuperscript{15} To calculate patient specific net monetary benefits we multiplied each patient’s QALYs by the assumed maximum value of a QALY and subtracted that patient’s costs. We used these patient specific net monetary benefits to derive cost effectiveness acceptability curves and estimated the aggregate benefit from the equation:

\[
\text{Probability cost effective} = \frac{\text{Max utility} - \text{utility}}{\text{Max utility}}
\]

where \(\lambda\) is the decision maker’s maximum willingness to pay for a QALY. For example, if treatment A has a mean cost of £100 000 and generates a mean of five QALYs with a QALY valued at £30 000, then the net
monetary benefit associated with treatment A is £50 000 (5×£10 000−£100 000).

Thus the net monetary benefit depends on the value of the QALY, and analysis shows how sensitive the results are to changes in this value. The uncertainty around the net monetary benefit can estimate the probability that a strategy is cost effective through the cost effectiveness acceptability curve. This is a graphical representation of the probability of an intervention being cost effective over a monetary range for a decision maker’s willingness to pay for an additional unit of health gain. We used the values 0 (which assumes outcomes are equivalent or not valued), £1000, £10 000, £20 000, £30 000, and £50 000 as examples of the decision maker’s willingness to pay for one extra QALY.

RESULTS

Missing data
At baseline utility data were missing for 184 patients (10%), either because they did not complete the questionnaire or because one of the items within the EQ-5D was missing. This figure increased to 489 patients (26%) at one month and 576 (31%) at one year. For data on resource use we examined the medical records of 1674 patients (89%); 606 patients (32%) did not report number of attendances in general practice. There were no significant differences in characteristics of patients with and without missing data.

Resource use
Table 3 presents mean levels of resource use. These estimates arise from responders to questionnaires without imputation. For most variables, the nurse based programme increased resource use (table 3). Endoscopy by nurses was followed by slightly more use of all primary care resources except home visits from general practitioners. In secondary care there were increased attendances at day hospital and outpatient clinics. These differences were small, however, and did not reach conventional levels of significance.

Health states
There was little effect in either group on usual activities or self care. Both groups showed an increased proportion of patients in the least severe pain and discomfort and anxiety and depression groups, and these differences favoured endoscopy delivered by doctors (table 4). Mobility deteriorated in both groups, with the nurse group again performing slightly less well than the doctor group. These results were consistent with the clinical effectiveness results as measured by the condition specific gastrointestinal symptom rating questionnaire. Three of the four factors measured at one year favoured the doctor group, though none of these differences reached significance.

QALYs
From these estimates, which included values imputed for missing values, we estimated changes in EQ-5D over one year (table 5). In turn we used these estimates to generate QALYs.

Though differences were small, the gain in QALYs was greater after endoscopy by doctors than by nurses. This is partly explained by difference in baseline characteristics, notably the EQ-5D score at baseline. Adjustment for these reduced the difference in QALYs to 0.0153 (95% confidence interval −0.008 to 0.039) (table 5). We used this estimate in the construction of cost effectiveness acceptability curves, which reflect the finding that EQ-5D score at baseline was higher for the doctor group.

While this difference in QALYs seems to be small in absolute terms, it equates to a difference of five to six days of additional perfect health each year. There are several plausible explanations for the improved QALY scores for doctors. The most likely scenario is that nurses requested more subsequent tests and investigations. Such tests and investigations might have a negative effect on patients’ wellbeing and it is possible that this manifested itself in lower QALY scores in the nurse group.

Table 3 | Mean resource use in two groups over 12 month study period

<table>
<thead>
<tr>
<th></th>
<th>Doctor group</th>
<th>Nurse group</th>
<th>Mean difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>General practitioner:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home visits</td>
<td>0.30</td>
<td>0.17</td>
<td>0.13 (−0.04 to 0.29)</td>
</tr>
<tr>
<td>Surgery visits</td>
<td>4.81</td>
<td>5.05</td>
<td>−0.24 (−0.72 to 0.24)</td>
</tr>
<tr>
<td>Practice nurse:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgery visits</td>
<td>1.43</td>
<td>1.56</td>
<td>−0.13 (−0.47 to 0.22)</td>
</tr>
<tr>
<td>Home visits</td>
<td>0.03</td>
<td>0.11</td>
<td>−0.08 (−0.22 to 0.05)</td>
</tr>
<tr>
<td>Day hospital attendances</td>
<td>0.27</td>
<td>0.35</td>
<td>−0.08 (−0.14 to 0.01)</td>
</tr>
<tr>
<td>Inpatient length of stay</td>
<td>1.10</td>
<td>1.10</td>
<td>0.00 (−0.53 to 0.53)</td>
</tr>
<tr>
<td>Outpatient attendances</td>
<td>1.34</td>
<td>1.46</td>
<td>−0.13 (−0.32 to 0.06)</td>
</tr>
<tr>
<td>Intervention time (minutes)</td>
<td>23.50</td>
<td>22.41</td>
<td>1.09 (−1.95 to 4.13)</td>
</tr>
</tbody>
</table>
There was little difference in the duration of endoscopy between groups. In addition patients allocated to doctors had slightly higher costs in both primary and secondary care. The differences in costs associated with resource use in primary and secondary care was small. Though in absolute terms there were (slightly) more contacts in both primary and secondary care in the nurse group than in the doctor group, the doctor group was associated with more costly resources, particularly salary costs. Adjustment for total costs for baseline differences in age and sex made no meaningful difference to the results.

Incremental cost effectiveness ratio (ICER)
The doctor group generated 0.0153 more QALYs than the nurse group, at a net cost of £56 per patient. This resulted in an incremental cost effectiveness ratio of £3660 per QALY. There was much uncertainty around these results, and neither the difference in patients’ outcomes nor that in costs approached traditional levels of significance. Therefore we used the net monetary benefit approach and generated cost effectiveness acceptability curves.

Net monetary benefits and cost effectiveness acceptability curves
Figure 1 shows the cost effectiveness acceptability curve for values of a QALY between zero and £50 000. Attaching no value to a QALY yields a probability of about 78% of the nurse group being cost effective, implying a chance of 78% that nurses reduced costs. The probability of nurses being cost effective, however, decreases as the value of a QALY increases and as doctors become more cost effective. At a value of £30 000 per QALY, often stated to be the borderline for the NHS, nurses have only a 13% chance of being cost effective. Indeed, for all plausible values of a QALY, doctors are more likely to be cost effective than nurses. There is, however, much uncertainty around this result; the cost effectiveness scatter in figure 2 shows the plots of incremental costs and incremental effects for doctors compared with nurses.

Sensitivity analysis
Though this stochastic analysis addresses intrinsic uncertainty, sensitivity analysis is still appropriate to check on methodological uncertainty, specifically uncertainty arising from data imputation and from different forms of endoscopy.

We had complete medical records and EQ-5D data for 440 patients, 227 (51.6%) in the nurse group and 213 (48.4%) in the doctor group, reflecting the distribution of patients between groups in the full trial. Most missing data omitted a single cost or a single EQ-5D dimension at baseline or follow-up. The resulting estimates were similar to those with imputed data: doctors cost £40 more than nurses (−£148 to £231) but generated 0.021 more QALYs (−0.02 to 0.06). This yields an incremental cost effectiveness ratio of £2060, lower than with imputed data. The smaller sample, however, leads to more uncertainty: the probability of doctors being cost effective never exceeds 84%, whatever the value of a QALY.

Sigmoidoscopy v oesophagogastroduodenoscopy
Sigmoidoscopy showed an incremental cost effectiveness ratio of £2600 per QALY for doctors.
Oesophagogastroduodenoscopy cost relatively more for doctors, resulting in a higher ratio of £7850. Both ratios, however, would be acceptable for most reasonable values of a decision maker’s willingness to pay for a QALY.

**DISCUSSION**

**Principal findings**

Patients undergoing endoscopy carried out by doctors gained 0.015 QALYs more than those treated by nurses, at an increased cost of around £56 per patient, yielding an incremental cost effectiveness ratio of £3660 per QALY. Doctor delivered endoscopy would therefore seem to be acceptably cost effective. The analysis also suggests that, for most reasonable values of a decision maker’s willingness to pay for an additional QALY, endoscopy delivered by nurses is unlikely to be cost effective compared with endoscopy delivered by doctors based on the evidence of this single trial analysis. There is considerable uncertainty around these estimates, which indicates that further research might be needed.

**Strengths and weaknesses**

We used a randomised trial to compare the cost effectiveness of doctors and nurses performing endoscopy. There were missing data both on resource use and patients’ utility. While imputing this data is not ideal, the results of that imputation are robust, as analysis limited to complete cases yields similar results. The study lasted only one year, though there is potential for later effects in this population—for example, missed diagnoses. A longer trial would be ideal, but the similarity of immediate and delayed complications between nurses and doctors suggests there is little difference in their long term performance.

It is possible that the use of the EQ-5D and the resulting estimates of QALYs are not sensitive enough in these patients to identify differences in their health related quality of life. The results of our economic analysis, however, are similar to those of the clinical analysis in that there was a non-significant effect in favour of doctors.

Hence the QALY gain was greater one year after endoscopy by doctors than by nurses. This is despite patients being more satisfied with endoscopies performed by nurses. This does not imply inconsistency, as short term gains in satisfaction probably reflect differences in process, while medium term gains in quality of life reflect patients’ perception of outcomes, perhaps in the form of the accuracy of the procedure and their confidence in the results.

Nurse endoscopists might not have reached “steady state” in experience and confidence. As their experience grows, they might become more confident and therefore order fewer follow-up tests. The higher frequency of tests and interventions in the nurse group, however, might reflect intrinsic differences between the professions in terms of attitudes to risk.

**Meaning of the study**

Though there is debate over the appropriate NHS threshold cost per QALY, Rawlins and Culyer argue that the National Institute for Health and Clinical Excellence would be unlikely to reject a technology with a cost of between £5000 and £15 000 per QALY on grounds of cost effectiveness. On the evidence of this trial, therefore, doctor delivered endoscopy seems cost effective.

This result might surprise some as upper gastrointestinal endoscopies and sigmoidoscopies by nurses cost slightly less than doctors and the difference in health outcomes is small and does not reach traditional levels of significance. Our economic analysis estimates the probability of cost effectiveness from the uncertainty around the estimates of costs and effects, rather than discarding differences that do not reach “significance.” Hence this methodological paradigm leads to a different interpretation of our results from that adopted in the clinical effectiveness paper.

Interpretation depends on the paradigm chosen and the factors under consideration by decision makers. Classic statistical inference fails to reject the null hypotheses that there is no difference in effectiveness or cost effectiveness between endoscopy delivered by doctors and nurses. Policy makers might therefore view nurse endoscopists as an acceptably safe and effective way of changing skill mix in health care, releasing medical resources and increasing the role of nurse specialists. In contrast with the classic statistical approach, economic inference in this context makes decisions by comparing the estimated cost per QALY

---

### Table 5: Mean EQ-5D scores, QALYs, and costs (£) per patient over 12 months by group

<table>
<thead>
<tr>
<th></th>
<th>Mean EQ-5D scores:</th>
<th>QALYs</th>
<th>(£) per patient over 12 months</th>
<th>Difference (95% CI)</th>
<th>%Difference in QALYs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Doctor group (n=931)</td>
<td>Nurse group (n=957)</td>
<td>Baseline</td>
<td>One month</td>
<td>One year</td>
</tr>
<tr>
<td>Mean EQ-5D scores:</td>
<td>0.700</td>
<td>0.689</td>
<td>0.011 (&lt;0.014 to 0.040)</td>
<td>0.713</td>
<td>0.697</td>
</tr>
<tr>
<td>Baseline</td>
<td>135</td>
<td>128</td>
<td>7 (&lt;3 to 15)</td>
<td>565</td>
<td>538</td>
</tr>
<tr>
<td>Total cost</td>
<td>739</td>
<td>683</td>
<td>56 (&lt;100 to 213)</td>
<td>957</td>
<td>931</td>
</tr>
</tbody>
</table>

*Difference in QALYs allows for baseline differences in EQ-5D, sex, and age.*
with a threshold equal to the most that a decision maker would pay for a QALY. In this trial this leads to the conclusion that upper gastrointestinal endoscopy and sigmoidoscopy delivered by doctors would be cost effective at a typical threshold. Bayesian analysis goes further by estimating the probability that the intervention is cost effective in the sense that the estimated cost per QALY exceeds a given threshold. In this trial this form of analysis leads to the conclusion that the average doctor endoscopist has a probability of 80-90% of being more cost effective than the average nurse endoscopist at commonly used values of willingness to pay for a QALY. Policy makers or hospital decision makers pursuing efficiency alone would therefore choose endoscopy delivered by doctors.

Unanswered questions

The choice of skill mix in endoscopy might be influenced by factors other than cost effectiveness, such as affordability, staff shortages, and access to health care, all of which enter into policy decisions. At the start of this trial there was concern about shortages of medical staff but, after the expansion of medical schools, concerns shifted to surpluses and potential unemployment of junior doctors. Endoscopy delivered by nurses, in the current state of their training and experience, is unlikely to be cost effective compared with endoscopy delivered by doctors. As nurses grow in experience over time it will be important to continue to monitor both effectiveness and cost effectiveness.

Contributors: GR was responsible for the design, conduct, and interpretation of the analysis, drafted the paper, and is guarantor. KB contributed to analysis, interpretation, and drafting the paper. JW led the trial team and was principal author of the final report. IR contributed to the design and implementation of the trial and drafting of the paper. DO developed, validated, and applied the method of assessing the upper GI endoscopy video recordings, validated the gastrointestinal symptom rating questionnaire, collected and analysed clinical data, notably from hospital records. WYC validated the gastrointestinal symptom rating and gastrointestinal endoscopy satisfaction questionnaires. AF contributed to the overall design and implementation of the trial and drafting of the paper. SC developed, validated, and managed the main database. All authors reviewed successive drafts of both papers.

Funding: The study was funded by the NIHR Evaluation Trials and Studies Coordinating Centre. All researchers are independent from this source of funding.

Competing interests: None declared.

Ethical approval: The study was approved by the Welsh multicentre research ethics committee and informed consent was given by all patients.


Accepted: 15 October 2008