The cost-effectiveness of psychotherapy and paroxetine for severe irritable bowel syndrome

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
Two treatments for patients with severe irritable bowel syndrome (IBS) were examined. The treatments were psychotherapy (PSY) and antidepressant therapy with paroxetine (PAR; a selective serotonin reuptake inhibitor). PSY consisted of psychodynamic interpersonal therapy. The patients received one long (approximately 2 hours) and 7 short (45 minutes) individual sessions over 3 months from three therapists. PAR was administered orally at a dose of 20 mg/day for 3 months. After 3 months of receiving PSY or PAR, the patients returned to their general practitioner (GP) who decided the course of further management.

Type of intervention
Treatment.

Economic study type
Cost-effectiveness analysis.

Study population
The study population comprised patients who fulfilled the following inclusion criteria:

Rome I criteria for IBS satisfied;

duration of symptoms more than 6 months;

failure to respond to usual medical treatment, including antispasmodics and laxatives or antidiarrhoeal medication administered for a minimum of 3 months;

severe abdominal pain, defined as greater than 59 on a visual analogue scale (VAS);

no contraindication to either PSY or PAR;

ability to complete a study questionnaire; and

age 18 to 65 years.

Setting
The setting was primary and secondary care. The economic study was conducted in the UK.

Dates to which data relate
The effectiveness and resource use data were gathered from October 1994 to February 1999. Prices relating to 1997/1998 were used.
Source of effectiveness data
The effectiveness evidence was derived from a single study.

Link between effectiveness and cost data
The costing was conducted prospectively on the same sample of patients as that used in the effectiveness analysis.

Study sample
Power calculations were conducted using the results of another study. A sample of 85 patients in each group was required to obtain a significant degree of improvement on a VAS of abdominal pain. Of an initial group of 552 patients, 295 were excluded (235 did not meet the inclusion criteria and 60 refused to participate). Therefore, the final sample consisted of 257 patients. There were 85 patients in the PSY group, 86 in the PAR group, and 86 in the TAU group. The mean age of the patients was 39.9 (+/- 1.36) years in the PSY group, 39.4 (+/- 1.15) years in the PAR group, and 40.6 (+/- 1.33) years in the TAU group.

Study design
This was a prospective, randomised clinical trial that was carried out at 7 centres around Manchester and Sheffield. The patients were stratified by hospital and by pain severity (using the VAS) as "severe" (> 59 to < 89) or "very severe" (>= 89). Randomisation was performed in blocks of 12 patients using randomisation lists derived from a computer-generated series of random numbers. After 3 months, patients receiving PAR or PSY returned to their GP who decided the course of further management over the next year. Therefore, the length of follow-up was 15 months. Only 59 patients in the PSY group and 43 in the PAR group received all treatments. At 3 months, 11 patients in the PSY group and 12 in the PAR group were lost to follow-up, 14 (PSY) and 29 (PAR) patients, respectively, discontinued the treatment, and 7 patients in the TAU group were lost to follow-up. At 15 months of follow-up, 13 patients in the PSY and PAR groups, and 9 patients in the TAU group were lost to follow-up. The researchers who assessed the clinical outcomes were blinded to the treatment group.

Analysis of effectiveness
The analysis of the clinical outcomes was conducted on an intention to treat basis. A per protocol analysis was also performed in a secondary analysis (fully compliant subjects only). The primary outcome measures were abdominal pain, as assessed using the VAS, and health-related quality of life, which was estimated using the SF-36 physical component score (maximum score 50, representing good health). The secondary outcome measures were the number of pain days per month, the SF-36 mental component score, the SCL-90 global severity index to assess psychological distress, and the patients' global rating of IBS symptoms. The relationship between SF-36 physical component score and ability to work was analysed. Sub-group analyses were also performed. Statistical analyses were carried out to assess the impact of baseline factors on the changes in the outcome measures.

The authors stated that there were no statistically significant differences at baseline between those patients who agreed to participate and those who declined to enter the trial. Similarly, in the PAR and PSY groups, there were no statistically significant differences at baseline between those patients who completed their course of treatment and those who did not. In general, the 3 groups were comparable at baseline in terms of their demographics, clinical conditions and quality of life.

Effectiveness results
Unless specified, only the statistically significant results of the intention-to-treat analysis will be reported here.

Abdominal pain, as assessed using the VAS, improved in all groups from baseline to the 3-month and 15-month assessments. No statistically significant differences among the groups were observed. In the per protocol analysis, the change in VAS score at 3 months was -15.9 with PSY, -24.8 with PAR, and 11.4 with TAU, (p=0.027 for PAR versus TAU).
Likewise, the change in SF-36 physical component score from baseline to 3-month assessment was comparable among the three groups. However, the changes of 5.2 (+/- 1.26) in the PSY group, 5.8 (+/- 1) in the PAR group, and -0.3 (+/- 1.17) in the TAU group from baseline to the 15-month assessment were statistically significant, (p<0.001 for both PSY and PAR versus TAU).

In terms of secondary outcomes, only the changes of -8.5 (+/- 1.70) in the PAR group and -4.3 (+/- 1.2) in the TAU group in the number of pain days from baseline to the 3-month assessment were statistically significant, (p=0.014 for PAR versus TAU). However, this significant difference disappeared at the 15-month assessment. In the per protocol analysis, the change in the number of pain days at 3 months was -9.1 with PSY, -10 with PAR, and -4.3 with TAU, (p=0.007 for PAR versus TAU).

Both PSY and PAR led to improvements in the change of SF-36 mental component score from baseline to 3-month assessment. The changes were 5.4 (+/- 1.68) for PSY and 3.7 (+/- 1.75) for PAR versus -0.5 (+/- 1.60) for TAU, (p=0.007 for PAR and PSY versus TAU). Again, the difference ceased to be significant at the 15-month assessment.

Only the changes of -0.24 (+/- 0.06) in the PAR group and 0.02 (+/- 0.06) in the TAU group in the SCL-90 global severity index from baseline to the 3-month assessment were statistically significant, (p=0.021 for PAR versus TAU). However, this difference was not statistically significant at the 15-month assessment.

At the 15-month assessment, the differences in the patients' global rating did not reach statistical significance.

There were no statistically significant differences in any of the primary measures when sub-group analyses were conducted (men and women analysed separately; severe and very severe pain groups analysed separately). However, this was presumably due to the small sample size.

The results at 15 months of the per protocol analysis were comparable with those reported above, which referred to the intention to treat analysis.

The analysis also showed that patients who were unemployed due to ill health had significantly lower SF-36 physical component scores than either those who were unemployed for other reasons, or those who were not unemployed.

The statistical analyses showed that, after allowing for the effect of age, gender, baseline SF-36 physical component score and change in SCL-90 global severity index, the differences between the three groups in terms of the improvement in SF-36 physical component score after 15 months of follow-up remained highly significant. The adjusted means were 5.53 (+/- 1.10) for PSY, 5.00 (+/- 1.09) for PAR, and 0.29 (+/- 1.09) for TAU, (p=0.001).

Finally, the multivariate analysis showed that the predictors of change in SF-36 physical component score were a reported history of sexual abuse, unemployment, and baseline SF-36 physical component for PSY, and baseline SF-36 physical component and unemployment for PAR.

**Clinical conclusions**
The effectiveness analysis showed that PSY and PAR showed no significant difference from TAU in terms of reduction in abdominal pain. However, both were significantly superior in terms of health-related quality of life.

**Measure of benefits used in the economic analysis**
The health outcomes were left disaggregated and no summary benefit measure was used in the economic analysis. In effect, a cost-consequences analysis was carried out.

**Direct costs**
An annual discount rate of 6% was applied to all costs incurred after the first year. The unit costs were not presented separately from the quantities of resources used. The health services included in the economic evaluation were study interventions, hospital services and non-hospital services. Hospital services covered inpatient days, outpatient, day-patient, and accident and emergency department attendances. Non-hospital services covered primary care contacts,
domiciliary care services, and day centres. Direct non-medical services (travel costs and additional patient expenditures related to other medications, housework, child care, and so on) were also included. The cost/resource boundary of the NHS was adopted. The costs were derived from relevant local service providers or national data sources. Resource use was estimated using actual patient-level data based on case notes and interviews derived from the sample of patients who were included in the effectiveness study. Data were available for 97% of patients. The prices related to 1997/1998.

**Statistical analysis of costs**
Bootstrap methods were used to compare the treatment groups in terms of the costs. The results were presented as 95% bias-corrected confidence intervals (CIs).

**Indirect Costs**
The authors considered the indirect costs so as to determine the impact of productivity losses, but the source of the patients’ wage rates was not reported. The workdays lost due to illness or clinic attendance were derived from the sample of patients included in the effectiveness study. Patients on invalidity benefits were not included, but the number of individuals receiving benefits at the beginning of the trial was recorded. Other information on the analysis of the indirect costs was not provided.

**Currency**
US dollars ($). The costs were assessed in UK pounds sterling (£) then converted to US dollars. The conversion rate was 1 = $1.6.

**Sensitivity analysis**
One-way and multi-way sensitivity analyses were performed to assess the impact of variations in the unit costs on the estimated total costs.

**Estimated benefits used in the economic analysis**
See the 'Effectiveness Results' section.

**Cost results**
From baseline to 3 months, the mean weekly direct health care costs were $36.67 (+/- 53.14) in the PSY group, $29.79 (+/- 24.82) in the PAR group, and $25.82 (+/- 37.01) in the TAU group. The bias-corrected 95% CI for the intervention against TAU was $1.41 - $17.60 for PSY and -$7.29 - $4.64 for PAR.

From 3 months to 15 months, the mean weekly direct health care costs were $18.77 (+/- 18.93) in the PSY group, $24.08 (+/- 31.07) in the PAR group, and $31.98 (+/- 51.10) in the TAU group. The bias-corrected 95% CI for the intervention against TAU was -$9.34 - $0.05 for PSY and -$6.39 - $4.50 for PAR.

From baseline to 15 months, the mean weekly direct health care costs were $23.12 (+/- 16.21) in the PSY group, $25.38 (+/- 25.84) in the PAR group, and $30.29 (+/- 52.10) in the TAU group. The bias-corrected 95% CI for the intervention against TAU was -$5.39 - $2.08 for PSY and -$5.29 - $3.58 for PAR. Variations in the unit costs did not affect the estimated costs.

The estimated non-medical and productivity costs were not reported.

The mean annual direct health care costs were $976 (+/- 984) for PSY, $1,252 (+/- 1,616) for PAR and $1,663 (+/- 3,177) for TAU. This difference was statistically significant for PSY compared with TAU, but not for PAR compared with TAU.
Synthesis of costs and benefits
A synthesis of the costs and benefits was not relevant as a cost-consequence analysis was carried out.

Authors' conclusions
From the perspective of the National Health Service (NHS), both psychodynamic therapy (PSY) and paroxetine (PAR) therapy led to improvements in health-related quality of life in patients with severe irritable bowel syndrome (IBS), without increasing the overall health care costs. PSY resulted in statistically significantly lower direct costs than treatment as usual (TAU). However, long-term improvements in abdominal pain were not observed.

CRD COMMENTARY - Selection of comparators
The choice of the comparator was appropriate since it reflected the traditional approach for treating patients with IBS in the UK. The two interventions under examination represented two alternative and more intensive strategies for patients with IBS. Also, the authors underlined the fact that the choice of a selective serotonin reuptake inhibitor, rather than a tricyclic antidepressant, was due to reasons of compliance. You should decide whether they are valid comparators in your own setting.

Validity of estimate of measure of effectiveness
The analysis of effectiveness was based on a clinical trial, which was appropriate for the study question. The internal validity of the study was further enhanced by several factors. First, the adequate sample size, determined using power calculations. Second, intention to treat analysis (although a treatment completers only analysis was also carried out). Third, the methods of sample selection and randomisation were described. Fourth, the blinded assessment of the outcomes. Fifth, the baseline comparability of the study groups. Sixth, the lack of statistically significant differences between patients who remained in the study and those who were lost to follow-up, or those who refused to participate. Finally, the use of multi-regression analyses to determine the impact of baseline factors on the estimated outcomes. Therefore, the effectiveness evidence would appear to be robust. The study sample is likely to have been representative of the patient population. The authors acknowledged that TAU was not standardised, owing to the pragmatic design of the trial.

Validity of estimate of measure of benefit
No summary benefit measure was used in the analysis because, in effect, a cost-consequences analysis was conducted.

Validity of estimate of costs
The authors stated which perspective was adopted in the study. As such, it appears that all the relevant categories of costs have been considered in the analysis. However, the estimated direct non-medical costs and productivity costs were not explicitly reported. Further, the inclusion of the indirect costs was not reported clearly and there limited information on the methods used to estimate them. The financial year during which the costs were assessed was reported, which will facilitate reflation exercises in other settings. However, the unit costs and the quantities of resources used were not presented separately, which limits the possibility of replicating the study. The source of the data was reported. Statistical tests were appropriately used to deal with the skewed distribution of the costs. The cost estimates were varied in the sensitivity analysis, but the results were not reported clearly.

Other issues
The authors made some comparisons of their findings with those from other studies. The issue of the generalisability of the study results to other settings was explicitly addressed. The authors noted three main limitations to the transferability of their findings. First, the study results referred only to patients with severe IBS. Second, the comparator was not standardised. Third, the cost data were specific to the UK.

Implications of the study
The authors suggested that further studies, to determine which patients would respond better to each treatment or a combination of both, should be conducted.

**Source of funding**
Supported by the Medical Research Council (UK) and the North Western Region Health Authority (UK) R&D Directorate.

**Bibliographic details**

**PubMedID**
12557136

**DOI**
10.1053/gast.2003.50055

**Other publications of related interest**


**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Absenteeism; Adult; Antidepressive Agents, Second-Generation /therapeutic use; Colonic Diseases, Functional /physiopathology /therapy; Cost-Benefit Analysis; Female; Forecasting; Health Care Costs; Health Status; Humans; Male; Middle Aged; Multivariate Analysis; Pain /physiopathology; Paroxetine /economics /therapeutic use; Patient Compliance; Psychotherapy /economics; Quality of Life; Severity of Illness Index; Treatment Outcome

**AccessionNumber**
22003000376

**Date bibliographic record published**
31/03/2005

**Date abstract record published**
31/03/2005