An economic evaluation of three physiotherapy treatments for non-specific neck disorders alongside a randomized trial

M. Lewis, M. James1, E. Stokes1, J. Hill, J. Sim, E. Hay and K. Dziedzic

Objectives. Cost-effectiveness and cost-utility analyses were conducted to compare advice and exercise plus manual therapy (MT) and advice and exercise plus pulsed shortwave diathermy (PSWD) with advice and exercise alone (A&E) in the treatment of non-specific neck disorders by experienced physiotherapists.

Methods. Between July 2000 and June 2002, 350 participants with neck disorders from 15 physiotherapy departments were randomized to: A&E (n = 115); MT (n = 114) and PSWD (n = 121). Outcome and resource-use data were collected using physiotherapist case report forms and participant self-complete questionnaires. Outcome measures were the Northwick Park Neck Pain Questionnaire (NPQ) and EuroQoL EQ-5D [used to derive quality-adjusted-life-year (QALY) utility scores]. Two economic viewpoints were considered (health care and societal). Cost-effectiveness acceptability curves were used to assess the probabilities of the interventions being cost-effective at different willingness-to-pay threshold values.

Results. Mean improvement in NPQ at 6 months was 11.5 in the A&E group, 10.2 in the MT group and 10.3 in the PSWD group; mean QALY scores were 0.362, 0.342 and 0.360, respectively. Mean health care costs were £105, £119 and £123 in the A&E, MT and PSWD groups, respectively. Mean societal costs were £373, £303 and £338 in each group, respectively. Depending on the viewpoint and the outcome measure, A&E or MT were most likely to be the cost-effective interventions. PSWD was consistently the least cost-effective intervention.

Conclusions. The cost-effective intervention is likely to be A&E or MT, depending on the economic perspective and preferred outcome, but not PSWD.

KEY WORDS: Randomized clinical trial, Economic evaluation, Cost effectiveness, Cost utility, Neck pain, Physiotherapy.

Introduction

Neck complaints are common, affecting 13% of adults at any one time and up to 30% of males and 50% of females in the course of a lifetime [1, 2]. Neck pain poses a large health and economic burden to society and is associated with significant utilization of health care resources, work absenteeism and disability. A cost-of-illness study in The Netherlands estimated the total cost of neck pain in 1996 to be US $686 million [3]. Direct health care costs amounted to US $160 million (23%)—most of this direct cost was attributable to paramedical care, for example physiotherapy. Indirect health care costs (work absenteeism and disability) accounted for the remaining US $526 million (77% of the total cost). A significant proportion of health care costs are attributed to a minority of the people with neck pain, and in particular, those with higher severity of pain, more persistent symptoms and poorer prognosis [2, 4, 5].

In the UK, patients are frequently referred to physiotherapy for their assessment and management because of reduced waiting times, lower costs and potentially quicker recovery compared with referral to secondary care specialists [6]. We conducted a trial to investigate whether advice and exercise plus manual therapy (MT) or advice and exercise plus pulsed shortwave diathermy (PSWD) provided greater clinical improvement than advice and exercise alone (A&E) in the treatment of non-specific neck disorders when delivered by experienced musculoskeletal physiotherapists [7]. Whilst providing evidence of no significant difference in effect [8, 9], differences in cost are not known, and hence which treatment provides best value for money is unknown. Because of limited resources and high demands on health care, economic evaluations have an important role in decision making and health policy [10]. Policy makers need to make decisions on how to optimize the allocation of resources available to them. This article presents the findings of a cost-effectiveness and cost-utility analysis comparing MT and PSWD with A&E in the treatment of non-specific neck disorders by physiotherapists.

Methods

Participants and randomization

This was a pragmatic multi-centre randomized controlled trial. The trial recruited patients aged 18 yrs and over who were referred to physiotherapy by their general practitioner (GP) with a clinical diagnosis of a non-specific neck disorder. Patients were eligible if they had not consulted any health care professional other than the primary health care team for this problem in the previous 6 months. A total of 55 experienced senior musculoskeletal therapists from 15 physiotherapy outpatient facilities in the Midlands region of the UK were involved in recruiting and treating participants according to standardized protocols. The West Midlands Multicentre Research Ethics Committee and 10 local research ethics committees granted ethical approval for the study. Written informed consent was obtained from all participants prior to randomization. A detailed description of the methodology has been reported elsewhere [8].

A total of 350 participants were recruited to the trial between July 2000 and June 2002 [8]. The sample size was pre-determined according to the ability to detect a clinically important difference of 5 points in the primary outcome measure, the Northwick Park Neck Pain Questionnaire (NPQ). Participants were randomized to one of the three treatment groups: advice and exercise with the addition of manual therapy (MT group, n = 115); advice and exercise with the addition of pulsed shortwave
diathermy (PSWD group, n = 121). Four randomized participants were found to have contraindications to the study interventions by the treating physiotherapists, and were excluded from follow-up assessment: participant follow-up based on eligibility was therefore 113 in the A&E group, 113 in the MT group and 120 in the PSWD group. Compliance with treatment was good—98% of participants received their allocated interventions.

In accordance with the protocol, physiotherapists were instructed to deliver an initial 40 min assessment and treatment session and a maximum of six further 20 min sessions. There was a minimal requirement in all arms to deliver education and advice, and individualized exercises. Treatment details have been reported previously [8]. The content of the interventions and exact number of sessions were at the clinical discretion of the physiotherapists. GPs were instructed to avoid prescribing co-interventions during the first 6 weeks of the trial, but participants already using non-steroidal anti-inflammatory drugs were asked to continue their stable dose. Participants were advised to return to their GP if needed for further possible treatment if symptoms persisted after the 6 weeks of physiotherapy treatment.

Outcomes

Neck pain and disability was measured using the NPQ (scale: 0–100; 0 = no neck pain and disability, 100 = maximum neck pain and disability) [11]; outcome was expressed as change in scores between baseline and 6 months. Utility was measured using QALYs, derived from weighted average summary scores for the EuroQol EQ-5D [12] across the 6 months of follow-up, calculated using area-under-the-curve methods [13]. The EQ-5D is the most commonly used generic measure of health status (scale = −0.59 to 1.00, where a score of −0.59 reflects the worst possible health state and 1.00 reflects perfect health). Although less responsive to disease-specific outcomes, generic measures are broad in their measurement of health status, making it possible to compare the outcomes of a range of interventions for different diseases [14]. Outcome measures were collected by means of a participant self-complete questionnaire at baseline, and 6 weeks and 6 months post-randomization.

Economic viewpoint

Two viewpoints were considered for costs. The health care perspective was based on primary and secondary health care costs alone (UK NHS). The societal perspective addressed the broader cost issues of the condition by totalling health care costs, patient costs and costs associated with loss of productivity. This perspective was included in light of the expected high indirect costs for neck pain [3], and to take account of the financial burden to the patients.

Resource measurement

Resources relating to the trial intervention were recorded by the treating physiotherapists on participants’ case report forms. These forms included a detailed breakdown of the modalities used in the treatment sessions for the neck disorder, such as advice and exercise, specific manual therapies and details of pulsed shortwave treatment. Case report forms also included information about whether the participant did not attend (DNA) or was unable to attend (UTA) for any of the sessions. The economic analysis included session costs for DNAs but not for UTAs, as UTAs can be reallocated to other patients whereas DNAs cannot, and therefore incur a cost.

Self-reported questionnaires recorded additional treatment (other than the trial treatment sessions) specifically relating to the participants’ neck disorder. The 6-week questionnaire covered the period from baseline to 6 weeks follow-up, and the 6-month questionnaire covered the period from 6 weeks to 6 months follow-up. Questions included, for example, visits to the GP, practice nurse, osteopath, chiropractor, acupuncturist; hospital outpatient attendances including visits to a rheumatologist and further physiotherapy; inpatient stays or day case visits; details of medications (prescriptions and over-the-counter medicines) and assistive devices or appliances. A distinction was made as to whether the resources were paid for on a health care or private basis. Participants were also asked about their current work status (full-time/part-time/not working) and occupation (if applicable), and whether they had taken time off work as a result of their neck problem, and if so on how many days during the 6-month period.

Validation of the participants’ responses to the self-complete questionnaires was undertaken for 10% of the study population. The questionnaire data were cross-validated against medical record data for neck disorders related to the following resources: (i) further consultations with primary care professionals; (ii) prescribed medications; (iii) referrals to NHS secondary care health professionals; and (iv) noted sick leave. The study nurse, who was kept blind to the self-report data, conducted the medical record reviews by cross-checking electronic and paper copies of GP consultations and repeat prescribing during the 6-month period of follow-up. The average number of GP consultations self-reported was 0.52 and by GP-records was 0.70; the intra-class correlation coefficient for the absolute agreement between the data was 0.84. For the other three measures (prescriptions, referrals and sick leave), the medical record data were categorized as ‘yes’ (use of resource at least once) or ‘no’ (resource not used) due to low occurrences of measures; a similar dichotomy was used for the self-reported data in order to cross-validate the findings. Agreements between self-reported data and GP-recorded data for these three measures ranged between 74% (20/27) and 93% (25/27).

Costs

Details of unit costs for the resources used in this study are given in Table 1 [15–20]. All costs were reported at 2003/04 UK prices in pound sterling (£). Costs of PSWD were calculated to be £0.10 per session based on purchase and maintenance costs associated with the Curapuls 670 machine [16]. Prescription costs were derived from the British National Formulary [21]. For devices and appliances supplied through the NHS, costs were derived from Physio-Med [16]. Patient costs associated with over-the-counter
Economic evaluation of neck pain treatments

1703

medicines and purchases of devices and appliances were reported on the self-completed questionnaires. Where patient costs were not reported, they were taken from the British National Formulary and Physio-Med. Productivity losses were calculated using a friction-cost approach based on time off work within the friction period of 6 months follow-up. Individuals without a contract of employment for reasons of long-term ill health, disability for whatever reason were precluded from having any productivity loss; the assumption being that they would have been replaced ‘within the system’ [22]. Indirect costs were estimated by multiplying the reported number of days off work during the 6-month follow-up period by the average daily wage (differentiating between full-time and part-time status) [19].

Economic analysis

The study was a superiority trial; hence the economic evaluation was carried out as a cost-effectiveness analysis (CEA) and cost-utility analysis (CUA) [10, 14, 23]. Costs were averaged across all patients in the three treatment groups, and represent 6-month (and not annualized) costs. Differences in mean outcomes and mean costs were compared between the treatment groups: (i) MT group vs A&E group; (ii) PSWD group vs A&E group. A&E, the active comparator, was used as the reference group in the analyses. Due to the skewed nature of costs associated with health care interventions [4, 24], bias-corrected and accelerated bootstrapping (BCa), was used to obtain confidence intervals for cost estimates (5000 replications) [24–26]. Confidence intervals for differences in effect were calculated using parametric methods.

Comparison of cost and effects were illustrated graphically using cost-effectiveness planes; 5000 bootstrap cost-effect pairs were plotted to illustrate the level of uncertainty surrounding the cost-effect estimates. Each cost-effect pair lies in one of four quadrants of the plane (with the exception that some pairs may lie on the x-y axes). An intervention is considered ‘dominant’ over its comparator (or vice versa) if it was both more effective and less costly. MT/PSWD is dominant (less costly and more effective) over A&E in the S-E quadrant of the plane; A&E is dominant over MT/PSWD in the N-W quadrant. Neither MT/PSWD nor A&E are dominant in the other two quadrants: N-E and S-W quadrants [14].

Incremental cost-effectiveness ratios (ICERs) evaluate the ratio of mean differences in cost relative to mean differences in effect for one treatment group compared with another [10, 14, 23]. When no treatment is dominant, the ICER summarizes the additional cost of achieving one extra unit of outcome (point change in NPQ or QALY): a judgement is then required as to whether the extra cost is worth the extra benefit. Cost-effectiveness acceptability curves (CEACs) were plotted to quantify, from the bootstrap data, the probabilities of the interventions being cost-effective across a range of ceiling ICER values (otherwise referred to as the willingness-to-pay threshold values) [27]. The probability curves for the three interventions were plotted together to allow direct comparison of which intervention had the greatest probability of being cost-effective [28]. A willingness-to-pay cost threshold of up to £30 000 per QALY gain for the more effective treatment (or conversely, a willingness to lose 1-QALY if this means a saving of £30 000 or more) has been suggested [29]. Although CEACs were plotted for both outcome measures, there is no such cost threshold of willingness-to-pay per unit of NPQ.

Analysis was carried out using an intention-to-treat approach. The base-case approach was to analyse all participants with complete data for outcomes and costs. Linear-regression analysis was carried out adjusting for the following baseline covariates: age; gender; social class (non-manual or routine/manual, derived from the Standard Occupational Classification 2000 (SOC 2000) [20]); time off work (in the last 3 months); NPQ scores and EQ-5D scores. For the baseline social-class variable, participants who were not employed at baseline were included as a distinct category (to overcome the problem of missing data).

Statistical analysis was carried out using SPSS version 14.0 and STATA version 9.0 [30, 31].

Sensitivity analysis

Outcome and cost comparisons were repeated after imputing missing NPQ change scores, QALYs and cost data (assumed to be missing at random [32]) using multiple imputation after appropriate data transformations had been carried out (predictors included: baseline treatment group; age; gender; social class; time off work; NPQ; EQ-5D; and 6-week NPQ and EQ-5D) [33].

The process adopted for multiple imputation was data augmentation (DA): multiple imputed values are drawn from a stochastic series of cyclical iterations of the posterior distribution—in this case a posterior normal distribution—given the observed data and assumed values for the a priori parameters.

Results

Participants

The average age of participants was 51 yrs and 63% were female. Baseline demographic and clinical characteristics (neck-specific and generic) were similar between the study groups [8]. However, higher mean NPQ scores and lower mean EQ-5D scores were recorded in the MT group (38.6 and 0.630, respectively) compared with the A&E group (36.6 and 0.683) and PSWD group (36.9 and 0.672), i.e. patients in the MT group had slightly worse health status at baseline.

Outcomes

The key outcomes at 6 months are presented in Table 2. These results show that NPQ change scores and QALY values were similar for the three interventions at 6 months follow-up. Effect differences for both outcomes were not statistically significant between neither MT nor PSWD and A&E. Most patient improvement in outcome (NPQ and QALYs) occurred within the first 6 weeks in all groups, although small gains in outcome continued up to 6 months (data not shown) [8].

Resource use and costs

For the purposes of acquiring a complete cost data set, participants were included for the primary cost analysis if case report forms were recorded and resource use questionnaires were returned at both 6 weeks and 6 months follow-up. As the follow-up period was <12 months, no discounting was applied to costs. Resource information was available at both 6 weeks and 6 months for 2S8 (75%) of eligible participants (77 for A&E; 87 for MT; 94 for PSWD). Responders were more likely than non-responders to be: older (mean age 53 yrs vs 47 yrs), female (65% vs 58%), have lower NPQ scores (mean 36.3 vs 39.7) and have higher EQ-5D scores (mean 0.68 vs 0.62), respectively.

Table 3 shows the utilization of resources by the participants during the study period. There were fewer visits to the physiotherapist in the A&E group than the other two groups. None of the participants recorded any hospital inpatient stays or day case visits. Participants in the MT group had taken less time off work on average than the respondents in the other two groups.

Resource costs are presented in Table 3. Total health-care and societal costs are shown in Table 4. Health-care costs were largely attributable to the number of physiotherapy intervention visits. Mean health-care costs were lower in the A&E group (£105) than the MT group (£119) and PSWD group (£123). The main cost driver for societal costs was productivity losses. Mean societal costs were lower in the MT group (£303) than the A&E group (£373) and PSWD group (£338). Cost differences were not statistically significant between neither MT nor PSWD and A&E.
of being cost-effective for all willingness-to-pay thresholds within based on the NPQ change score, A&E had the greatest probability indicated by the wide scatter of cost-effect pairs across the four estimating the true cost-effectiveness of the interventions as Table 2. Change in outcomes after 6 months between treatment groups

<table>
<thead>
<tr>
<th></th>
<th>A&amp;E (n=110)</th>
<th>MT (n=104)</th>
<th>PSWD (n=78)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPQ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numbers analysed</td>
<td>97</td>
<td>104</td>
<td>110</td>
</tr>
<tr>
<td>Mean (s.d.) score</td>
<td>11.5 (15.7)</td>
<td>10.2 (14.1)</td>
<td>10.3 (15.0)</td>
</tr>
<tr>
<td>Mean difference (95% CI)&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted&lt;sup&gt;d&lt;/sup&gt;</td>
<td>–</td>
<td>–1.0 (–5.1, 3.1)</td>
<td>–1.7 (–5.7, 2.4)</td>
</tr>
<tr>
<td>Imputed&lt;sup&gt;d&lt;/sup&gt;</td>
<td>–</td>
<td>–0.8 (–4.8, 3.2)</td>
<td>–0.7 (–4.6, 3.3)</td>
</tr>
<tr>
<td>QALY</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numbers analysed</td>
<td>78</td>
<td>86</td>
<td>91</td>
</tr>
<tr>
<td>Mean (s.d.) score</td>
<td>0.362 (0.114)</td>
<td>0.342 (0.114)</td>
<td>0.360 (0.094)</td>
</tr>
<tr>
<td>Mean difference (95% CI)&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted&lt;sup&gt;d&lt;/sup&gt;</td>
<td>–</td>
<td>0.002</td>
<td>–0.011</td>
</tr>
<tr>
<td>Imputed&lt;sup&gt;d&lt;/sup&gt;</td>
<td>–</td>
<td>(–0.023, 0.026)</td>
<td>(–0.023, 0.026)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Numbers analysed were the number of responders to the respective outcome measures at 6 months follow-up; except for the imputation analyses, which were based on data for all 346 participants.
<sup>b</sup>Difference in mean scores calculated as mean score of MT group minus mean score of A&E group and mean score of PSWD group minus mean score of A&E group.
<sup>c</sup>Adjusted for baseline age, gender, social class, time off work, NPQ scores and EQ-5D scores.
<sup>d</sup>Based on imputed 6-month NPQ/QALY data (numbers analysed were 113 in the A&E group, 113 in the MT group and 120 in the PSWD group).

A&E, advice and exercise with no further addition to treatment; MT, advice and exercise plus manual therapy; PSWD, advice and exercise plus pulsed shortwave diathermy.

TABLE 3. Resource utilization and costs (£) over 6 months follow-up by treatment group

<table>
<thead>
<tr>
<th>Resource use</th>
<th>A&amp;E (n=77)</th>
<th>MT (n=87)</th>
<th>PSWD (n=94)</th>
<th>A&amp;E (n=77)</th>
<th>MT (n=87)</th>
<th>PSWD (n=94)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health-care resources</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean no. of intervention sessions&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4.49</td>
<td>5.79</td>
<td>6.63</td>
<td>67.81</td>
<td>83.76</td>
<td>99.72</td>
</tr>
<tr>
<td>Mean no. of General Practice consultations</td>
<td>0.39</td>
<td>0.89</td>
<td>0.72</td>
<td>7.01</td>
<td>15.10</td>
<td>13.02</td>
</tr>
<tr>
<td>Mean no. of outpatient attendances</td>
<td>0.58</td>
<td>0.69</td>
<td>0.16</td>
<td>17.78</td>
<td>11.65</td>
<td>4.49</td>
</tr>
<tr>
<td>Percentage who received prescriptions</td>
<td>25%</td>
<td>34%</td>
<td>32%</td>
<td>1.32</td>
<td>3.40</td>
<td>5.53</td>
</tr>
<tr>
<td>Percentage who received devices/appliances</td>
<td>8%</td>
<td>8%</td>
<td>4%</td>
<td>11.44</td>
<td>4.90</td>
<td>0.44</td>
</tr>
<tr>
<td>Patient resources</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean no. of Health Professional consultations&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.17</td>
<td>0.49</td>
<td>0.12</td>
<td>6.42</td>
<td>20.83</td>
<td>4.94</td>
</tr>
<tr>
<td>Percentage who purchased OTC medicines</td>
<td>44%</td>
<td>38%</td>
<td>41%</td>
<td>4.16</td>
<td>3.55</td>
<td>4.85</td>
</tr>
<tr>
<td>Percentage who purchased devices/appliances</td>
<td>36%</td>
<td>40%</td>
<td>28%</td>
<td>36.64</td>
<td>19.83</td>
<td>23.00</td>
</tr>
<tr>
<td>Mean no. of days off work&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3.5</td>
<td>2.3</td>
<td>3.1</td>
<td>220.14</td>
<td>140.29</td>
<td>182.40</td>
</tr>
</tbody>
</table>

<sup>a</sup>Includes initial 40 min combined assessment/treatment session and subsequent 20 min treatment sessions and is inclusive of DNAs and additional treatments considered to be part of the intervention.
<sup>b</sup>Patient costs for health-care professional consultations was based on the cost of the sessional time of the private consultation only and did not include travel costs.
<sup>c</sup>Costs were based on a usual 7 day working week in the UK.
<sup>d</sup>Costs were averaged across all patients in the three treatment groups, and represent 6-month (not annualized) costs.

A&E, advice and exercise with no further addition to treatment; MT, advice and exercise plus manual therapy; PSWD, advice and exercise plus pulsed shortwave diathermy.

TABLE 4. Health-care and societal costs (£) over 6 months follow up between treatment groups

<table>
<thead>
<tr>
<th></th>
<th>A&amp;E (n=77)</th>
<th>MT (n=87)</th>
<th>PSWD (n=94)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health-care costs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (95% CI)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>105.36 (78.92, 143.34)</td>
<td>118.81 (103.32, 138.33)</td>
<td>123.21 (110.76, 138.45)</td>
</tr>
<tr>
<td>Difference (95% CI)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5.97 (–37.24, 38.84)</td>
<td>17.98 (–20.12, 48.43)</td>
<td>21.17 (–6.66, 44.15)</td>
</tr>
<tr>
<td>Societal costs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (95% CI)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>372.72 (179.97, 786.33)</td>
<td>303.31 (175.94, 533.95)</td>
<td>338.40 (196.02, 623.32)</td>
</tr>
<tr>
<td>Difference (95% CI)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>–76.34 (–449.58, 232.53)</td>
<td>7.59 (–328.47, 333.67)</td>
<td>0.8 (–15.7)</td>
</tr>
</tbody>
</table>

Health-care costs relate to the total costs of health-care resources. Societal costs relate to the total costs of health-care resources plus patient resources plus productivity losses.
<sup>a</sup>Costs were averaged across all patients in the three treatment groups, and represent 6-month (not annualized) costs.
<sup>b</sup>Difference in mean scores calculated as mean score of MT group minus mean score of A&E group and mean score of PSWD group minus mean score of A&E group.
<sup>c</sup>Adjusted for baseline age, gender, social class, time off work, NPQ scores and EQ-5D scores.
<sup>d</sup>Based on imputed cost data (numbers analysed were 113 in the A&E group, 113 in the MT group and 120 in the PSWD group).

A&E, advice and exercise with no further addition to treatment; MT, advice and exercise plus manual therapy; PSWD, advice and exercise plus pulsed shortwave diathermy.

Cost-effectiveness and cost-utility analyses

**Health-care perspective.** There was much uncertainty in estimating the true cost-effectiveness of the interventions as indicated by the wide scatter of cost-effect pairs across the four quadrants of the CEA/CUA-planes (Fig. 1). Figure 2A shows that, based on the NPQ change score, A&E had the greatest probability of being cost-effective for all willingness-to-pay thresholds within the range investigated, i.e. up to £500 per unit gain in NPQ change score. In terms of QALY’s, MT had the greatest probability of being cost-effective for the majority of willingness-to-pay thresholds investigated, but at low threshold values (below £7000 per QALY), A&E had the greatest probability of being cost-effective. For QALY’s, uncertainty regarding the intervention of choice increased as the willingness-to-pay threshold increased. This was illustrated by the convergence of the probability curves (Fig. 2A).
At the £30 000 per QALY threshold the probability of each intervention being cost-effective was very similar: 0.32 for A&E, 0.37 for MT and 0.31 for PSWD.

Societal perspective. In terms of NPQ change score, A&E had the greatest probability of being cost-effective for the majority of willingness-to-pay thresholds investigated, but at low threshold values (below £50 per unit gain), MT had the greatest probability of being cost-effective (Fig. 2B). For QALYs, MT had the greatest probability of being cost-effective across all willingness-to-pay threshold values investigated. There was much uncertainty in the data in relation to an increased willingness-to-pay per unit QALY gain. At the £30 000 per QALY threshold the probability that A&E, MT and PSWD were cost-effective was 0.30, 0.44 and 0.26, respectively.

Sensitivity analyses
Mean differences in NPQ change scores, QALYs and costs were little changed after imputation of missing data (Tables 2 and 4).

Discussion
Our study showed that A&E and MT were more likely to be cost-effective than PSWD in treating patients with non-specific neck disorders. The intervention of choice differed according to the economic perspective chosen and the outcome of interest.

Economic decision making is based on a comparison of costs and effects. If an intervention is dominant (i.e. has greater effects and lower costs) then the decision is clear, the dominant treatment would be the preferred treatment. However, most evaluations are not as straightforward as this, often one intervention is more effective and more costly. There is also the added problem of uncertainty in the estimates due to variation in the data. We assessed the cost-effectiveness of the interventions through plotting CEACs. The CEAC is a decision analytic method that assesses the probabilities (taking into account the variability in the data) of interventions being cost-effective at ceiling willingness-to-pay values [27, 28].

When adopting a health-care perspective, A&E was clearly the intervention of choice in relation to the disease-specific measure (NPQ). In terms of the QALY, A&E had the highest probability of being cost-effective at low willingness-to-pay thresholds (less than £7000 per QALY); above this threshold MT had the highest probability of being cost effective, but there was much uncertainty as to which intervention was cost-effective. At the suggested health-care willingness-to-pay threshold of £30 000 per QALY [29] each intervention had approximately equal probability of being cost-effective. Manca et al. [34] reported that a brief physiotherapy intervention was more cost-effective, from a health-care perspective, than usual physiotherapy which could include manual therapy, electrotherapy and advice [34]. In our study, the average cost of health-care treatment over 6 months was estimated to be £5.97 more in the MT group than the A&E group, and £17.98 more for PSWD compared with A&E. In terms of resource
use, the MT and PSWD groups received more treatment sessions within the study protocol, reported higher numbers of general practice consultations and more prescribed medication; the key health-care cost driver was the treatment cost of the trial intervention. This difference in physiotherapy resource use may be of importance if there are constraints on the provision of physiotherapy services. Intervention costs do not reflect a true incremental cost of adding one treatment to another e.g. manual therapy or pulsed shortwave diathermy sessions, to advice and exercise sessions. The actual time spent on components of a combined treatment package was not measured and may vary in each treatment option up to the maximum sessional time; however, a minimum requirement was to deliver advice and exercise in all treatment arms.

Neck disorders are not only a significant medical problem but also a socioeconomic burden [3]. In our study, the societal perspective looked at overall combined costs taking into account health-care, patient and lost productivity costs. From this perspective, MT was the least costly intervention. On average, the reduction in societal cost through using MT in contrast to A&E was estimated to be £76.34. This difference was largely attributable to smaller productivity losses in the MT group. In terms of the NPQ, at low willingness-to-pay thresholds, MT was most likely to be the intervention of choice; at higher thresholds, A&E was most likely to be cost-effective. MT was cost-effective across all threshold values studied with respect to the QALY. At a willingness-to-pay threshold of £30 000 per QALY, the probability that MT was cost-effective was 0.44. However, we must be cautious in our interpretation of lost productivity as this estimate was crude, because our questions regarding work loss took into account only the discrete number of days lost due to patients’ neck disorder, and did not address or encompass other aspects of lost work productivity such as reduced functional capacity at work or change in roles or responsibilities attributed to their incapacity.

Although there is now a growing acknowledgement of the use of a £30 000 cost-effectiveness threshold [29] for decisions on health-care interventions using the QALY as the outcome measure, no such evidence exists for other health-care outcomes. Therefore, in terms of the NPQ, we do not know how much decision makers would be willing to pay for a one-unit increase. The NPQ is sensitive to very small changes in neck pain status. Hence, in reality, commissioners may be prepared to pay very little for a one-unit change in this score. It may be that the meaningful clinical difference in NPQ score is a 5-point change, or a 25% reduction in score from baseline [35]. Thus, a one-unit change alone may not be important. In practice, the QALY is a global measure of health outcome. Measures such as the NPQ are disease-specific instruments and as such cost-effective thresholds for such measures are not readily available or interpretable across studies.

Our cost findings can be compared with those of Korthals-de Bos et al. [36]. In our study, health-care costs were £119 for MT, £123 for PSWD and £105 for A&E over 6 months follow-up. In comparison, Korthals-de Bos reported 12-month health-care costs of 222 euros (£155) (based on the average exchange rate during October 2003 [37]) for manual therapy alone, 390 euros (£272) for a package of physiotherapy and 316 euros (£220) for routine GP care. Societal costs were similar for manual therapy in the two studies: £303 in our study and 447 euros (£312) in the Dutch study. Our costs were also comparable to those of Manca et al. [34] who reported average total 1-yr health-care costs of £152 and societal costs of £408 for usual physiotherapy.

A strength of this study was the detailed and systematic approach to data collection. Gathering of resource data used specifically designed questionnaires capturing data on health...
service resources and costs, costs to the participants and indirect costs (losses due to work absenteeism). Study participation was good—the intended sample size was achieved and high follow-up success reduced the potential for censorship bias. Nearly all participants were treated according to protocol [8]. However, the study was not powered to detect significant differences in cost (a common problem in health economic evaluations linked to clinical trials). Validation of the participants’ responses to the self-complete questionnaires was undertaken and the results indicated moderate to very strong reliability across the four resource measures according to the classification of Landis and Koch [38]. Due to the pragmatic nature of the design and interventions it was not possible to blind the physiotherapists and patients, which may have introduced a bias in therapists’ actions and patients’ responses to questions. A sensitivity analysis, based on imputed data, was carried out to test the robustness of the results: the findings showed similar effect and cost differences thus affirming our inferences.

Hoving et al. [39] pointed to a paucity of evidence from primary studies on neck pain, suggesting that more research is needed to allow systematic reviews to formulate stronger conclusions regarding conservative treatments for neck disorders. Bronfort et al. [40] highlighted the need for more CEAs in neck and low back pain research. Our study compared the cost-effectiveness of common treatments for neck disorders and included a broad range of participants with neck problems, namely those classified as having non-specific neck disorders who were referred to physiotherapy from GPs in primary care. There is evidence that outcomes for the type of interventions described in this study do not change greatly between 6 and 12 months follow-up [41, 42]. In our study, on average the majority of resource consumption occurred in the immediate post-randomization period (79% of total health-care costs occurred within the first 6 weeks). Therefore, follow-up periods of 6 months is likely to have captured the relevant costs and benefits. The relevant comparator could have been no physiotherapy since the GP could have given advice and exercise. Limited available evidence suggests that advice and exercise may be similarly effective whether administered by physiotherapists or GPs [43]. However, the advice and exercise package in this study was a comprehensive package, which was more substantial than simple advice and exercise given by the GP.

Although the study was not powered to detect pre-defined differences in costs, conceptually the differences in costs between interventions were small. On an individual patient basis, cost differences in the range of £5 to £80 over 6 months represent minimal outlays. However, by extrapolating these figures to the general population, where neck disorders represent a common problem [1, 2], this may represent an important difference in cost overall. Larger/refined studies would be needed to establish whether differences apply across specific patient subgroups e.g. by disease duration, age group, sex and pain severity.

In conclusion, the intervention of choice is likely to be A&E or MT, depending on the economic perspective and outcome measure. There was much variability in the data, which is reflected in the uncertainty in the results. However, it is unlikely that PSWD is a cost-effective intervention. Policy makers may wish to consider additional factors, such as patient preference, when planning provision of physiotherapy for people with non-specific neck disorders.

Acknowledgements

The team would like to acknowledge our funders, the Arthritis Research Campaign (arc) and the West Midlands R & D NHS (Executive). We would like to thank our collaborators: participating physiotherapists, therapy managers and physiotherapists involved in protocol development. We would also like to thank Richard Gray, Jane Daniels, Louise Hiller, the Primary Care Sciences Research Centre administrative staff, the members of the independent Data Monitoring and Ethics Committee and Birmingham Clinical Trials Unit for the randomization service. We thank Chan Vohora, the study nurse, for her assistance with the GP record review. Finally, we would like to thank the anonymous reviewers for their helpful comments.

Disclosure statement: The authors have declared no conflicts of interest.

References


