How long should outpatient pulmonary rehabilitation be? A randomised controlled trial of 4 weeks versus 7 weeks

L Sewell, S J Singh, J E A Williams, R Collier, M D L Morgan

Background: The evidence of benefit for pulmonary rehabilitation (PR) programmes is established. However, the optimal duration of a PR programme is not known. A randomised controlled trial was undertaken in patients with chronic obstructive pulmonary disease (COPD) to assess whether a 4 week PR programme was equivalent to our conventional 7 week PR programme at equivalent time points of 7 weeks and 6 months.

Methods: One hundred patients (56 men) with stable COPD of mean (SD) age 70 (8) years and forced expiratory volume in 1 second (FEV₁) 1.13 (0.50) litres were randomised to either a 7 week (n = 50) or 4 week (n = 50) supervised PR programme. Patients were assessed at baseline, at completion of the supervised PR programme, and 6 months later. Patients randomised to the 4 week group were also assessed at the 7 week time point. Outcome measures were the Incremental Shuttle Walk Test, Endurance Shuttle Walk Test (ESWT), Chronic Respiratory Questionnaire-Self Reported, and the Breathing Problems Questionnaire.

Results: Forty one patients in each group completed the PR programme. Patients made significant within group improvements after supervised rehabilitation. There were no statistically significant differences between the groups for any other measure at the 7 week or 6 month time points, except that patients in the 4 week group attained higher ESWT times (mean difference 124 seconds (95% CI 17.00 to 232.16), p=0.024) at the 7 week time point.

Conclusions: A shortened 4 week supervised PR programme is equivalent to a 7 week supervised PR programme at the comparable time points of 7 weeks and 6 months.
Breathing Problems Questionnaire (BPQ). The CRQ-SR is a reliable and valid measure of health status in patients with COPD. It is comprised of four domains: dyspnoea, fatigue, emotional impact of respiratory disease. A total score produced ranging from 0 to 30 with lower scores indicating better health status. This measure has previously been shown to be sensitive to change following a short course of pulmonary rehabilitation.

**Outcome measures**

The Incremental Shuttle Walking Test (ISWT) and the Endurance Shuttle Walk Test (ESWT) were primary outcome measures for this study. The ISWT was used to measure maximal exercise performance. All patients completed one practice walk. The ESWT measured submaximal exercise performance. This test is completed following the ISWT. Patients walked at a constant speed equating to 85% of predicted peak oxygen consumption (V̇O₂) as calculated from the ISWT. Total continuous walking times were recorded and patients also completed daily training walks at home. These were recorded on a home training diary. These home sessions were unsupervised; patients were advised to increase the duration of their walk rather than the intensity. This was monitored weekly in the supervised sessions. During the supervised strength training sessions, patients completed four separate exercises (bicep curls, sit to standing, pull-ups, and step-ups). Patients completed these exercises using hand weights and were instructed to complete three sets of 10 repetitions of each exercise. Perceived exertion scores were recorded and hand weights increased when these scores were equal to or below 13.

**Pulmonary rehabilitation (PR) programme**

The PR programme took place at Glenfield Hospital, University Hospitals of Leicester NHS Trust. The aim of the programme is to improve both physical performance and health status. Patients attended twice weekly with each session lasting for 2 hours; our institute offers a rolling programme. The session was divided into 1 hour of supervised exercise and 1 hour of education. All patients also completed a home training programme.

The education programme consisted of a rotation of 14 sessions of seminars and discussions covering the following topics: relaxation, disease education, dietary advice, benefits advice, energy conservation, medication advice, chest clearance, and breathing control techniques. Those patients randomised to the 4 week PR programme would miss a proportion of these talks. Therefore, all patients in the 4 week group were given a manual of comprehensive written information covering all the education topics.

During each week patients received 1 hour of supervised aerobic training and 1 hour of supervised strength training exercises. Supervised aerobic training sessions consisted of walking and cycling. Progression was achieved through weekly increases in duration. Patients were instructed to walk at a speed equal to 85% of predicted peak V̇O₂ as calculated from the ISWT. Total continuous walking times were recorded and patients also completed daily training walks at home. These were recorded on a home training diary. These home sessions were unsupervised; patients were advised to increase the duration of their walk rather than the intensity. This was monitored weekly in the supervised sessions. During the supervised strength training sessions, patients completed four separate exercises (bicep curls, sit to standing, pull-ups, and step-ups). Patients completed these exercises using hand weights and were instructed to complete three sets of 10 repetitions of each exercise. Perceived exertion scores were recorded and hand weights increased when these scores were equal to or below 13.

**Analysis of data**

A power calculation for an equivalence trial was completed; it was calculated that 40 patients would be required in each group based on a difference in the mean improvement on the ISWT of 20 metres between the two groups. Recruitment of 100 patients would therefore allow for a study dropout rate of 20%.

Statistical analysis was carried out using Statistical Package for Social Science (SPSS) software Version 13. Baseline characteristics and exercise data are presented as mean (SD). Mean and range values were identified for quality of life scores. The within group differences were compared using repeat measures analysis of variance, post hoc analysis identifying where significant differences occurred. Between group comparisons were compared using analysis of variance for repeat measures with weeks on rehabilitation being the dependent factor. Analysis between
groups was performed at equivalent time points—that is, baseline, 7 weeks, and 6 months. Further analysis was completed between groups at comparable time points, comparing the difference at the end of the supervised component of training. To take into account possible baseline differences in exercise performance, univariate analysis of variance was performed with baseline exercise performance as a covariant. Where analysis of variance identified a significant difference, post hoc tests were completed with correction for multiple measures. Mean difference and 95% confidence intervals (CI) are presented for comparisons within and between groups. A p value of <0.05 was considered statistically significant.

RESULTS
One hundred patients were recruited to the study, 50 of whom were randomised to the 4 week PR programme and 50 to the 7 week PR programme. Baseline characteristics of these patients are presented in Table 1. Eighteen patients withdrew from the study (nine from each group) before they completed the 7 week time point of the study; the reasons are detailed in Fig 1. This level of dropouts is similar to that experienced in our routine clinical service. No baseline differences in age, sex distribution, or pulmonary function were found between the two groups. There were also no differences for either group when study dropouts were compared with those who completed the study in terms of age, sex distribution, pulmonary function, and baseline ISWT score.

Of those patients who completed rehabilitation, a total of 71 patients were reassessed 6 months after completing the rehabilitation programme; 36 from the 4 week PR programme and 35 from the 7 week PR programme. The reasons for non-attendance at the 6 month follow up assessment were exacerbation of COPD (n = 5), died (n = 2), moved away from area (n = 1), patient did not attend but no reason given (n = 6).

Within group analysis identified that there was a significant difference in ISWT performance among the four time points. Post hoc analysis revealed that the difference did not reach significance between the 4 and 7 week measurements (mean difference 6.7 m (95% CI 22.8 to −9.5). All other comparisons were statistically significant. The improvement in the ISWT from baseline to 4 weeks was 56.9 m (95% CI 41.2 to 72.2). At 6 months the improvement from baseline was 34.2 m (95% CI 15.8 to 52.7; Fig 2A).

The pattern of response was similar for the ESWT. Analysis revealed that there was a significant difference in performance over the four time points. Post hoc analysis showed that there was a difference between all time points except between 4 weeks and 6 months. Mean improvement between baseline and 4 weeks was 222.9 seconds (95% CI 152.4 to 293.5). Unlike the ISWT, performance continued to improve after 4 weeks of rehabilitation with a mean increase from 4 to 7 weeks of 115.0 seconds (95% CI 41.2 to 72.2).

Table 1 Baseline demographic data, exercise assessments, and health status

<table>
<thead>
<tr>
<th></th>
<th>Four week group</th>
<th>Seven week group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 50)</td>
<td>(n = 50)</td>
</tr>
<tr>
<td>M:F</td>
<td>29:21</td>
<td>27:23</td>
</tr>
<tr>
<td>Age (years)</td>
<td>68.2 (7.7)</td>
<td>71.96 (8.3)</td>
</tr>
<tr>
<td>Current smokers</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>FEV₁ (litres)</td>
<td>1.16 (0.52)</td>
<td>1.11 (0.47)</td>
</tr>
<tr>
<td>FVC (litres)</td>
<td>2.45 (0.97)</td>
<td>2.17 (0.57)</td>
</tr>
<tr>
<td>FEV₁/FVC</td>
<td>0.50 (0.17)</td>
<td>0.51 (0.14)</td>
</tr>
<tr>
<td>MRC score</td>
<td>3.6 (2–9)</td>
<td>3.8 (2–9)</td>
</tr>
<tr>
<td>ISWT (metres)</td>
<td>196.0 (117.1)</td>
<td>166.3 (96.5)</td>
</tr>
<tr>
<td>End Borg breathlessness score</td>
<td>4.6 (2–9)</td>
<td>4.5 (1–9)</td>
</tr>
<tr>
<td>End perceived exertion score</td>
<td>14.4 (11–17)</td>
<td>13.8 (9–17)</td>
</tr>
<tr>
<td>ESWT (seconds)</td>
<td>183.8 (108.6)</td>
<td>165.5 (104.1)</td>
</tr>
<tr>
<td>CRQ-SR dyspnoea</td>
<td>2.6 (1.5)</td>
<td>2.8 (1.1)</td>
</tr>
<tr>
<td>CRQ-SR fatigue</td>
<td>3.5 (1.1)</td>
<td>3.5 (1.2)</td>
</tr>
<tr>
<td>CRQ-SR emotion</td>
<td>4.4 (1.1)</td>
<td>4.2 (1.2)</td>
</tr>
<tr>
<td>CRQ-SR mastery</td>
<td>4.4 (1.3)</td>
<td>4.3 (1.4)</td>
</tr>
<tr>
<td>BPQ</td>
<td>13.3 (4.9)</td>
<td>13.9 (4.6)</td>
</tr>
</tbody>
</table>

Values are mean (SD) or mean (range).

FEV₁, forced expiratory volume in 1 second; FVC, forced vital capacity; ISWT, Incremental Shuttle Walking Test; ESWT, Endurance Shuttle Walking Test; CRQ-SR, Chronic Respiratory Questionnaire-Self Reported; BPQ, Breathing Problems Questionnaire.
The mean increase from baseline to 7 weeks was 337.8 seconds (95% CI 245.9 to 429.8; fig 2B).

The CRQ-SR for the dyspnoea domain identified significant differences. Post hoc analysis identified improvements between baseline and 4 weeks (p<0.05); further modest improvements were achieved at 7 weeks for these three domains but they were not statistically significant. At 6 months there was no significant difference from baseline. This pattern was repeated for the domains of emotion and fatigue. All four domains achieved clinically important changes at the 4 week time point. Interestingly, statistical significance was delayed until the 7 week assessment time point for mastery. The mean values for the difference in mastery increased significantly higher at 6 months (mean 1.1, 95% CI −0.01 to 1.0) was double the mean value at 4 weeks (mean 0.5, 95% CI 0.7 to 1.5; fig 3).

Analysis of the BPQ results also identified significant differences across the time points. There were significant improvements between baseline and weeks 4 and 7, but no important changes were observed between weeks 4 and 7. The score declined over the subsequent 6 months and was not significantly higher than baseline.

**Seven week group**

There was a significant difference in performance on the ISWT over the three assessment times (baseline, 7 weeks, and 6 months). The mean increase between baseline and 7 weeks was 52.4 m (95% CI 41.9 to 62.9). The difference at 6 months remained statistically significant at 28.8 m (95% CI 13.4 to 44.2; fig 2A). Analysis of endurance performance yielded a similar pattern. Post hoc analysis identified a significant improvement from baseline to 7 weeks with a mean improvement of 216.6 seconds (95% CI 132.0 to 301.2) which remained significantly higher at 6 months (mean difference 121.2 seconds (95% CI 64.5 to 177.7); fig 2B). We secured clinical and statistically important changes in the CRQ-SR (all domains p<0.001; fig 3) and the BPQ from baseline to 7 weeks.

**Between group comparisons**

At the 7 week point only differences for the ESWT were statistically significant with at 4 week group attaining greater improvement. The mean difference between the 4 and 7 week group was 124.6 seconds (95% CI 17.0 to 232.2), p = 0.024. There were no other statistically or clinically significant differences for any other measure between the 4 and 7 week groups at the 7 week or 6 month time points.

**Comparison at the end of supervised rehabilitation**

Univariate analysis of variance was used to identify the changes observed at the end of a period of exercise supervision. The results were further analysed by including baseline performance as a covariant. Improvement of performance on the ISWT and the ESWT were compared between baseline and 4 weeks for the 4 week group and between baseline and 7 weeks for the 7 week group. The mean improvement in the ISWT in the 4 week group at 4 weeks was 57.5 m (95% CI 44.2 to 70.8) and 50.0 m (95% CI 38.9 to 62.1) for the 7 week group at 7 weeks. There was no statistically significant difference between these improvements. The results were not influenced by baseline performance of the ISWT. There was a statistically significant difference in performance on the ESWT in favour of the 4 week group. The mean difference was 116.9 seconds (95% CI 11.7 to 222.3), which takes into account any differences at baseline. The CRQ-SR data did not reveal any statistically significant differences between the two groups at this equivalent time point. There was a trend for the mastery

and fatigue score to be lower, but this did not reach statistical significance. The BRQ results were similar.

**DISCUSSION**

These results indicate that a 4 week supervised PR programme is equivalent to a 7 week supervised PR programme at comparable time points of 7 weeks and 6 months following completion of the programme. Importantly, the overall response to PR is similar to that documented in other trials. Interestingly, the 4 week group attained higher improvements in endurance walking times when compared with the 7 week group at the 7 week time point. There was also a trend for the 4 week group to attain greater improvements in all other outcome measures, although this did not reach statistical significance. A combination of hospital and home training appears to be as effective as supervised hospital based training. However, it is accepted that further study is needed to support this view.

At the 4 week time point the 4 week group largely matched the 7 week group with regard to improvements in both measures of exercise performance and health status. Patients in this trial could clearly not be blinded to the length of their supervised rehabilitation programme, but the assessor was blinded to the intervention. It could be that the patients in the 4 week group knew that they only had 4 weeks in which to gain from the supervised element of the PR programme. Importantly, because only the 4 week group were reassessed after 4 weeks of supervised training, we do not know if this response to PR occurred in the 7 week group at the 4 week time point. We believe that it may be that this is the general response in rehabilitation and not simply a phenomenon seen in patients who know they only have 4 weeks of supervised PR available to them.

Changes in health status were similar to the results we reported previously. This earlier study failed to attain clinically important improvements in CRQ scores at 4 weeks. In the current study the data acquired at the end of the supervised programme revealed statistically and clinically significant changes in all domains except mastery; by the 7 week time point dyspnoea improved again (p<0.05). Mastery improved by the 7 week assessment to a comparable level to the 7 week supervised group, making large improvements in the 3 weeks of unsupervised home activity. The mastery domain reflects the level of control and confidence that patients feel they have with regard to managing their respiratory disease. It may be that patients gained confidence from realising that their level of exercise performance improved during the home training phase of the study, away from any hospital support or intervention.

It is also important to note that, at 6 months, patients in both groups were able to maintain some of the improvements in exercise performance noted at the 7 week time point. However, improvements in exercise performance and health status were not wholly maintained 6 months after completion of PR. It is accepted that there is no control group within this study, but this pattern of deterioration has been noted in other trials. It has previously been accepted that the progression of the underlying respiratory disease and comorbidity are very important factors. These patients have a progressive disease and exacerbations may hinder compliance with the home training advice that patients receive upon discharge from the rehabilitation programme. Compared with other studies, the immediate outcome to rehabilitation is favourable despite the relatively short intervention in both groups. Ries et al secured a modest improvement in 6 minute walking test distance, less than the acknowledged minimal clinically important difference, again over a 6 week period. A recent review catalogued the immediate improvement in rehabilitation over varying
lengths of programmes; most programmes identified in this review failed to achieve the MCID for the 6 minute walking test. The randomised controlled trial reported by Griffiths et al10 included a 1 year follow up after rehabilitation and a control period. Data showed that patients were significantly better at 12 months than a control group, which would inevitably decline with time. The ISWT data revealed that patients had declined in performance to baseline values at 12 months. Our data show that patients have retained about half of the performance benefit at 6 months, suggesting that our population had declined in capacity in a similar way to the population studied by Griffiths et al.10

One of the methodological problems with this study was that all of the patients in the 4 week group knew that they were going to be reassessed at the 7 week time point. It could be argued that this knowledge of the impending reassessment may have artificially influenced compliance with their home training programme. It is also accepted that, because of the dropout rate, this study was underpowered at the 6 month time point. Larger studies may be needed to confirm the equivalence of a 4 week programme in the longer term.

It is also accepted that there were some non-responders to rehabilitation in both treatment groups. This is clearly important to recognise, but was not the primary focus of this study. Further large prospective trials are needed to investigate the length of rehabilitation programmes, and could be competency driven in contrast to completing previously predefined PR programme lengths.

In summary, this study has shown that a shortened 4 week supervised PR programme is capable of achieving similar results to a 7 week supervised PR programme at the comparable time point of 7 weeks and at the end of supervised rehabilitation. It is recognised that the provision of PR services in the UK and other developed countries is poor and is currently only provided to a small minority of people with disabling lung disease. A shorter supervised programme may facilitate a more effective use of resources and result in PR being offered to a greater number of patients.

ACKNOWLEDGEMENTS
The authors thank Dr Ruth Green for her invaluable statistical advice.